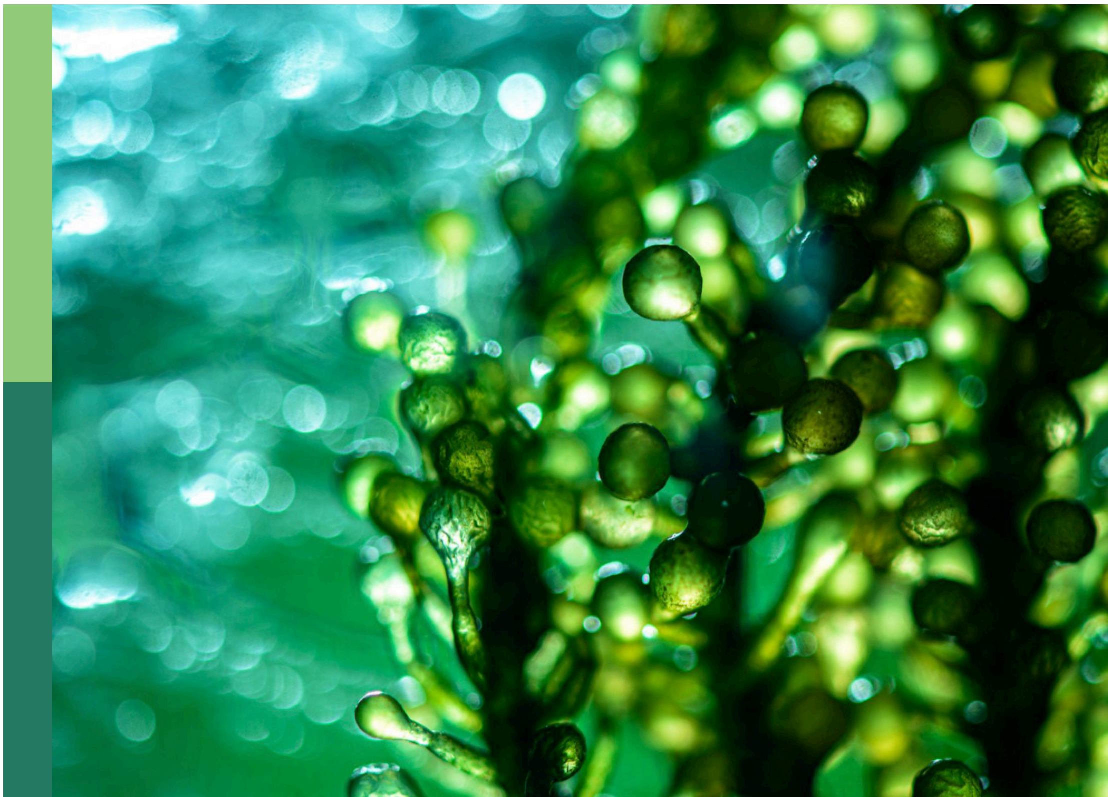
frontiers | Research Topics

Sustainametrics - envisioning a sustainable future with data science

Edited by
Shutaro Takeda, Alexander Ryota Keeley, Shunsuke Managi
and Thomas Gloria

Published in
Frontiers in Sustainability



ISSN 1664-8714
ISBN 978-2-83251-698-0
DOI 10.3389/978-2-83251-698-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Sustainametrics - envisioning a sustainable future with data science

Topic editors
Shutaro Takeda — Kyoto University, Japan
Alexander Ryota Keeley — Kyushu University, Japan
Shunsuke Managi — Kyushu University, Japan
Thomas Gloria — Independent researcher Industrial Ecology Consultants, United States

Citation
Takeda, S., Keeley, A. R., Managi, S., Gloria, T., eds. (2023). Sustainametrics - *envisioning a sustainable future with data sciencee*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-698-0

TG was employed by Industrial Ecology Consultants, a private consultancy firm of his own.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Table of contents

frontiers | Frontiers in Sustainability

TYPE Editorial
PUBLISHED 02 February 2023
DOI 10.3389/frsus.2023.1130622

EDITED AND REVIEWED BY
Maria Molinos-Senante,
Pontificia Universidad Católica de Chile, Chile

*CORRESPONDENCE
Shutaro Takeda
✉ takeda.shutarou.55r@kyoto-u.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 23 December 2022
ACCEPTED 25 January 2023
PUBLISHED 02 February 2023

CITATION
Takeda S, Keeley AR, Gloria T and Managi S (2023) Editorial: *Sustainametrics*—Envisioning a sustainable future with data science. *Front. Sustain.* 4:1130622. doi: 10.3389/frsus.2023.1130622

Editorial: *Sustainametrics*—Envisioning a sustainable future with data science

Shutaro Takeda[1]*, Alexander R. Keeley[1,2], Tom Gloria[3] and Shunsuke Managi[1,2]

[1]Urban Institute, Kyushu University, Fukuoka, Japan, [2]Department of Civil Engineering, Kyushu University, Fukuoka, Japan, [3]Industrial Ecology Consultants, Newton, MA, United States

KEYWORDS

sustainability, data science, artificial intelligence, life cycle assessment, circular economy, environment social and governance

Editorial on the Research Topic
Sustainametrics—Envisioning a sustainable future with data science

What is *sustainametrics*?

Sustainability study provides a roadmap to a better future. The global community is seeking a way to protect the people and the environment while supporting our economic prosperity, embodied by the UN Sustainable Development Goals (SDGs). Regrettably, however, progress toward the sustainable future has lagged through the pandemic; worse, there are some early signs that governments are reducing financing toward reaching the SDGs in the post-COVID era.

Today, researchers have access to an unprecedented amount of data in the field of sustainability, and data science may hold the key to changing this tide. We can urge immediate actions for the global community by objectively analyzing the environmental, social, and economic progress toward SDGs—to open a new way to a sustainable future.

It is time for us, researchers and practitioners, to reinvigorate global strides toward a sustainable future through coordinated efforts to utilize data in analyzing every aspect of sustainability. What is particularly needed today is the quantitative analysis of global development based on the empirical development of sustainability theory and observation. With this approach—which we named *sustainametrics*—academics may be able to present a viable path to a sustainable future (Figure 1).

The *sustainametrics* approach

Pioneers have long strived to pursue *sustainametrics* over the past generations, even before the advent of data science. We can trace the origin of *sustainametrics* to the *Essay on the Principle of Population* by Thomas Robert Malthus (1798), where Malthus argued that exponential population growth would eventually outpace linear agricultural production, leading to famine or war. This argument had become the underlying proposition of the famous report *The Limits to Growth* by the Club of Rome (1972), which showcased the groundbreaking application of the system dynamics in the field of sustainability.

10.3389/frsus.2023.1130622

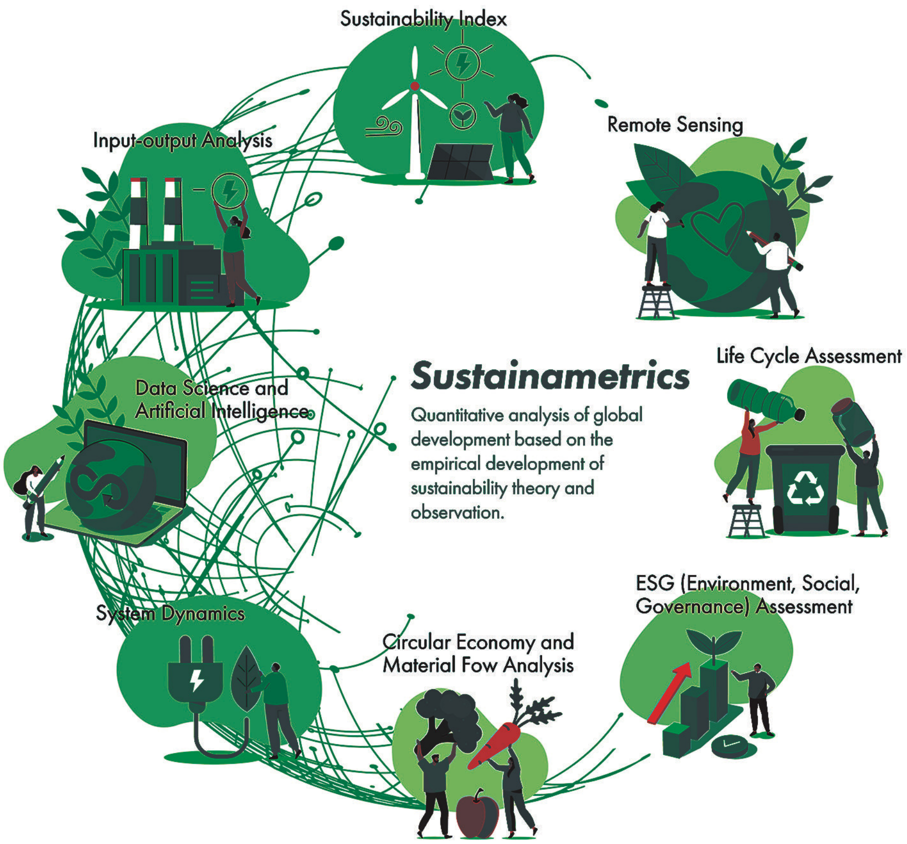


FIGURE 1
Sustainametrics-envisioning a sustainable future with data science.

Recent developments in data science have provided academics and practitioners the ability to analyze the progress of sustainable development more holistically and swiftly than these pioneers—a critical ability in a rapidly changing post-COVID landscape. Artificial intelligence (AI) is already proving useful in assessing the environmental impact from remote sensing data. Machine learning algorithms are being deployed to track poverty in real-time from alternative data including remittance. Further, existing quantitative methodologies, such as life cycle assessment (LCA), input-output analysis, material flow analysis, not to mention system dynamics, are also benefiting greatly from enhanced access to data supported by applied AI.

About this Research Topic

There are challenges to overcome toward establishing sustanametrics as an academic field. Some practitioners are still skeptical of the usefulness of data science in the actual world due to the incompatibility, uncertainty, and gaps in real-life data; others have suspicions about the capabilities of simple analytical models on understanding the inherently indeterminate real-life consequences. However, revealing new insights and gaining experience to bolster confidence is not enough. As Okui and Takeda conclude, catalyzing civic action through disclosure of what sustainametric professionals see and understand that other do not is essential to making real progress. Therefore, we see the need to bring together the wisdom on how data science and other quantitative methodologies can push the sustainability development forward now more than ever.

Our first collection on sustanametrics has gathered the 15 most insightful and inspiring contributions from like-minded researchers worldwide that is truly worthy of a start of a new academic field.

Keeley, Chapman et al. provides a closer look at the ESG metrics, which plays a crucial role as an enabler of investment strategies that consider ESG factors. The study confirms that ESG investments can be expected to provide stable and high returns especially over the long term. However, the study also identifies that there is significant divergence among the different ESG metrics in the elements assessed, and the weak correlation of ESG ratings of widely used ESG metrics.

Plugge proposes a new set of metrics that provides companies with a tool to measure and report on progress toward a circular economy. The proposed set of metrics is composed of quantitative risk and hazard metric combining chemical exposure and environmental health hazard within the supply chain. Similarly, Betts et al. propose a new set of metrics to measure the circularity and impact of reusable packages in supply chains, by connecting the existing research on circular economy metrics with reuse strategies in supply chains. The authors categorize these metrics as product-level or system-level based on the level of detail they incorporate and demonstrate their application with a case study from an omnichannel retail company.

Keeley, Li et al. calculates the cumulative ESG ownership based on the ESG scores of invested companies, the total market price of invested companies, and the investor history portfolio report. The study identifies the major players in the field, differences in the trend by type of investor and country and expands the study by investigating the relation between calculated ESG ownership score and investor's ESG commitment and ESG performance.

Aboginije et al. measure the life-cycle sustainability performance of a waste management system in South Africa and shows that the South African construction industry is hitherto to fully adopt and implement a sustainable waste management system for effective waste minimization, although the overall performance shows that the waste sectors are thriving and improving in their approach to waste management. Rinawati et al. focuses on LCA studies of hydrogen-based power generation and provide thorough review of the technological and methodological choices made in hydrogen-based power generation LCAs. The study points out that no studies that addressed social impact were identified through the systematic review and assert the importance of applying social LCA methodology to the research and development process for understanding the future impact of the hydrogen-based power generation more holistically. Zulfhazli et al. shed light on techno-economic aspect of hydrogen production from power generation through comparatively investigating the energy sources, feedstocks, and various methods of hydrogen production in detail.

Schubert et al. illustrates the complementary value of multidisciplinary inferential models in informing large predictive models by applying structural equation modeling to investigate the relationships that delineate the underlying mechanisms for energy consumption behaviors in the case of private transportation.

Kitsuki and Managi proposes a framework for weighting priority for the multidimensional domains of slum development from the viewpoint of residents and demonstrate this approach by accessing residents' needs for slum development focusing on India with employing a large-scale questionnaire data. The study sheds light on the importance of information on marginal utilities of each domain, as well as satisfaction scores for designing sustainable path for slum development.

Jin and Ialnazov also investigates the elements that are important for sustainable rural development, focusing on house-hold level solar energy projects in rural and remote areas. The study applies analytic hierarchy process and fuzzy comprehensive evaluation method to assess sustainability performance of the solar energy projects conducted in Jinzhai County, China.

Sakaguchi and Fujii investigate the impact of renewable energy on wholesale electricity prices. The result of the study implicates that compared with solar power, wind power has stronger merit order effect, which is the price reducing effect of renewable energy on wholesale electricity prices. Hao et al. sheds light on the solar PV locations' hazard risks at a national scale employing satellite data. The study investigates the risks stemming from landslides and floods for the existing solar PV power plants and finds that the shares of medium and large-scale solar PV power plants found in areas where landslides and floods are likely to occur are about 8.5 and 9.1% respectively in Japan.

Edwards et al. provides thorough review of satellite data application in the study and practice of sustainable energy system development and shows that satellite data are increasingly applied to a wide range of energy issues with varying information needs, from planning and operation of renewable energy projects, to tracking changing patterns in energy access and use, to monitoring environmental impacts and verifying the effectiveness of emissions reduction efforts. Similarly, Kazawa et al. emphasizes the effectiveness of using satellite data for assessing urban environment and urban design and understanding industrial clusters. The study examines whether nighttime light can be a proxy for building height and finds a high correlation between night light and building height.

Finally, Okui and Takeda revisit the measure of development and provide a critique of *sustainametrics*. Following Heidegger's and Arendt's threads of thought, the authors stress that any measures of development must be fundamentally grounded in disclosure through speech and action in the public realm; thus if the discipline of *sustainametrics* is to devise and propose measures for the sustainability of the world, then it must be conscious of the limit and prevent the reproduction of its thoughtless adoption by constantly exposing its outcomes to the scrutiny of political debate in the public realm.

Author contributions

ST conceptualized sustainametrics. AK, TG, and SM made a substantial, direct, and intellectual contribution to the concept. All authors contributed to the editorial and approved it for publication.

Conflict of interest

TG was employed by Industrial Ecology Consultants, a private consultancy firm of his own.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Sustainability

ORIGINAL RESEARCH
published: 30 November 2021
doi: 10.3389/frsus.2021.770045

The Impact of Variable Renewable Energy Penetration on Wholesale Electricity Prices in Japan Between FY 2016 and 2019

*Makishi Sakaguchi and Hidemichi Fujii**

Faculty of Economics, Kyushu University, Fukuoka, Japan

The merit order effect (MOE), which renewable energy sources can decrease wholesale electricity prices, plays an important role in establishing low-carbon societies. After the liberalization of the electricity market, the trade volume of the Japan Electric Power Exchange (JEPX) day-ahead spot market drastically increased between 2016 and 2019; however, price spikes still occur often. Ordinary least squares and quantile regression analyses were applied in this study to investigate how wind and solar photovoltaics (PV) energy generation affect the JEPX day-ahead spot price by time, price range, and area, and we concluded that the MOE of wind increased between 2016 and 2019 while that of PV decreased during this time. In regard to the high price ranges, although wind generation is not significant in terms of reducing price spikes, PV had this effect in 2016 and 2017 but not during the other years covered. The study area was divided into four regions, and each area followed trends that were different from those of the national analysis. Overall, the key finding of our study is that wind power has more potential to reduce electricity prices than PV.

Keywords: merit order effect, wholesale electricity market, renewable energy, Japan, electricity price

OPEN ACCESS

Edited by:
Shutaro Takeda,
Kyoto University, Japan

Reviewed by:
Dimiter Ialnazov,
Kyoto University, Japan
Edwin Samir Pinto Maquilon,
University of Zaragoza, Spain
Sumie Nakayama,
Kyoto University, Japan

***Correspondence:**
Hidemichi Fujii
hidemichifujii@econ.kyushu-u.ac.jp

Specialty section:
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

Received: *03 September 2021*
Accepted: *08 November 2021*
Published: *30 November 2021*

Citation:
Sakaguchi M and Fujii H (2021) The Impact of Variable Renewable Energy Penetration on Wholesale Electricity Prices in Japan Between FY 2016 and 2019. Front. Sustain. 2:770045. doi: 10.3389/frsus.2021.770045

INTRODUCTION

Man-made greenhouse gases (GHGs) from burning fossil fuels are causing climate crisis worldwide. The Intergovernmental Panel on Climate Change (IPCC, 2018) said that global net anthropogenic CO_2 emissions must decline by ~45% from 2010 levels by 2030 and reach net zero by ~2050 to limit global warming to 1.5 degrees Celsius. Since then, many countries have set a target of net zero emissions by 2050, as Japan did in 2020 (Cabinet of Japan, 2020). Japan has a major responsibility to cut emissions because it ranks fifth in terms of GHG emissions worldwide. Approximately 40% of Japan's GHG emissions in 2018 came from the electricity generation sector. Therefore, the 5th Japanese strategic energy plan identifies renewable energy sources as the main power supplies that will be used to achieve the net zero emissions goal (Ministry of Economy, 2018). The Japanese Ministry of the Environment (Ministry of the Environment, 2020) reported that variable renewable energy (VRE), such as wind and solar photovoltaics (PV), has greater future growth potential than other renewable energy sources. However, wind and PV contribute a small part of electricity generation (1 and 8%, respectively, in 2020). Since the implementation of the feed-in-tariff system in 2012, the use of PV power has drastically increased; however, wind power is used less often.

In terms of the electricity market, in 1995, the first institutional reform launched the wholesale electricity market, bringing competition to the power generation sector; this was followed by a partial liberalization of electricity retailing in 2000, and the full liberalization of retailing began in April 2016. In addition, the JEPX, which is the only wholesale power exchange in Japan, was established in 2003, and trading began in 2005. Most trading happens in the day-ahead spot market, although the JEPX comprises several markets, such as intraday and forward markets. The day-ahead spot market has two types of prices: system (national) prices and area prices. First, the supply and demand curves are set to determine the system price by considering only bids from all over Japan. The country is then divided into nine areas, and the price for each area is determined after taking into account the capacity of each interregional interconnection line. If the system can operate without exceeding the capacity of the interregional interconnection lines, the system price will equal the respective area price, but if the capacity of an interconnection line is exceeded, market divisions occur, and each area aggregates bids and sets its own price. These market divides often occur; for example, only 5% of the day-ahead spot prices of 2019 did not divide a market price. During the research period, two policies were applied to increase trade volume (see **Appendix**). The first was gross bidding, which was launched on 1st April 2017. This represents an initiative to buy and sell electricity on the JEPX, including that of the former general electric utility's own supply (in-house trading), and engage in exchange trading (net bidding), which was previously conducted mainly for surplus electricity. Second, implicit auctions were introduced on 1st October 2018; implicit auctions allocate power that passes through interregional interconnection lines to inexpensive power sources through the spot market, and they have enabled the efficient use of power sources across wide areas. As a result of these policies, the trade volume on the JEPX day-ahead spot market increased, and in 2019, 30–40% of all Japan's generated electricity was sold through JEPX transactions. Additionally, price spikes, namely, temporary and sudden increases in electricity prices mainly due to demand increases or supply shortages, became a problem, and this phenomenon can represent a major risk for market customers.

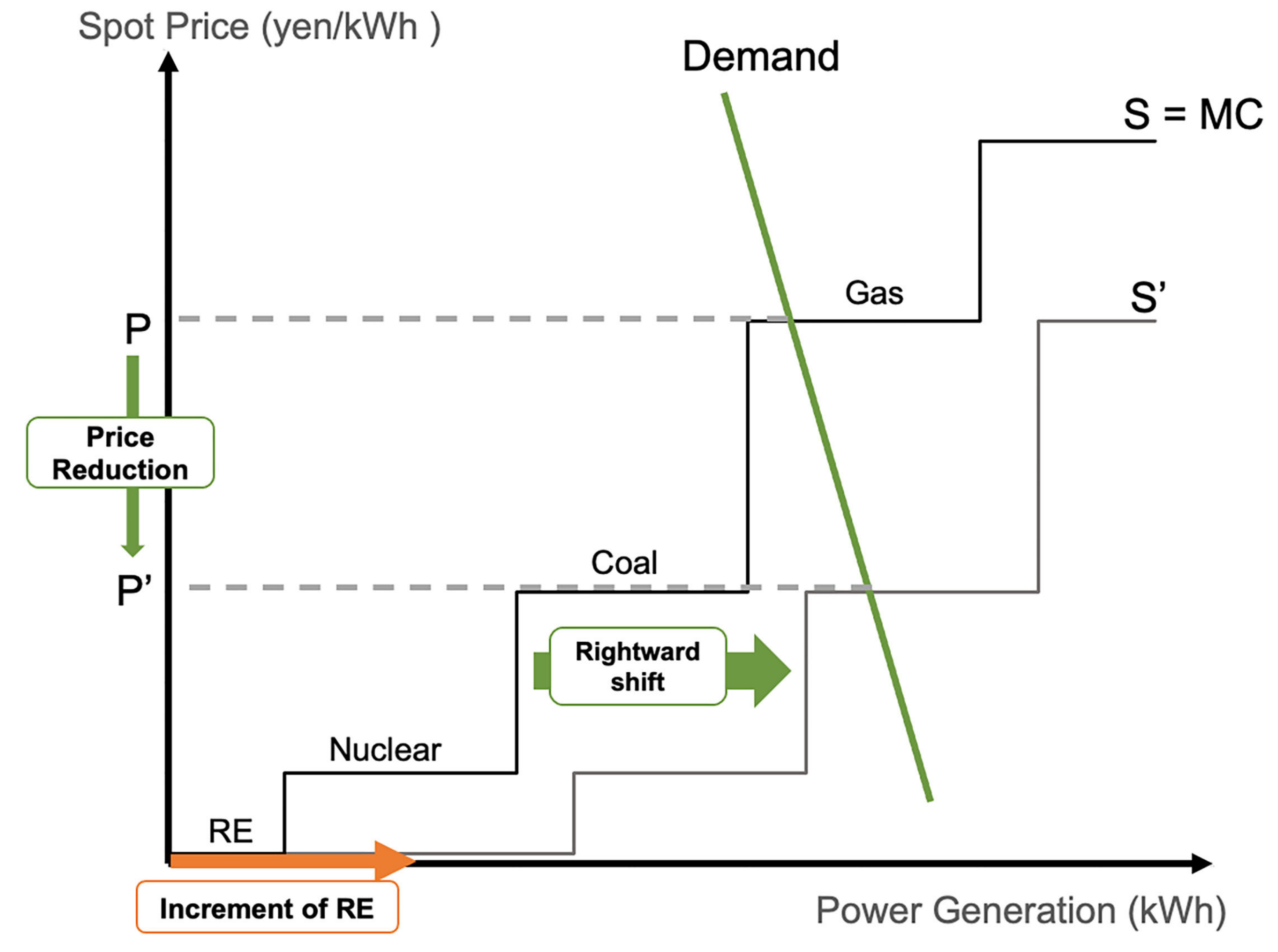


FIGURE 1 | Conceptional graph of the merit order effect (MOE). (Source: created by authors).

In the context of the effect of renewable energy penetration, the merit order effect (MOE) has been researched around the world, as shown in **Figure 1**. According to Marshall (1890), in a perfectly competitive market, it is socially optimal when marginal costs create supply curves, and prices are equal to the points where supply curves and demand curves intersect. Incidentally, the marginal cost of renewable energy is theoretically zero because, unlike thermal plants, it does not require fuel. Therefore, the more renewable electric energy is generated, the lower the electricity price becomes because the supply curve (or marginal cost of each power plant) shifts to the right, which lowers the equilibrium price. The point is that the penetration of renewable energies has a price reduction effect on wholesale electricity market, and this effect is called MOE.

BACKGROUND AND LITERATURE REVIEW

To the best of the authors' knowledge, there are few studies about the MOE in Japan; Maekawa et al. (2018) found that the MOE exists in Japan using ordinary least squares regression analysis. However, their study has some issues, such as the use of climatic data as a proxy variable, and the analysis was conducted when the trade volume of the JEPX was still below 10% (30–40% in 2019). Yoshihara and Ohashi (2017) found that based on the 2030 scenario proposed by the government, further installations of renewable energy sources lower kWh prices, reduce fuel costs, and reduce the volume of CO_2 emissions.

Zipp (2017) found time-series changes in the MOE of other countries; for example, the MOE of wind power in Germany increased between 2011 and 2013. According to Keeley et al. (2021), electricity from wind and solar sources reduced the German/Austrian spot market price by 9.64 €/MWh on average between 2010 and 2017; these authors analyzed this phenomenon using generalized least squares regression and a machine learning approach. Csereklyei et al. (2019) analyzed the MOE of wind and PV in Australia and found that with a 1 GW increase in 30-min dispatched capacity, wind has a wholesale price reduction effect of 11 AUD/kWh while solar has a wholesale price reduction effect of 14 AUD/kWh. In addition, they measured the MOE of each of the four subdivided areas of Australia, namely, NSW, QLD, VIC, and SA, and found that there is significant variation across these areas. In terms of different price ranges, using quantile regression analysis, Hagfors et al. (2016) found that the MOE of PV is higher at higher price ranges in Germany, reducing the effect of extreme price spikes, while the MOE of wind is greater at lower price points. Maciejowska (2020) used quantile regression analysis to examine German electricity market prices and found that wind has a greater price-reducing impact on the lower tail of the price distribution while solar has a greater price-reducing impact on the higher tail of the price distribution. As these studies show, the MOE was investigated not only in terms of its overall effect but also at the area, year and price levels.

In Japan, the use of solar power has dramatically increased due to the feed-in-tariff policy while the utilization of wind power has been low. In addition, capacity factors of VRE in japan are relatively low compared to other countries, with the average capacity factor in Japan being 23% for wind and 15.8% for PV (Renewable Energy Institute, 2018, 2019). However, wind and PV have different characteristics in terms of generating electricity; to be more specific, PV only works during the daytime when sunlight reaches the ground, but wind works regardless of the time of day. They also have seasonal differences; wind power generation increases during the winter and decreases during the summer, while solar power generation increases during the spring and summer and decreases during the winter. Furthermore, regional differences in the implementation of VRE influence the price in each area due to the area price system.

Thus, we choose three research objectives that have not been investigated in Japan thus far. (1): We investigate how the MOE has changed overtime, aiming to identify the changes that have occurred each year as the amount of electricity generated and the trade volume of the JEPX have changed. (2): We examine how the MOE varies across price ranges, analyzing whether VRE penetration reduces price spikes. (3): We consider how the MOE differs across the four areas of Japan, identifying the differences between these areas (Japan is classified into four areas in this study based on price correlations listed on **Appendix**), as each area tends to have its own spot price and power supply composition (**Figure 2**).

METHODOLOGY AND DATA

In this study, two regression models are employed. First, an ordinary least squares regression analysis is performed on the mean value of the data set to analyze the impact of the examined factor on overall energy prices. Second, quantile regression analysis is used to analyze the effect within each price range by performing a regression analysis for each quantile. Using these two methods, we aim to analyze how wind and PV affect the day-ahead spot price of the JEPX. The first step is to analyze national data to identify the changes that occur over time, and then differences in the impact across different areas are captured using data subdivided according to the nine identified areas; the Okinawa area is excluded from this analysis, as this area is not included in the JEPX market. The overall research framework is shown in **Figure 3**.

Models

We utilize the following models in both our national analysis and our area analysis.

Ordinary Least Square Regression Model (OLS Regression)

The independent variable is the JEPX day-ahead spot price. Variables for wind and PV electricity generation and control variables are included. Time autocorrelation effects must be reduced due to the use of time-series data. Therefore, we include the lagged price for the same hour of the day and for the previous seven days and the average day-ahead spot price across the 24 h of the previous day to eliminate autocorrelation in the data set. Moreover, we add a spot price volatility variable for historic price instability following Paraschiv et al. (2014) and Sirin and Yilmaz

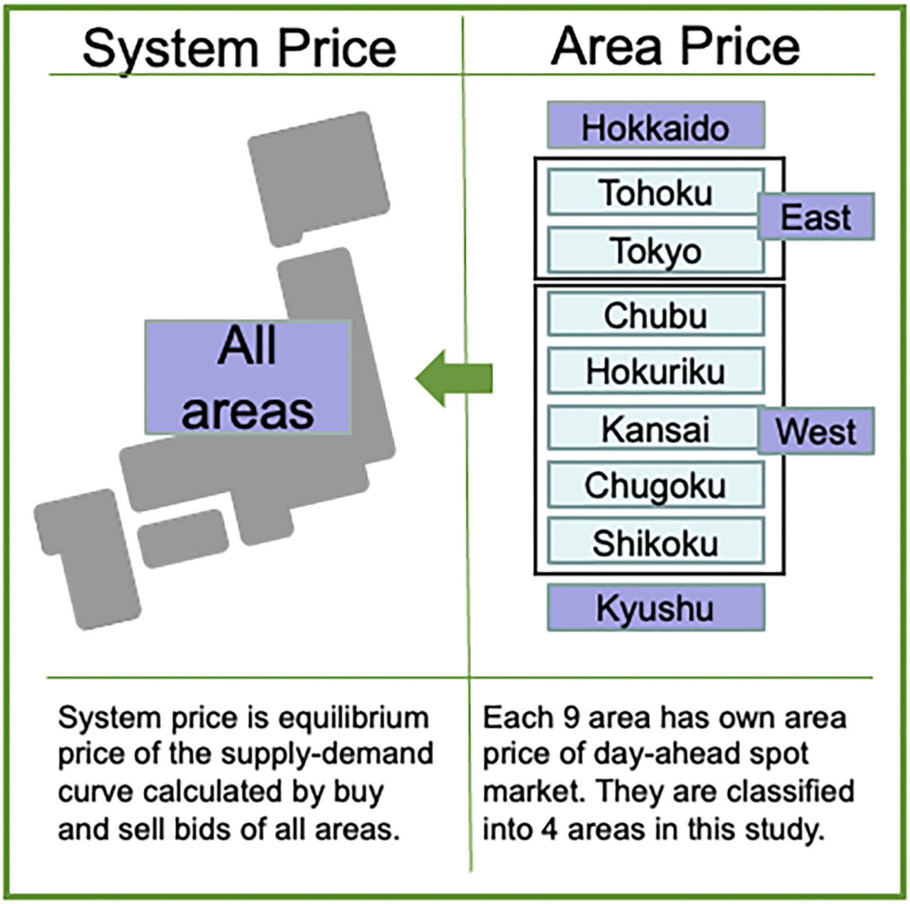


FIGURE 2 | Classification of the four areas. (Source: created by authors).

(2020), which is the standard deviation of the day-ahead spot prices of the same hour of the previous five days.

According to the above procedures, a linear regression model is estimated as shown in Equation (1). In this model, $Price_t$ is the JEPX day-ahead spot price at time t (one h).

$$\begin{aligned} Price_t = \alpha_t + \beta_1 Wind_t + \beta_2 PV_t + \beta_3 Demand_t \\ + \beta_4 Price_{t-24} + \beta_5 Price_{t-168} + \beta_6 AverageLag_t + \beta_7 Volatility_t \\ + \beta_8 Oil_t + \beta_9 Summer_t + \beta_{10} Winter_t + \beta_{11} Daytime_t \\ + \beta_{12} Holiday_t + \beta_{13} Policy_t + \gamma_t \end{aligned} \quad (1)$$

Quantile Regression Model

Using the same data as used for the OLS regression model, we estimate the following quantile regression model.

$$\begin{aligned} Q_q\left(Price_t\right) = \alpha_t^q + \beta_1^q Wind_t + \beta_2^q PV_t + \beta_3^q Demand_t \\ + \beta_4^q Price_{t-24} + \beta_5^q Price_{t-168} + \beta_6^q AverageLag_t + \beta_7^q Volatility_t \\ + \beta_8^q Oil_t + \beta_9^q Summer_t + \beta_{10}^q Winter_t + \beta_{11}^q Daytime_t \\ + \beta_{12}^q Holiday_t + \beta_{13}^q Policy_t + \gamma_t \end{aligned} \quad (2)$$

where $q \in (0, 1)$ represents the 5, 10, 25..., and 95% quantiles.

Data

National Analysis

The system price data cover the period from 1st April 2016 to 31st March 2020 according to the Japanese fiscal year, which runs from April to March of the following year. Each JEPX day-ahead spot price is determined every 30 min; however, in this study, they are converted to hourly prices. The prices are weighted averages of the JEPX day-ahead spot prices for each time period according to trade volume. The amount of electricity generation is calculated as the sum of the generation of each of the nine areas per h, which is reported by the respective transmission system operators. All the data used in this study are listed in **Table 1**.

Area Analysis

The area price data cover the period from 8th April 2016 to 31st March 2020 due to data limitations. Although we use weighted average prices in our national analysis, a simple average of each pair of 30-min price periods is employed in our area analysis because the JEPX discloses trade volume only at the national level. Demand, wind and PV variables for each area are used, and these values are published by each transmission system operator.

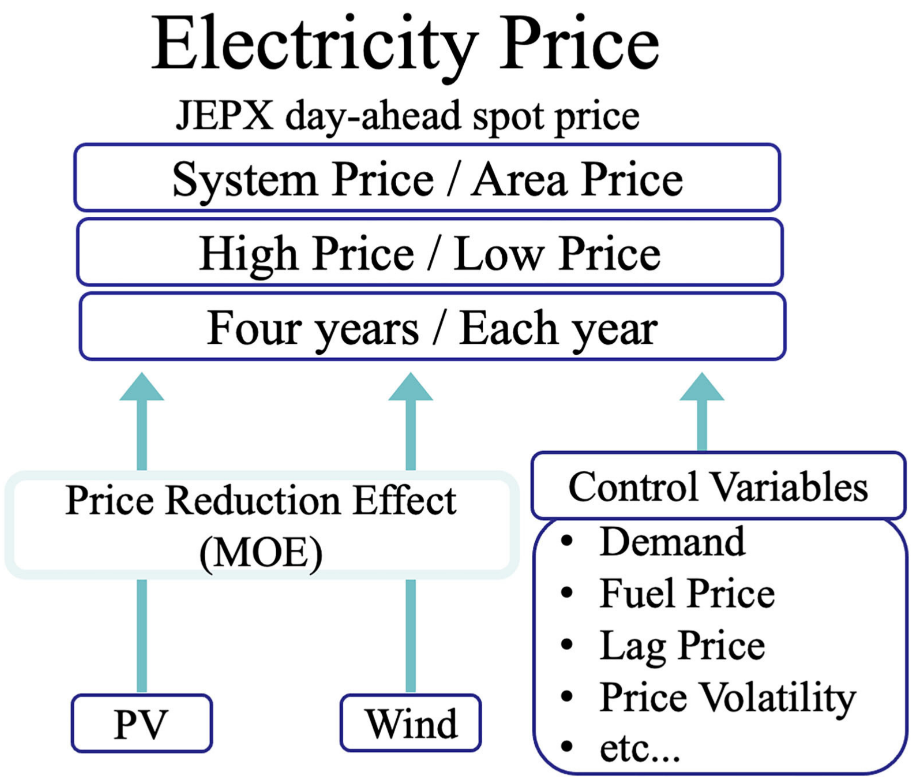


FIGURE 3 | Research framework of this study. (Source: created by authors).

Robustness Check

As a unit root may be present due to our use of time-series data, we perform an augmented Dicky-fuller (ADF) test to determine whether the variables are stationary according to Dickey and Fuller (1979). We reject the null hypothesis of a unit root at the significance level of 1%, which means that the JEPX spot price data are stationary. Then, we apply Durbin-Watson statistics and find that our data exhibit a positive autocorrelation. It is reasonable to consider the use of lagged data for our OLS and quantile regression models to eliminate the effect of this autocorrelation; thus, these data are added to equations (1) and (2).

RESULTS AND DISCUSSION

National Analysis

National Analysis (OLS Regression Analysis)

First, an MOE was indicated by negative coefficient values within the regression analysis results. The larger a negative coefficient was, the larger the corresponding MOE was. First, five models with different combinations of control variables were calculated (**Table 2**), and it was revealed that each coefficient was significantly negative, i.e., the MOEs of wind and PV were significant during the years covered.

In addition, the MOE of wind was a few times larger than that of PV in all five models. Next, since the adjusted coefficient of determination of model 5 was the highest, we split it into four fiscal years and examined the time series changes. Different chronological variations in the MOEs of wind and PV were found. The MOE of PV decreased from 2016 to 2019; on the other hand, the MOE of wind increased each year. In addition, the adjusted coefficient of determination decreased over the years examined, which implies that the impacts of other factors that determine the JEPX prices became stronger. The Japanese electricity market is still an oligopoly market because the former general electric utilities used to monopolize each area market before the liberalization of retail electricity market started in March 2000. This can be a factor to determine the JEPX spot price due to imperfect competitive market. Considering the effect of the policy change dummies: gross biding and implicit auction, the coefficient of implicit auction is positive while the one of gross bidding is negative. The reason why the implicit auction helps the electricity price to decrease would be the result of market integration.

National Analysis (Quantile Regression Analysis)

Figure 4 shows the results of a quantile regression used to capture the MOEs of each price range. The MOEs of wind were not significant in most quantiles, but the number of significant quantiles increased over time. The MOEs of PV, on the other hand, were significant for all the quantiles over the four years examined. Notably, the MOEs of the 95th quantile decreased from 2016 to 2019, demonstrating the price-decreasing effect of price spikes. Additionally, there were sharp drops between the 90th quantile and the 95th quantile in 2016 and 2017. These results indicated that the price-reduction effect of PV electricity

TABLE 1 | Data description.

Variable	Unit	Description	Source
Price	yen/kWh	System Price: hourly weighted average of the JEPX day-ahead spot price by trade volume. Area price: hourly simple average of the JEPX day-ahead spot price.*	Japan Electric Power Exchange
Demand	GWh	Total hourly electricity generation	Website of nine electricity companies
Wind	GWh	Hourly wind power electricity generation	
PV	GWh	Hourly solar PV electricity generation	
Lagged spot price	yen/kWh	Day-ahead spot price of the same h of the previous day	
One-week lagged spot price	yen/kWh	Day-ahead spot price of the same h seven days before	
Average lagged spot price	yen/kWh	Average day-ahead spot price across all 24 h of the previous day	
Spot price volatility	yen/kWh	Standard deviation of the day-ahead spot prices of the same h of each of the last five days	
Fossil fuel price	1,000 yen/KL	Monthly CIF (cost insurance and freight) price of crude oil	Ministry of Finance, Trade Statistics of Japan
Summer	Dummy	1 for months between July and September; otherwise, zero	
Winter	Dummy	1 for months between December and February; otherwise, zero	
Gross bidding	Dummy	1 for days between 2017/4/1 and 2020/3/31; otherwise, zero	
Implicit auction	Dummy	1 for days between 2018/10/1 and 2020/3/31; otherwise, zero	
Daytime	Dummy	1 for hours between 8AM and 10PM; otherwise, zero	
Holiday	Dummy	1 for Saturday, Sunday and Japanese national holidays; otherwise, zero	

**Area price in East and West region is estimated using hourly weighted average of the JEPX day-ahead spot price by each generation area after a simple 30-min average of each price. Nine electricity companies website are 1. Chubu electric power grid, demand data; 2. Chugoku electric power transmission and distribution company, demand data; 3. Hokkaido electric power network, demand data; 4. Hokuriku electric power transmission and distribution company, demand data; 5. Kansai transmission and distribution, demand data; 6. Kyushu electric power transmission and distribution, demand data; 7. Shikoku electric power transmission and distribution company, demand data; 8. TEPCO power grid, demand data; 9. Tohoku electric power network, demand data.*

generation is particularly strong during price spikes. However, the trade volume on the JEPX was relatively quite small during 2016 and 2017, and it is possible that the price spike reduction effect weakened as the transaction volume on the JEPX increased.

Discussion National Analysis

Certainly, the decline in the MOE of PV over time could be due to the corresponding increase in PV energy generation: solar PV accounted for 4.5% of the total energy generation of Japan in 2016, and it increased to 7.4% in 2019; this is because the impact of a single unit of PV generation decreases as PV installations increase. Since an overall upward trend in the MOE of wind was found, especially in 2018 and 2019, which had relatively significant quantiles, we considered those two years separately for daytime and night-time hours (**Figure 5**); we found that the magnitude of the MOE of wind did not change much during night-time hours, but it was dramatically larger during daytime hours. However, we were not able to infer the reason for this from our study; thus, we would like to leave this as an issue for future research.

Area Analysis

Area Analysis (OLS Regression)

Second, the study area was broken down into four regions, and the same calculation used in the national analysis was run. The results of this OLS regression are given in **Table 3**. Regarding wind, MOEs were not observed for the areas of West and Kyushu in 2016 and 2017, but they were significant for the other areas. Additionally, the MOEs of PV were significant in all the areas and years. In contrast to the national analysis, consistent time-series variation was not found in the area analysis. That is, the changes observed in the national analysis were not caused by the characteristics of a certain area; rather, they represent an overall trend. Notably, the wind and PV MOEs of Hokkaido were much higher than those of the other areas.

Area Analysis (Quantile Regression)

Figure 6 shows the results of the quantile regression analyses for each of the four areas. The MOE of wind became larger for the highest quantiles only in Hokkaido which means wind power can reduce price spikes. Similar to the national analysis, considering the overall trends of the PV MOEs, the lower the price range is, the larger the MOE is. Looking at the lowest quantile, a sudden increase for PV in Kyushu is seen. This might be because the JEPX Kyushu area price often set 0.01 yen/kWh which is the lowest price determined, and further PV curtailment also often occurs then. The fact that Kyushu is the area where the penetration of PV generation is the most advanced is the most conducive to this phenomenon.

Discussion Area Analysis

It should be noted that the area analysis did not take into account the effect of electricity passing through interconnection lines, which are used to trade electricity with other areas. For instance, Kyushu has only one interconnection line with another area, and Kyushu exports electricity throughout the year because it has the highest percentage of PV generation. That is, in the context of

TABLE 2 | The results of the OLS regression analysis of system prices.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5			
						2016	2017	2018	2019
Wind	−0.464***	−0.314***	−0.383***	−0.316***	−0.345***	−0.061	−0.194***	−0.325***	−0.322***
PV	−0.113***	−0.127***	−0.128***	−0.074***	−0.085***	−0.106***	−0.087***	−0.085***	−0.077***
Demand	0.147***	0.174***	0.168***	0.070***	0.092***	0.12***	0.11***	0.099***	0.078***
One-day Lagged Price				0.383***	0.352***	0.133***	0.253***	0.409***	0.497***
One-week Lagged Price				0.158***	0.121***	0.147***	0.171***	0.0287***	0.0574***
Average Lag Price				0.0393***	−0.0118*	−0.179***	−0.0712***	−0.0737***	0.0154
Price Volatility				0.273***	0.337***	0.475***	0.524***	0.280***	0.175***
Oil			5.21e-09***		2.59e-09***	0.0521***	−0.0241***	0.108***	0.0733***
Summer		−1.066***	−1.098***		−0.827***	−1.052***	−1.022***	−0.843***	−0.459***
Winter		−1.340***	−1.565***		−0.943***	−0.574***	−0.303***	−0.413***	−1.257***
Daytime		−0.312***	−0.179***		−0.354***	−0.314***	−0.707***	−0.459***	−0.276***
Holiday		0.634***	0.572***		−0.0482**	−0.0868**	−0.0637	0.142***	0.0308
Implicit Auction		−1.111***	−1.302***		−0.541***			−0.674***	
Gross Bidding		1.547***	0.864***		0.472***				
Constant	−4.994***	−7.887***	−9.906***	−2.996***	−5.285***	−5.525***	−3.652***	−7.982***	−6.526***
Observations	35,064	35,064	35,064	35,064	35,064	8,760	8,760	8,760	8,784
R-squared	0.538	0.598	0.619	0.736	0.758	0.807	0.813	0.723	0.706
Adj R-squared	0.538	0.598	0.619	0.736	0.758	0.807	0.812	0.723	0.706

****, **, * indicate significance at the 1, 5, and 10% levels, respectively.*

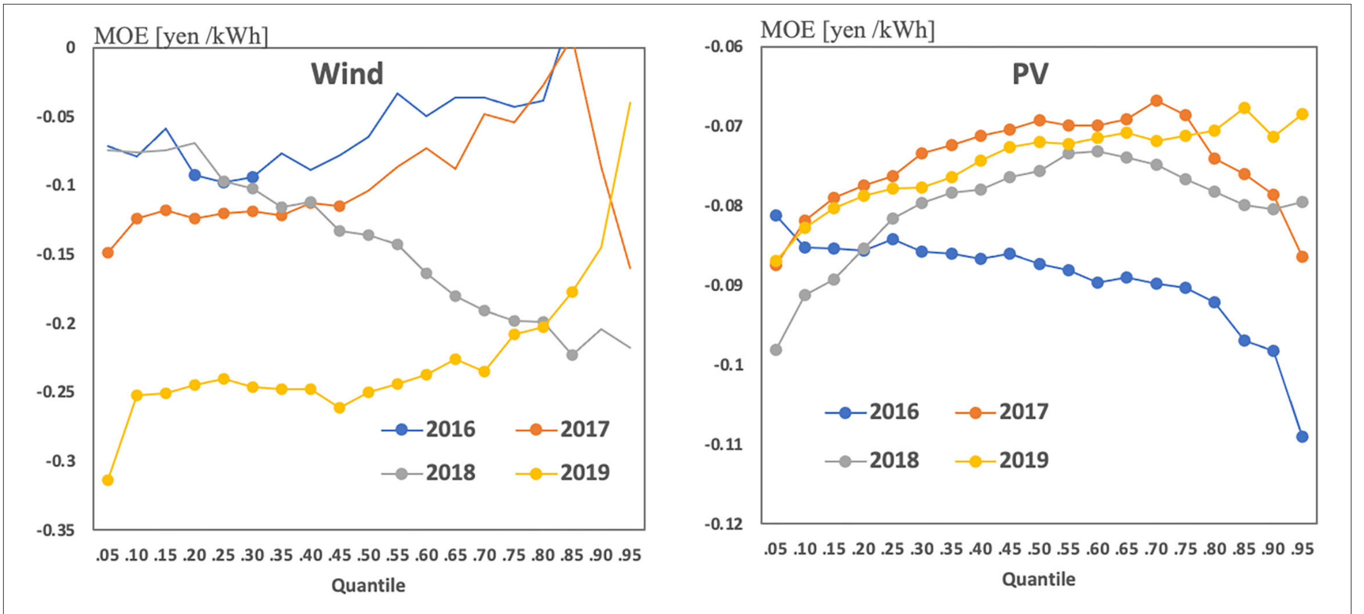


FIGURE 4 | The results of the quantile regression analysis of system prices. (Source: Created by author). The dots in the graph indicate significance at the 1% level.

this study, these lines are likely to cause the actual consumption of an area to be unequal to the electricity that is generated in the area, so errors exist due to this phenomenon. As with the results of the analyses of the four areas, the MOE trends of the various areas were not very consistent. There may be other factors in each area that determine the JEPX spot price. Additionally, the electricity mix, average prices, and the percentage of the market in each area within the high price range vary greatly, as shown in **Table 4**. This study revealed that a certain amount of VRE electricity generation is necessary to receive MOE benefits. Since Hokkaido has the largest share of electricity generation from wind in the four areas, the MOE for wind power was significant in all quantiles analyzed. On the other hand, the MOE of wind power was not significant in some quantiles in West and Kyushu

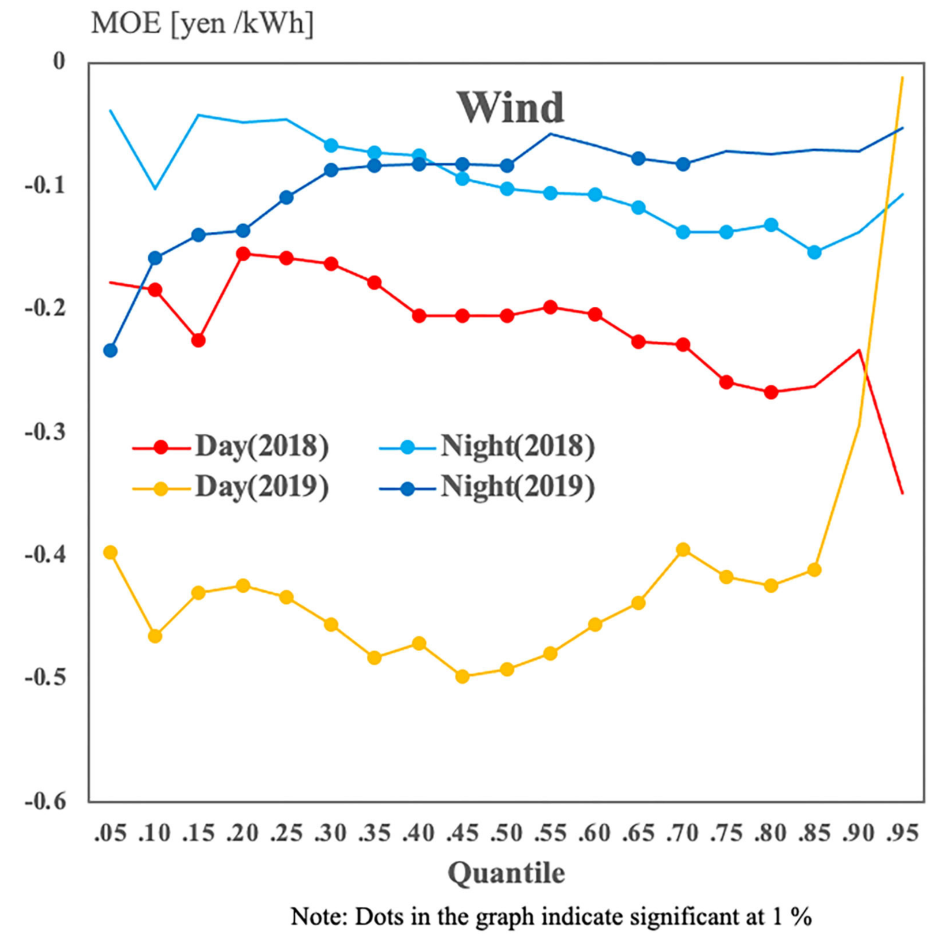


FIGURE 5 | MOE differences between daytime and night-time hours by price range. (Source: created by authors). The dots in the graph indicate significance at the 1% level.

TABLE 3 | Wind and PV MOE changes by year and area.

		All years	2016	2017	2018	2019
Wind	Hokkaido	−4.825***	−4.269***	−3.109***	−3.598***	−7.019***
	East	−0.760***	−0.529***	−0.324***	−0.699***	−0.677***
	West	−0.337***	−0.00623	0.0549	−0.586***	−0.672***
	Kyushu	0.270*	2.022***	1.558***	−1.707***	−0.651**
PV	Hokkaido	−1.846***	−1.810***	−1.231***	−2.338***	−1.710***
	East	−0.212***	−0.296***	−0.139***	−0.220***	−0.225***
	West	−0.211***	−0.238***	−0.260***	−0.187***	−0.201***
	Kyushu	−0.498***	−0.414***	−0.537***	−0.433***	−0.589***

****, **, * indicate significance at the 1, 5, and 10% levels, respectively.*

regions. Less electricity is generated by wind power in these regions than in other regions.

Comparison With Previous Research

In terms of comparing our overall MOE to those of other studies, although some previous research estimated the MOE by multiplying a calculated coefficient by a load-weighted average to quantify the actual amount of decrement due to RE penetration, we compared how much the price of electricity was reduced for each 1 GWh of additional VRE (wind and PV) generation per hour. The total number of VRE installations differs greatly by country. The results of a comparison of our MOE to those of other studies is given in **Table 5**, considering currency

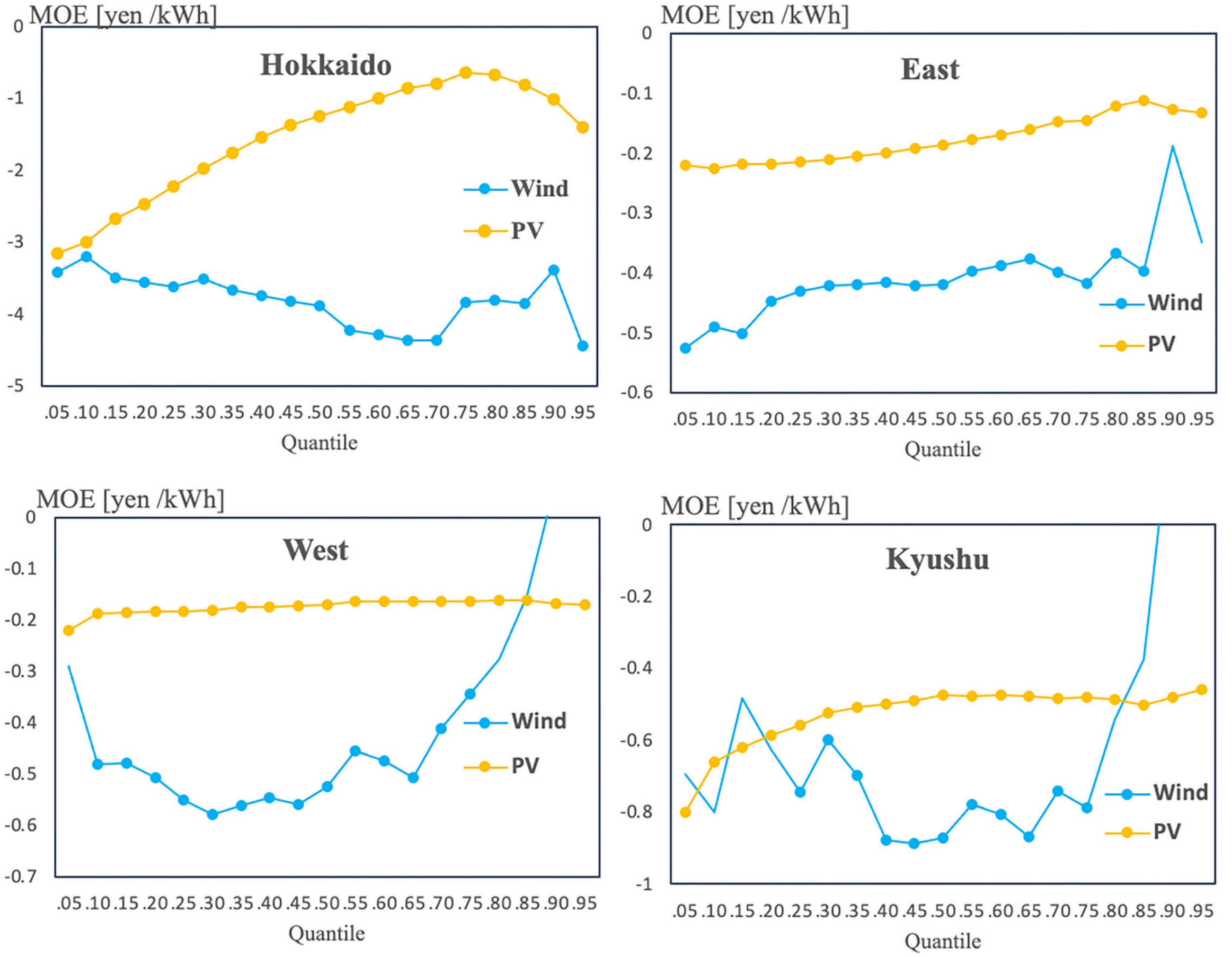


FIGURE 6 | The results of the quantile regression analyses of the price of each area. (Source: created by authors). The dots in the graph indicate significance at the 1% level.

TABLE 4 | Descriptive statistics of the electricity prices and generation sources in each area during 2019.

	Mean price (yen/kWh)	Ratio of price above 15 yen/kWh	Share of wind generation	Share of PV generation
Hokkaido	10.7	9.7%	3.7%	6.5%
East	9.1	2.9%	1.1%	6.5%
West	7.2	0.9%	0.5%	7.3%
Kyushu	6.8	0.9%	0.8%	12.4%

exchange rates into euros. The MOE of VRE generation in Japan is quite similar to that in other countries. However, Japan is unique in that wind power has a much larger MOE than PV, while both wind and PV have almost the same magnitude effect in Germany. The reason could be the differences of both VRE penetration percentage of the total generation. According to IEA (2021), electricity generation from wind and solar power in 2019 is about 20.7 and 7.6% in Germany while 0.7 and 6.6% in Japan. An examination of the MOE in the context of price spikes shows another difference. Although Hagfors et al. (2016) revealed that PV reduces the chance of extreme spikes, this effect weakened over time in Japan as trade volumes increased. Since this is the first study to the authors' knowledge to investigate the MOE of VRE in Japan considering price ranges, more research in this field is needed.

TABLE 5 | The MOE for each GWh of additional VRE generation.

Research	Country	Period	MOE for Wind	MOE for PV
Cludius et al. (2014)	Germany	2008–2012	0.94 to 2.27 €/MWh	0.84 to 1.14 €/MWh
Clò et al. (2015)	Italy	2005–2013	4.2 €/MWh	2.3 €/MWh
Zipp (2017)	Germany	2011–2013	1.08 to 1.54 €/MWh	1.03 to 1.45 €/MWh
Quint and Dahlke (2019)	US	2008–2016	1.07 to 2.61 €/MWh	
Keeley et al. (2021)	Germany	2010–2017	0.88 €/MWh	0.93 €/MWh
This Study	Japan	FY 2016–FY 2019	2.77 €/MWh	0.68 €/MWh

Exchanged to euros from other currencies by using the average exchange rate during the research period.

CONCLUSIONS AND POLICY IMPLICATIONS

The results of this study confirmed the MOE of both wind and PV electricity generation through the use of actual generation data, showing that the overall MOE of wind is greater than that of PV and that the former has been increasing over time while the latter has not. Japan can benefit from a greater price reduction by prioritizing the introduction of wind power. It was also evident that MOEs vary significantly across price ranges. In addition, price spike suppression was observed in 2016 and 2017 in the case of PV. Other factors may have started to affect the MOE during the price spike, as JEPX traded less in this period and solar PV generated less electricity than in 2019. For wind, the MOE was not significant at high price ranges in all year covered. It is also interesting that the MOE of wind is a few times higher than the one of PV, unlike previous research for other countries. In terms of price spikes, JEPX has experienced the long- continued high market price from December 2020 to January 2021 with unprecedented level. During the term, it is not appropriate to use these high price data unconditionally with the data we used in this study because a lot of externalities could be concerned. However, investigating the factors that caused these extreme high prices and the MOE of VRE during the period is important to the improvement of JEPX. And we leave this to future research.

In Japan, "Act on Promoting the Utilization of Sea Areas for the Development of Marine Renewable Energy Power Generation Facilities" came into effect in April 2019, aiming to further promote offshore wind power. Although the global weighted-average levelised cost of electricity (LCOE) for offshore wind is higher than PV according to IRENA (2020), the MOE for wind revealed in this study makes its implement easier.

The fact that MOEs vary across areas is also an important finding; indeed, the characteristics of a given area are taken into account in terms of the impact of this effect on electricity market prices when the government tries to promote renewable energy. As our results simply implied that electricity prices decrease and nearly reach zero as more renewable energy facilities are installed, recouping the initial investments of these facilities is important. The value of zero emission electricity generation should be prioritized to attract renewable energy investors. Even though the Japanese feed-in-tariff scheme is almost finished, the government should apply other options to eliminate the risks involved in developing renewable energy.

We found that renewable energy has a significant price-reducing effect in the wholesale electricity market of Japan. However, the variables used in this study did not fully explain the price changes observed. Thus, identifying other factors that determine electricity prices is a future research topic.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Japan Electric Power Exchange http://www.jepx.org/.

AUTHOR CONTRIBUTIONS

MS was involved in study design, dataset making, data analysis, and writing manuscript. HF critically revised the report, commented on drafts of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Grant-in-Aid for Scientific Research (JP18H03433) and (JP20K12283) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. The results and conclusions of this article do not necessarily represent the views of the funding agencies.

REFERENCES

Cabinet of Japan (2020). *Press Conference by the Chief Cabinet Secretary*. https://japan.kantei.go.jp/tyoukanpress/202010/26_p.html (accessed October 12, 2021).

Chubu Electric Power Grid, Demand Data. Available online: https://powergrid.chuden.co.jp/denkiyoho/ (accessed October 12, 2021).

Chugoku Electric Power Transmission and Distribution Company, Demand Data. Available online at: https://www.energia.co.jp/nw/service/retailer/data/area/index.html (accessed October 12, 2021).

Clò, S., Cataldi, A., and Zoppoli, P. (2015). The merit-order effect in the Italian power market: the impact of solar and wind generation on national wholesale electricity prices. *Energy Policy* 77, 79–88. doi: 10.1016/j.enpol.2014.11.038

Cludius, J., Hermann, H., Matthes, F. C., and Graichen, V. (2014). The merit order effect of wind and photovoltaic electricity generation in Germany 2008–2016: estimation and distributional implications. *Energy Econ.* 44, 302–313. doi: 10.1016/j.eneco.2014.04.020

Csereklyei, Z., Qu, S., and Ancev, T. (2019). The effect of wind and solar power generation on wholesale electricity prices in Australia. *Energy Policy* 131, 358–369. doi: 10.1016/j.enpol.2019.04.007

Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* 74, 427–431. doi: 10.1080/01621459.1979.10482531

Hagfors, L. I., Paraschiv, F., Molnar, P., and Westgaard, S. (2016). Using quantile regression to analyze the effect of renewables on EEX price formation. *Renew. Energy Environ. Sustain.* 1:32. doi: 10.1051/rees/2016036

Hokkaido Electric Power Network, Demand Data. Available online at: https://www.hepco.co.jp/network/renewable_energy/fixedprice_purchase/supply_demand_results.html (accessed October 12, 2021).

Hokuriku Electric Power Transmission and Distribution Company, Demand Data. Available online at: http://www.rikuden.co.jp/nw_jyukyudata/area_jisseki.html (accessed October 12, 2021).

IEA (2021). IEA World Energy Statistics and Balances 2021 (database). Available online at: https://www.iea.org/data-and-statistics/data-product/world-energy-balances (accessed October 12, 2021).

IPCC (2018). *Summary for Policymakers of IPCC Special Report on Global Warming of 1.5°C Approved by Governments*. Available online at: https://www.ipcc.ch/2018/10/08/summary-for-policymakers-of-ipcc-special-report-on-global-warming-of-1-5c-approved-by-governments/ (accessed October 12, 2021).

IRENA (2020). *Renewable Power Generation Costs in 2019*. Abu Dhabi: International Renewable Energy Agency.

Japan Electric Power Exchange. Available online at: http://www.jepx.org/market/index.html (accessed October 12, 2021).

Kansai Transmission and Distribution, Demand Data. Available online at: https://www.kansai-td.co.jp/denkiyoho/area-performance.html (accessed October 12, 2021).

Keeley, A. R., Matsumoto, K. I., Tanaka, K., Sugiawan, Y., and Managi, S. (2021). The impact of renewable energy generation on the spot market price in germany: ex-post analysis using boosting method. *Energy J.* 41:S11. doi: 10.5547/01956574.42.S12.akee

Kyushu Electric Power Transmission and Distribution, Demand Data. Available online at: https://www.kyuden.co.jp/td_service_wheeling_rule-document_disclosure.html (accessed October 12, 2021).

Maciejowska, K. (2020). Assessing the impact of renewable energy sources on the electricity price level and variability – A quantile regression approach. *Energy Econ.* 85:104532. doi: 10.1016/j.eneco.2019.104532

Maekawa, J., Hai, B. H., Shinkuma, S., and Shimada, K. (2018). The effect of renewable energy generation on the electric power spot price of the Japan electric power exchange. *Energies* 11:2215. doi: 10.3390/en11092215

Marshall, A. (1890). *Principles of Economics*. Amherst, NY: Palgrave Macmillan.

Ministry of Economy, Trade, and Industry (2018). *The Strategic Energy plan*. Available online: https://www.enecho.meti.go.jp/en/category/others/basic_plan/ (accessed November 04, 2021).

Ministry of Finance, Trade Statistics of Japan. Available online at: https://www.customs.go.jp/toukei/info/index_e.htm(accessed October 12, 2021).

Ministry of the Environment (2020). *Entrusted Work Concerning the Development and Disclosure of Basic Zoning Information Concerning Renewable Energies (FY 2019)*. Available online at: https://www.renewable-energy-potential.env.go.jp/RenewableEnergy/report/r01.html (accessed October 12, 2021).

Paraschiv, F., Ermi, D., and Pietsch, R. (2014). The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy* 73, 196–210. doi: 10.1016/j.enpol.2014.05.004

Quint, D., and Dahlke, S. (2019). The impact of wind generation on wholesale electricity market prices in the midcontinent independent system operator energy market: an empirical investigation. *Energy* 169, 456–466. doi: 10.1016/j.energy.2018.12.028

Renewable Energy Institute (2018). *Analysis of Wind Power Costs in Japan. 50–Translated From Japanese*. Available online at: https://www.renewable-ei.org/en/activities/reports/img/pdf/20180125/JapanWindPowerCostReport_EN_20180124.pdf (accessed October 12, 2021).

Renewable Energy Institute (2019). *Solar Power Generation Costs in Japan Current Status and Future Outlook*. Available online at: https://www.renewable-ei.org/pdfdownload/activities/Report_SolarPVCostJapan_EN.pdf(accessed October 12, 2021).

Shikoku Electric Power Transmission and Distribution Company, Demand Data. Available online at: https://www.yonden.co.jp/nw/renewable_energy/data/supply_demand.html (accessed October 12, 2021).

Sirin, S. M., and Yilmaz, B. N. (2020). Variable renewable energy technologies in the Turkish electricity market: quantile regression analysis of the merit-order effect. *Energy Policy* 144:111660. doi: 10.1016/j.enpol.2020.111660

TEPCO Power Grid, Demand Data. Available online at: https://www.tepco.co.jp/forecast/html/area_data-j.html (accessed October 12, 2021).

Tohoku Electric Power Network, Demand Data. Available online at: https://setsuden.nw.tohoku-epco.co.jp/download.html (accessed October 12, 2021).

Yoshihara, K., and Ohashi, H. (2017). *Assessing the Impact of Renewable Energy Sources: Simulation Analysis of the Japanese Electricity Market. Discussion Paper Series 17-E-063, RIETI.* Tokyo: Research Institute of Economy, Trade and Industry (RIETI).

Zipp, A. (2017). The marketability of variable renewable energy in liberalized electricity markets – An empirical analysis. *Renew. Energy* 113, 1111–1121. doi: 10.1016/j.renene.2017.06.072

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

APPENDIX

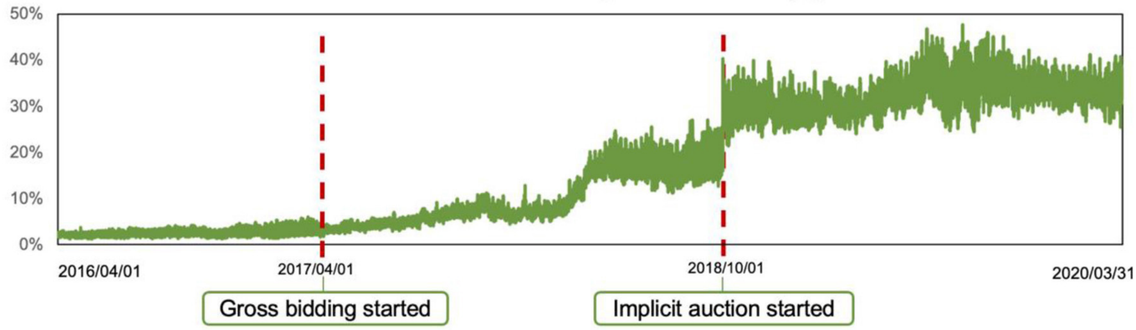


FIGURE A1 | Share of JEPX trade volume by total electricity generation. (Data source: JEPX day ahead spot price).

TABLE A1 | Results of correlation analysis for each area price.

	Hokkaido	Tohoku	Tokyo	Chubu	Hokuriku	Kansai	Chugoku	Shikoku	Kyushu
Hokkaido	1.000								
Tohoku	0.801	1.000							
Tokyo	0.791	0.989	1.000						
Chubu	0.556	0.657	0.644	1.000					
Hokuriku	0.557	0.657	0.644	0.997	1.000				
Kansai	0.557	0.657	0.644	0.997	1.000	1.000			
Chugoku	0.557	0.657	0.644	0.997	1.000	1.000	1.000		
Shikoku	0.549	0.651	0.639	0.988	0.991	0.991	0.991	1.000	
Kyushu	0.521	0.619	0.604	0.942	0.944	0.944	0.944	0.940	1.000

frontiers in Sustainability

ORIGINAL RESEARCH
published: 21 December 2021
doi: 10.3389/frsus.2021.815986

GIS Analysis of Solar PV Locations and Disaster Risk Areas in Japan

Kazuki Hao, Dimiter Ialnazov and Yosuke Yamashiki*

Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto, Japan

Following the global trend of climate change mitigation, Japan has been rapidly increasing its share of renewable energy, in particular, its share of solar energy. However, Japan has limited flat land area that is suitable for solar photovoltaic (PV) power generation and a high risk of natural disasters. There is a possibility that some of its newly built solar power plants are located in areas where landslides and floods are likely to occur. Therefore, it is important to study the locations for solar PV from the perspective of disaster risk management. Previous studies have reported a number of incidents where solar PV installations were damaged as a result of natural disasters. One study utilized geographical analysis technology to reveal the overlapping of solar PV powerplant locations and disaster-prone areas in Fukuoka prefecture in Japan. However, to our best knowledge, no previous research about the solar PV locations' hazard risks has been done on a national scale. This paper investigates the risks stemming from landslides and floods for the existing solar PV power plants in Japan. We compare the geographical data of disaster risks in Japan with the location data of solar PV power plants to investigate the number of solar PV power plants located in disaster risk areas. Our results show that the shares of medium and large-scale solar PV power plants located in areas where landslides and floods are likely to occur are about 8.5 and 9.1% respectively.

Keywords: solar PV, location of solar PV power plants, hazard map, disaster risks, GIS

Edited by:
Alexander Ryota Keeley,
Kyushu University, Japan

Reviewed by:
Yuki Ishimoto,
Institute of Applied Energy, Japan
Junya Kumagai,
Kyushu University, Japan

****Correspondence:***
Kazuki Hao
hao.kazuki.75w@st.kyoto-u.ac.jp

Specialty section:
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

Received: *16 November 2021*
Accepted: *23 November 2021*
Published: *21 December 2021*

Citation:
Hao K, Ialnazov D and Yamashiki Y (2021) GIS Analysis of Solar PV Locations and Disaster Risk Areas in Japan. Front. Sustain. 2:815986. doi: 10.3389/frsus.2021.815986

INTRODUCTION

The number of solar photovoltaic (PV) power plants in Japan has been increasing in recent years. Since the introduction of the Feed-in Tariff system in 2012 to promote renewable energy, solar PV power generation, which has a relatively short set-up process, in particular has played a central role (**Figure 1**).

Although the maximum power generation potential is limited now due to the grid connection capacity, the government plans to significantly increase renewable energy, including solar PV power generation by 2030, toward its target to achieve carbon neutrality by 2050. The 6th Strategic Energy Plan, which was approved by the Japanese government on October 22, 2021, included a goal of doubling the share of renewable energy sources in the country's energy mix from 2019 to 2030. According to that plan, the share of renewable energy in total power generation will increase from 18% in 2019 to 36–38% in 2030. Solar power generation will increase from 6.7% in 2019 to 14–16% in 2030, and wind power generation will increase from 0.7% in 2019 to 5% in 2030 (Agency for Natural Resources Energy, 2021b).

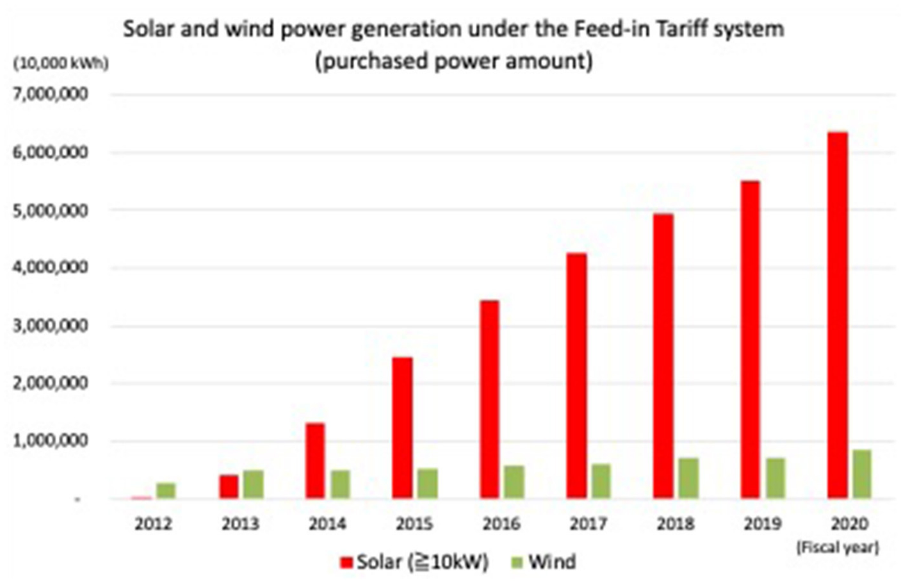


FIGURE 1 | Solar and wind power generation under the Feed-in Tariff system in Japan in 2012–2020. *Source*: Data from the Agency for Natural Resources Energy (2021a) (https://www.fit-portal.go.jp/PublicInfoSummary). Created by the authors.

Solar power generation accounts for a large part of the renewable energy introduced since the start of the Feed-in Tariff system in 2012, while the amount of wind power generation is not as large. This is thought to be due to the relatively long lead time required for the environmental impact assessments of wind power generation sites, including offshore plants (Li and Xu, 2019; Li et al., 2019; Obane et al., 2020). If the current trend continues, it is expected that the expansion of renewable energy in the near future will continue to be centered on solar PV power generation.

A major issue regarding the expansion of solar power generation is where to install the solar panels. The locations of the solar PV power plants that have been introduced so far are places where it's relatively easy to acquire land and install the solar panels, but the land areas in Japan that have potential to be used for solar PV are limited (Esteban et al., 2012; Sahu et al., 2016).

Therefore, when looking into future solar panel installation locations, the use of agricultural land, water surfaces such as lakes, and mountainous areas could all be possible candidates.[1] Regarding mountainous areas, solar PVs are being installed after deforestation and land reclamation (Ministry of Environment of Japan., 2019; Itaka, 2021). For those cases, the risk of disasters caused by heavy rains or landslides needs to be considered carefully.

However, only a few studies have so far examined the overlapping of the location of solar power plants and disaster risk areas. Solar power plants of 50 kW or higher are obliged to report accidents under the Electricity Business Act, and according to the Ministry of Economy, Trade and Industry (METI), there were a total of 57 accidents in the solar PV power plants in 2018 (**Table 1**). Ohgami (2019) from METI introduces the accidents reported under the Electricity Business Act. According to Ohgami (2019), for example, because of the heavy rains which hit vast areas of western Japan in July 2018, eight solar PV power plants were damaged by flooding and 11 by landslides. In addition, damages caused by strong winds and storm surges of the same year were also reported. The paper also reported the breakdown of damages by the size of the solar PV capacity. As a result of the heavy rains mentioned above, 7 out of 8 submerged solar power plants were <500 kW. Of the 11 solar PV power generation facilities affected by the landslides, 10 were 500 kW or more.

Tabata (2019) identified the damage caused by natural disasters to solar PV power generation from information available online and estimated the amount of damage to solar PV power plants in those cases. Tabata's paper called to the attention of solar power plant operators that a considerable amount of solar panels might be damaged by natural disasters. Kwon et al. (2020) used the geological information system (GIS) to identify solar power plants installed in areas which have risks of inundation and sediment-related disasters in Nogata City, Fukuoka Prefecture.

From an environmental point of view, the installation of solar power plants has caused a great loss of natural habitats and semi-natural habitats (Kim et al., 2021). The study of Kim et al. (2021) also found, by using simulation methods, that if the solar PV power plants are installed closer to a city, less habitats would be lost. Although all previous studies mentioned above discuss the risk of disasters related to solar PV power generation, they focus on case study analysis, or are confined to one region or to the effects on natural habitats.

[1]Since this paper targets medium-scale (between 500 KW and 10 MW) and large-scale (above 10 MW) solar power generation, it does not cover small installation locations such as the rooftops of houses or factories.

TABLE 1 | Accidents of solar PV powerplants due to natural disasters in Japan in 2018.

		Heavy rain in July 2018	21st typhoon in the season	Hokkaido Eastern Iburi earthquake	24th typhoon of the season
	Number of powerplants	19	23	3	12
Causes*	Submergence	8	–	–	–
	Landslide	11	–	–	–
	Wind	–	20	–	12
	Storm surge	–	3	–	–
Damaged parts*	Panel	10	21	2	12
	Power conductor	9	5	1	4
	Cubicle	4	1	–	–
	Others	9	7	2	9

*There are duplications.
Source: Ohgami (2019) and Ministry of Economy (2019).

In our paper, we estimate the significance of two types of disaster risks for the location of solar PV power plants in Japan. In particular, we examine the risk of sediment-related disasters and the risk of floods. We focus on these two types of disaster risks because they are the main causes of the damage on solar PV facilities as reported in Ohgami (2019). Although solar power plants can be also damaged by strong winds caused by typhoons, we did not calculate risk values based on past average wind speeds because the course of typhoons is wide and difficult to estimate.

Below, we first explain the data that we have used and then our research methods. Finally, we discuss our findings.

DATA

We used three types of datasets. The first is the area information of the solar PV power plants used in the research by Kim et al. (2021). This information is based on the location information of the power plants published in Electrical Japan (Electrical Japan, 2021).[2] The data from Electrical Japan is point data, which shows only one point in the area of the solar PV power plant. For our analysis, it was necessary to use polygon data, which is the location information showing the entire solar PV area where the solar panels are located.

The data of Kim et al. (2020) is a plot of the area based on the data from Electrical Japan. The data is also available online[3] (**Figure 2**). According to that dataset, the location information of Electrical Japan is updated periodically, while Kim et al. (2021) uses information from December 30, 2020. In addition, their study targeted both Japan and South Korea, while our study focuses only on Japan.

In Japan, there is no official GIS database showing where solar power plants are currently located. The Ministry of Economy, Trade and Industry has the information on the operators and date of approval that licenses solar power plant operators, but it is not disclosed in association with the GIS location information.

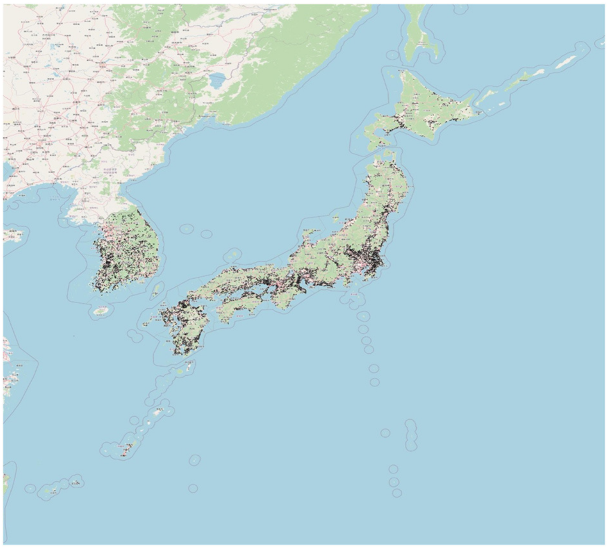


FIGURE 2 | Data of solar PV power plant locations (displayed on the OpenStreetMap with QGIS). *Source*: Kim et al. (2020).

Therefore, we made the judgement that the data used in Kim et al. (2021) was optimal for our analysis.

The data includes only medium-sized and larger solar PVs. The capacity of medium-scale solar PVs are larger than 500 kW and <10 MW, and that of large-scale solar PVs are >10 MW. As this study targets the disaster risks of medium- and large-scale solar PVs, small-scale solar PVs, including those on the rooftops of houses and factories are omitted.

The second data set used in this paper is the information about sediment-related disaster-prone areas of the Digital National Land Information provided by the Ministry of Land, Infrastructure, Transportation and Tourism of Japan[4] (MLIT) (National Land Information Division, 2021b) (**Figure 3**). The

[2]http://agora.ex.nii.ac.jp/earthquake/201103-eastjapan/energy/electrical-japan/type/8.html.ja.
[3]Available online: https://doi.org/10.5281/zenodo.4644036.
[4]Available online: https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A33-v1_4.html (in Japanese).

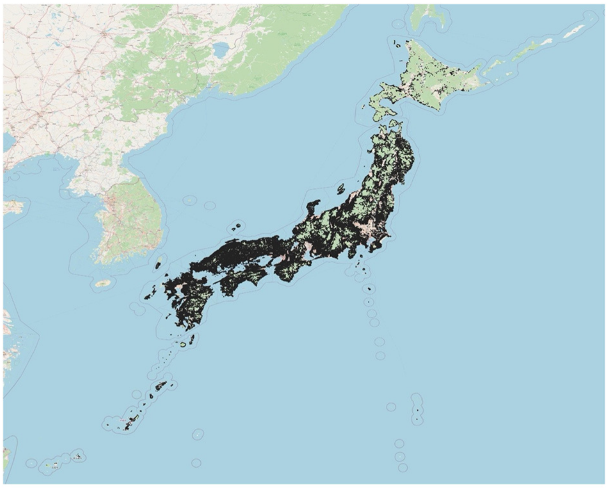


FIGURE 3 | Data of sediment disaster hazard areas (displayed on the OpenStreetMap with QGIS). *Source*: Sedimental disaster prone areas by National Land Information Division, MLIT (2021).

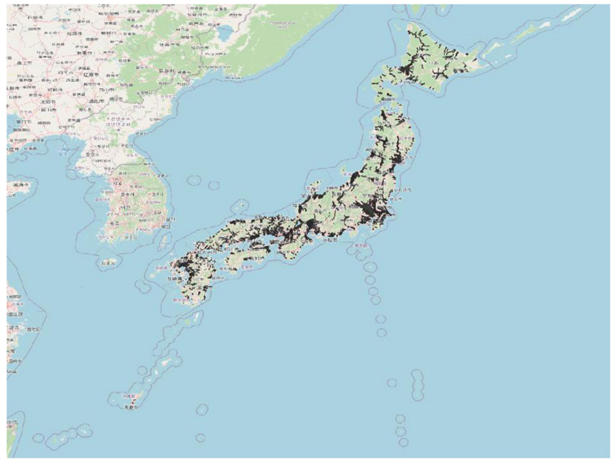


FIGURE 4 | Data of floods and inundation hazard areas (displayed on the OpenStreetMap with QGIS). *Source*: Floods and inundation hazard areas by National Land Information Division, MLIT (2021).

information is in polygon style and is provided in the format of a shapefile. There are two types of areas covered, sediment disaster hazard areas and special hazard areas. According to the Act on Sediment Disaster Countermeasures for Sediment Disaster Prone Areas, sediment disaster hazard areas are areas where sediment disasters are likely to occur, with incidents of slope failures, landslides, and debris flows. The dataset includes all these three types of areas in a single dataset, and as such, there are some overlapping areas among them. Sediment disaster special hazard zones are areas with the risk of buildings being destroyed, which may cause severe injuries or fatalities to the residents (Disaster Management Bureau Cabinet Office of Japan., 2018). Sediment disaster hazard zones include all the areas of special hazard zones, so in this paper, we utilized the data of sediment disaster hazard zones. Data is provided by prefecture, and we used data from all 47 prefectures of Japan. The dataset we used was the latest dataset available, which was of August 2020.

The third data set is the floods and inundation hazardous area information of the Digital National Land Information from the MLIT (National Land Information Division, 2021a), in which the flood warning area in polygon data is included as a shapefile[5] (**Figure 4**).

RESEARCH METHODS

As explained in the previous section, we used two polygon data, a sediment disaster hazard areas and flood and inundation hazard areas, as indicators related to disaster risks. We also used the location data of the solar power plants on GIS to find the number of points where the polygons overlap.

[5] Available online: https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A31-v2_2.html (in Japanese).

FIGURE 5 | An example of overlapping solar PV area (black lines) with sedimental disaster prone area (purple) on Google Maps. *Source*: Created by the authors.

```
def PlysOverlap(plys1, plys2):
    l = []
    for i in range(len(plys1)):
        if any(plys2.overlaps(plys1[i])):
            l.append(plys1[[i]])
    return(len(l))
```

FIGURE 6 | Basic structure of the function used. *Source*: Created by the authors.

For the calculation, we used the overlap function in the GeoPandas library (https://geopandas.org/) of Python, a programming language often used for data analysis (Zhou et al., 2021). The function takes two parameters of polygon data and returns true when the two polygon data are combined and overlapped (**Figure 5**). We accumulated the number in which the function returned true, as seen in the basic function shown below (**Figure 6**). The two disaster datasets of each prefecture and the location data of the solar power plants were input as two

TABLE 2 | Number of solar power plant locations overlapping with high landslide risk areas.

Prefecture code	Prefecture name	Results	Prefecture code	Prefecture name	Results
01	Hokkaido	9	25	Shiga	16
02	Aomori	6	26	Kyoto	15
03	Iwate	3	27	Osaka	15
04	Miyagi	13	28	Hyogo	69
05	Akita	2	29	Nara	18
06	Yamagata	3	30	Wakayama	22
07	Fukushima	3	31	Tottori	1
08	Ibaraki	19	32	Shimane	12
09	Tochigi	17	33	Okayama	15
10	Gunma	25	34	Hiroshima	71
11	Saitama	7	35	Yamaguchi	19
12	Chiba	10	36	Tokushima	8
13	Tokyo	1	37	Kagawa	21
14	Kanagawa	10	38	Ehime	10
15	Niigata	3	39	Kochi	14
16	Toyama	1	40	Fukuoka	35
17	Ishikawa	2	41	Saga	17
18	Fukui	7	42	Nagasaki	27
19	Yamanashi	9	43	Kumamoto	21
20	Nagano	19	44	Oita	9
21	Gifu	42	45	Miyazaki	18
22	Shizuoka	33	46	Kagoshima	38
23	Aichi	9	47	Okinawa	1
24	Mie	37		**Total**	**782**

Source: Authors' own calculations.

variables. We also calculated the number of solar PV power plant areas overlapping with both the sediment disaster hazard areas and the flood and inundation hazard areas.

As explained above, the disaster data is separated for each prefecture, thus when formulating the program, we used the prefecture code starting from Hokkaido and ending with Okinawa to compare the disaster areas and the location of the solar power plants in order.

For data processing, both disaster-related data and solar power plant-related data used the shapefile format. For geocoding, EPSG: 4326, one of the standard geocoding systems, was used.

RESULTS

As a result, we found the number of points where the sediment disaster hazard areas and the flood and inundation hazard areas in each prefecture overlap with the solar PV power plant locations. Our results are shown in **Tables 2–4** for each prefecture with the total number calculated at the bottom.

In the dataset, the total number of the solar power plants in Japan was counted as 9,250 points. Among them, 782 points were found to overlap with the sediment disaster hazard areas in Japan as a whole, which is ∼8.5% of the total (**Table 2**). Also, among all the solar PV power plant locations in Japan, the number of points overlapping with the flood and inundation hazard areas was 846. This represents about 9.1% of the total solar PV power plant locations in Japan (**Table 3**).

Also, there are 30 solar PV locations which overlap with both the landslides hazard areas and the floods and inundation hazard risk areas (**Table 4**). This means that about 0.3% of the total solar PV power plants are facing both the landslides risk and the floods and inundation risk.

DISCUSSION

Where to construct solar power plants is an important issue because the Japanese government policies aim to significantly increase the capacity of the solar power plants in the future. However, due to prevalence of mountainous areas in Japan, the flat land that could be used as a location for the solar PV power plants is limited. If the flat land where solar power plants can be easily built is exhausted, the next potential candidates will include mountainous or steep zones. In that case, before we start to select the sites for the future solar PV power plants, it is necessary to take into consideration not only the expected deforestation and other damages to the natural environment, but also the risks of sediment-related and of flood-related disasters.

In our study, we have shown that a number of the solar power plants in Japan are located in the sediment-related disaster and flood inundation hazard areas. Although it has been pointed

TABLE 3 | Number of solar power plant locations overlapping with high flooding and inundation risk areas.

Prefecture code	Prefecture name	Results	Prefecture code	Prefecture name	Results
01	Hokkaido	111	25	Shiga	31
02	Aomori	1	26	Kyoto	11
03	Iwate	3	27	Osaka	19
04	Miyagi	31	28	Hyogo	33
05	Akita	3	29	Nara	4
06	Yamagata	14	30	Wakayama	1
07	Fukushima	13	31	Tottori	4
08	Ibaraki	45	32	Shimane	4
09	Tochigi	37	33	Okayama	12
10	Gunma	21	34	Hiroshima	6
11	Saitama	65	35	Yamaguchi	2
12	Chiba	56	36	Tokushima	7
13	Tokyo	2	37	Kagawa	4
14	Kanagawa	9	38	Ehime	4
15	Niigata	10	39	Kochi	7
16	Toyama	9	40	Fukuoka	20
17	Ishikawa	8	41	Saga	8
18	Fukui	7	42	Nagasaki	0
19	Yamanashi	8	43	Kumamoto	19
20	Nagano	11	44	Oita	11
21	Gifu	29	45	Miyazaki	22
22	Shizuoka	33	46	Kagoshima	4
23	Aichi	28	47	Okinawa	0
24	Mie	59		**Total**	**846**

Source: Authors' own calculations.

out that site planning in consideration of disasters is important (Rediske et al., 2019), there have been few studies based on concrete calculations. If we look at the entire country of Japan, there has been no study done on the number of overlapping points between the location of solar power plants and the disaster risk areas. The findings of our research results show that the locations of a number of power plants overlap with the areas where sediment-related disasters and floods are highly likely to occur.

In relation to sediment-related disasters, it is necessary to consider the above risks when the demand for electricity derived from renewable energy sources increases in the future. In relation to flood disasters, the classification of solar PV power generation, i.e., whether it is classified as a rooftop solar PV or as an utility-scale solar PV, becomes an important issue. In areas with flood disaster risks, solar PV panels that are located on the rooftops of buildings are less likely to be damaged. On the other hand, for large-scale solar power plants located in flat areas, the flood disaster risks should be considered as an important factor to prevent future potential damage.

In our research, we focused on the solar PV power plants that have already been constructed. We have shown that a certain number of power plants that have already been constructed are located at points that are susceptible to sediment-related disasters or to flood hazards. We believe that our results are useful in suggesting that disaster risks should be taken into consideration when planning the locations of future solar power PV plants. As extreme weather events are likely to occur more frequently in the future, it will become even more important to consider disaster risk factors when selecting a power plant's location.

What could be some possible solutions regarding the location of solar power plants in Japan in the future? One possible solution is to build all new power plants in areas where there is no risk of disaster. However, that is not a realistic answer for future solar PV plant installations, as the future demand for renewable sourced energy is expected to rapidly increase. Another solution is to keep building solar power plants in disaster-risk areas but add some form of insurance against potential disasters. Based on the data about which areas could be highly impacted by disasters, we propose an insurance scheme for solar power plants in these high-risk areas. Our result potentially supports the development of insurance system for the solar PV power plant operators, as our study revealed that some solar PV power plants already locate in the disaster risk areas, and that the system can potentially be used the evaluation of the risk type and magnitude for solar PV upon conducting relevant improvement. The insurance could be provided by the government (central or local) or the private sector. The insurance scheme could promote a swift recovery from the damage caused by disasters and thus reduce the risks faced by solar power operators. We support the argument by Fraser (2019) that local governments can play an important role for an equitable and locally engaged energy transition.

TABLE 4 | Number of solar power plant locations overlapping with both high landslides risk areas and high flooding and inundation risk areas.

Prefecture code	Prefecture name	Results	Prefecture code	Prefecture name	Results
01	Hokkaido	0	25	Shiga	0
02	Aomori	0	26	Kyoto	0
03	Iwate	0	27	Osaka	0
04	Miyagi	0	28	Hyogo	5
05	Akita	0	29	Nara	0
06	Yamagata	0	30	Wakayama	1
07	Fukushima	0	31	Tottori	0
08	Ibaraki	4	32	Shimane	0
09	Tochigi	1	33	Okayama	1
10	Gunma	1	34	Hiroshima	2
11	Saitama	0	35	Yamaguchi	0
12	Chiba	3	36	Tokushima	1
13	Tokyo	0	37	Kagawa	0
14	Kanagawa	0	38	Ehime	1
15	Niigata	0	39	Kochi	1
16	Toyama	0	40	Fukuoka	1
17	Ishikawa	0	41	Saga	0
18	Fukui	0	42	Nagasaki	0
19	Yamanashi	0	43	Kumamoto	2
20	Nagano	1	44	Oita	0
21	Gifu	2	45	Miyazaki	0
22	Shizuoka	0	46	Kagoshima	2
23	Aichi	0	47	Okinawa	0
24	Mie	1		**Total**	**30**

Source: Authors' own calculations.

In addition, as our results in **Tables 2–4** show, the number of solar PV power plants located in high-risk areas displays a wide range of differences among prefectures. The prefectures have a variety of characteristics in their size, terrain, and amount of investment into solar PV. The local governments as well as private companies should properly identify these differences and consider expanding solar PV where suitable by taking into consideration the perspective of disaster prevention.

The risk of typhoons and strong winds was not included in our study. The reason is that, due to the wide range of typhoon potential paths, it is difficult to show the overlap between the typhoon risk and the location of the installed solar panels. However, as Ohgami (2019) shows, typhoons and strong winds can also lead to solar power plant accidents. Studies have also been conducted on the location and wind strength of solar power plants to conflict with potential wind farm locations (Obane et al., 2020). In future research, we believe that the risk of typhoons or strong winds for the location of the solar PV power plants should be studied as well.

The Law Concerning the Promotion of the Measures to Cope with Global Warming, which was revised in May 2021, has a provision saying that local prefectural governments should set goals and specify areas to promote renewable energy (Ministry of Environment of Japan, 2021a). It has been decided by an Ordinance of the Ministry of the Environment what locations should become the promotion areas for renewables. However, the Ministry of the Environment Ordinance does not exclude areas with high risk of disasters. At the time of writing (Nov. 2021), the Ministry of the Environment is planning to discuss with experts and decide a policy to exclude disaster risk areas from the promotion area of renewables (Ministry of Environment of Japan, 2021b). Therefore, our analysis of the overlap between the locations of the solar PV power plants and the disaster risk areas, could be also relevant for those forthcoming policy discussions.

Japan is a country where many natural disasters occur and where it is absolutely necessary to study the impact of disasters. However, we believe that research on the impact of natural disasters in Japan will be also relevant for countries where there have been few disasters so far due to the expected changes in the climate in the future. It will also serve as an important reference when promoting the solar PV power generation in developing countries where natural disasters due to climate change are expected to increase.

CONCLUSION

In this study, we performed a geographically-integrated analyses seeking for evaluating risks stemming from landslides and floods for the existing solar PV power plants in Japan. We compared the geographical data of disaster risks in Japan with the location data of solar PV power plants to investigate the number of solar PV power plants located in disaster risk areas by GIS-based integrated survey developed by authors. We focused on how many points overlap in the locations of the solar PV power plants constructed so far in Japan with the areas where there is a high risk of sediment-related disasters and of floods or inundation.

As a result, we counted the total number of the solar power plants in Japan as 9,250 points. Among them, 782 points were found to overlap with the sediment-related disaster hazard areas as a whole, which is ∼8.5% of the total solar PV locations in Japan. Also, among all the solar PV locations, the number of points overlapping the flood and inundation hazard areas was 846. This is about 9.1% of the total number. Our results not only show that a significant number of solar power plants are located in the disaster risk areas, but also contribute to the discussion on future potential sites for solar PV installations.

A limitation of our study is that not all disaster risks were included in the analysis. We did not include the risks of typhoons and strong winds although typhoons and strong winds cause a large number of accidents at solar power plants. In addition, it could be said that areas with strong winds are also areas suitable for wind power generation. Further research on those issues is needed. For example, if we assume that the areas where maximum wind speed in the past exceeded a certain speed, which might damage solar PV power generation, are risky areas, the method used in this paper can be utilized to identify the risks from strong winds for the solar PV power plants.

Another possible direction for future research is to create framework that evaluates the degree of risk if the solar PV power plant is located in the landslides and floods hazard areas. In this research, we didn't evaluate the various degrees of risks. Future studies are needed, for example, to quantify the risks or calculate the potential costs of constructing solar PV in hazardous areas.

In addition, when applying this research to other regions or countries, it is necessary to have information on the locations of solar PV power plants and disaster risk information in those regions or countries. The information on the locations of solar power plants was limited even in the case of Japan. We believe that further data construction on solar power generation and disasters will be useful for future research, not only in Japan but around the world.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://nlftp.mlit.go.jp/ksj/index.html, https://doi.org/10.5281/zenodo.4644036.

AUTHOR CONTRIBUTIONS

KH: conceptualization, methodology, data analysis, and writing original draft. DI and YY: conceptualization, reviewing, and editing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank Dr. Kim Ji Yoon and other co-researchers who contributed to create the solar PV location dataset in Japan, and also thank Prof. Kitamoto Asanobu, who is the author of Electrical Japan, and his team for sharing an outstanding database of GIS based power plant information on their web site. The authors thank all those who worked to prepared disaster risk datasets on the digital national land information system, prepared by the Ministry of Land, Infrastructure, Transportation and Tourism of Japan. The authors also thank Ms. Cassandra Ling for the English proofreading.

REFERENCES

Agency for Natural Resources and Energy. (2021b). *Website for Feed-in Tariff Information*. Available online at: https://www.fit-portal.go.jp/PublicInfoSummary (accessed on November 15, 2021).

Agency for Natural Resources and Energy. (2021a). *The 6th Strategic Energy Plan*. Available online at: https://www.enecho.meti.go.jp/en/category/others/basic_plan/pdf/6th_outline.pdf (accessed on November 15, 2021).

Disaster Management Bureau Cabinet Office of Japan. (2018). *Report on Flood and Sediment Disaster*. Available online at: http://www.bousai.go.jp/kyoiku/pdf/h30_tebikisho_english.pdf (accessed on November 15, 2021).

Electrical Japan (2021). "Solar power plant locations," in *Japanese*. Available online at: http://agora.ex.nii.ac.jp/earthquake/201103-eastjapan/energy/electrical-japan/type/8.html.ja (accessed on November 15, 2021).

Esteban, M., Zhang, Q., and Utama, A. (2012). Estimation of the energy storage requirement of a future 100% renewable energy system in Japan. *Energy Policy* 47, 22–31. doi: 10.1016/j.enpol.2012.03.078

Fraser, T. (2019). How governance and disasters shape renewable energy transitions: the case of Japanese mega-solar. *Soc. Sci. Q.* 100, 975–990. doi: 10.1111/ssqu.12603

Itaka, S. (2021). Economic Analysis of Cedar Plantation Management and Mega-Solar Replacement. *Forests* 12:361. doi: 10.3390/f12030361

Kim, J. Y., Koide, D., Ishihama, F., Kadoya, T., and Nishihiro, J. (2020). "Solar PV data: Current site planning of medium to large solar power systems accelerates the loss of the remaining semi-natural and agricultural habitats [Data set]," in *Science of the Total Environment* Zenodo, 146475.

Kim, J. Y., Koide, D., Ishihama, F., Kadoya, T., and Nishihiro, J. (2021). Current site planning of medium to large solar power systems accelerates the loss of the remaining semi-natural and agricultural habitats. *Sci. Total Environ.* 779:146475. doi: 10.1016/j.scitotenv.2021.146475

Kwon, D., Bunme, P., and Mitani, Y. (2020). "Extraction of solar photovoltaic power plants in disaster areas using GIS and construction of a method for proposing safe and efficient solar photovoltaic power plants installation locations," in *Japanese. Joint Conference of Electrical, Electronics and Information Engineers in Kyushu*, 470–471.

Li, A., and Xu, Y. (2019). The governance for offshore wind in Japan. *Energy Procedia* 158, 297–301. doi: 10.1016/j.egypro.2019.01.092

Li, A., Xu, Y., and Shiroyama, H. (2019). Solar lobby and energy transition in Japan. *Energy Policy* 134:110950. doi: 10.1016/j.enpol.2019.110950

Ministry of Economy, Trade and Industry (METI). (2019). *Report on the Recent Trends and Future Directions Regarding Solar cell Power Generation Facilities*. Available online at: https://www.meti.go.jp/shingikai/sankoshin/hoan_shohi/denryoku_anzen/newenergy_hatsuden_wg/pdf/015_03_00.pdf (accessed on November 15, 2021).

Ministry of Environment of Japan. (2019). *Report of the Study Group on the Basic Concept of Environmental Impact Assessment for Solar Power Generation Facilities*. Available online at: https://www.env.go.jp/press/files/jp/110948.pdf (accessed November 15, 2021).

Ministry of Environment of Japan. (2021a). *About the Cabinet Decision to Amend a Part of the Law Concerning the Promotion of the Measures to Cope with Global Warming*. Retreived from: http://www.env.go.jp/press/109218.html (accessed November 15, 2021).

Ministry of Environment of Japan. (2021b). *Expectations and Issues Regarding Future Local Government Implementation Plans*. Retreived from: https://www.env.go.jp/policy/council/52keikaku-manual/y520-01b/mat04_1-1-3.pdf (accessed November 15, 2021).

National Land Information Division, National Spatial Planning, and Regional Policy Bureau MLIT of Japan. (2021b). "Sedimental disasters prone areas," In *Japanese*. Available online at: https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A33-v1_4.html (accessed November 15, 2021).

National Land Information Division, National Spatial Planning, and Regional Policy Bureau MLIT of Japan. (2021a). "Floods and inundation hazardous areas," in *Japanese.* Available online at: https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A31-v2_2.html (accessed November 15, 2021).

Obane, H., Nagai, Y., and Asano, K. (2020). Assessing land use and potential conflict in solar and onshore wind energy in Japan. *Renewable Energy* 160, 842–851. doi: 10.1016/j.renene.2020.06.018

Ohgami, H. (2019). Incidents of accident damages caused by natural disasters at solar power plants. *J. Japan Soc. Mater. Cycles Waste Manage*. 30, 408–412. doi: 10.3985/mcwmr.30.408

Rediske, G., Siluk, J. C. M., Gastaldo, N. G., Rigo, P. D., and Rosa, C. B. (2019). Determinant factors in site selection for photovoltaic projects: A systematic review. *Int. J. Energy Res.* 43, 1689-1701. doi: 10.1002/er.4321

Sahu, A., Yadav, N., and Sudhakar, K. (2016). Floating photovoltaic power plant: a review. *Renew. Sustain. Energy Rev.* 66, 815–824. doi: 10.1016/j.rser.2016.08.051

Tabata, T. (2019). Damage to the solar photovoltaic power generation system due to natural disasters. *Proceed. Japan Soc. Mater. Cycles Waste Manag*. 30:13.

Zhou, S., Li, Y., Chi, G., Yin, J., Oravecz, Z., Bodovski, Y., et al. (2021). GPS2space: an Open-source Python Library for Spatial Measure Extraction from GPS Data. *J. Behav. Data Sci.* 1, 1–29. doi: 10.35566/jbds/v1n2/p5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Sustainability

ORIGINAL RESEARCH
published: 18 March 2022
doi: 10.3389/frsus.2022.837427

Structural Equation Modeling as a Route to Inform Sustainable Policies: The Case of Private Transportation

Iljana Schubert[1], Sylvain Weber[2,3], Adan L. Martinez-Cruz[4], Paul Burger[1] and Mehdi Farsi[2]*

[1] Sustainability Research Group, Department of Social Sciences, University of Basel, Basel, Switzerland, [2] Institute of Economic Research, University of Neuchâtel, Neuchatel, Switzerland, [3] University of Applied Sciences and Arts of Western Switzerland (HES-SO), Haute école de gestion de Genève, Geneva, Switzerland, [4] Department of Forest Economics, and Centre for Environmental and Resource Economics (CERE), Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden

OPEN ACCESS

Edited by:
Alexander Ryota Keeley,
Kyushu University, Japan

Reviewed by:
Kiriaki M. Keramitsoglou,
Democritus University of Thrace, Greece
Shouheng Sun,
Université du Québec à Chicoutimi, Canada

***Correspondence:**
Iljana Schubert
iljana.schubert@unibas.ch

Specialty section:
This article was submitted to Quantitative Sustainability Assessment,
a section of the journal Frontiers in Sustainability

Received: 16 December 2021
Accepted: 01 February 2022
Published: 18 March 2022

Citation:
Schubert I, Weber S, Martinez-Cruz AL, Burger P and Farsi M (2022) Structural Equation Modeling as a Route to Inform Sustainable Policies: The Case of Private Transportation. Front. Sustain. 3:837427. doi: 10.3389/frsus.2022.837427

The availability of big data allows a wide range of predictive analyses that could inform policies for promoting sustainable behaviors. While providing great predictive power, adopted models fall short in explaining the underlying mechanisms of behavior. However, predictive analyses can be enhanced by complementary theory-based inferential analyses, guiding tailored policy design to focus on relevant response mechanisms. This paper illustrates the complementary value of multidisciplinary inferential models in informing large predictive models. We focus on Structural Equation Modeling, an approach suitable for a holistic examination of different pathways and hypotheses from multiple disciplines. Drawing on an interdisciplinary theoretical framework we develop an empirically tractable model and apply it to a sample of household data from Switzerland. The model focuses on the relationships that delineate the underlying mechanisms for energy consumption behaviors in the case of private transportation. The results are discussed in light of possible contributions to policies aiming at the promotion of sustainable travel behavior as well as data requirements for analyses relying on big data.

Keywords: Structural Equation Modeling (SEM), interdisciplinary models, big data, intervention pathways, sustainable transport choices

INTRODUCTION

Widespread digitisation in various sectors of advanced economies brings about a fundamental change in the availability of consumption data. Smart meters, connected appliances and electric vehicles, to mention a few, allow unprecedented access to individual- and household-level data. Such data is increasingly used to predict various consumption behaviors and to tailor marketing messages to specific segments of the population. However, so far, the potential arising from big data is not often harnessed for promoting sustainable consumption and/or pro-social behavior. This lag can be partly explained by the distinctive methodologies used by experts in separate fields. On the one hand, focusing on predictive analyses, marketing experts rely on models with a great number of variables in order to identify patterns of behavior across different groups. On the other hand, focusing on inferential analyses, social scientists usually rely on parsimonious models to identify underlying mechanisms and to explain behavior. While the latter use theoretical premises to provide a relatively rigid structure to meet their empirical models, the former use models with less theoretical structure, allowing predictions to be mainly data-driven. Both types of analyses are essential and complementary. We need to understand behavioral mechanisms, in particular for promoting sustainable consumption, while at the same time utilizing the predictive power of the emerging big data. There is, however, a methodological tension hindering mutual feedback, as revealed by some large-scale studies (e.g., O-Power study, Allcott and Rogers, 2014) where reductions in energy consumption were achieved, but underlying mechanisms for change remained unclear, hence impeding widespread usage beyond a specific context.

In order to tap into the emerging data potentials for promoting sustainability, we need to identify adoption tendencies. In addition, we need to understand the barriers and drivers for different groups in the population. While predictive models are sufficient for the former objective, the latter requires testing specific hypotheses derived from theories. However, theories usually originate from different disciplines, and fail to provide a holistic picture of the consumption behavior of interest. Moreover, by focusing on a single aspect of behavior, usually dictated by a rigorous causality analysis, the analyst inevitably leaves out many variables that might be relevant for a comprehensive analysis.

Inference and prediction can complement each other. Inferential analysis, using statistical models, provides a basis for a sound and theory-driven interpretation whereas predictive models, based for example on machine learning, are less interpretable but provide a powerful framework for data-driven predictions. Recognizing the complementarity of predictive and inferential analyses in large data sets, we put forward the notion that comprehensive structural models can be used to bridge the chasm between the two types of analyses. Instead of zooming into a single aspect of behavior, comprehensive models include a multitude of variables integrated into a relatively rigid structure. The adopted structure can be based on a comprehensive framework rooted in several disciplinary theories. Such an empirical analysis can be conducted by structural equation models (SEM) which offer several advantages. First, compared to other statistical models such as linear regression, SEM have a greater flexibility to accommodate a multitude of pathways for a given outcome. Regression models could also be used for causality analysis, but their focus on a specific aspect restricts their ability for considering multiple hypotheses from various disciplines. Second, SEMs not only provide a holistic picture with a relatively large number of variables, they can also be used to assess the relative importance of various causal pathways. Finally, as opposed to predictive models based on data mining and machine learning methods, SEMs can provide an overall picture of behavior, used for generating relevant hypotheses to be tested with further regression models. Therefore, SEM can be used as a "prime language for causal analysis", as put by Pearl (2012), to provide a conceptual structure to predictions purely driven from data.

Our empirical analysis is based on a broad dataset that contains "distance driven by each household (HH)", but also additional information on HH decisions, like socioeconomic and demographic characteristics, norms and values. The objective is primarily to show how SEMs, including multiple pathways, can play a complementary role to predictive models and disciplinary theory-based analyses. Our empirical illustration in the field of private transportation further provides insights into the challenges in translating a comprehensive framework into empirically applicable models and thus also highlights data requirements. The insights gained into behavioral mechanisms driving HH transportation decisions do not constitute the main focus of this paper. They are used to exemplify the added value of applying SEM for modeling different pathways. Our contribution is to highlight the interplay of determinants behind consumer decisions, and the extent to which SEMs based on an interdisciplinary framework can play a complementary role to predictive and inferential models. As such, it is primarily directed to inform a sustainametrics conversation—i.e., a discussion on how increasingly available data can support transitions to sustainable societies and limitations to such a role—rather than a transportation behavior and policy discussion.

This paper is structured as follows. In Section Background we provide rationales for choosing yearly distance traveled by private car as a relevant issue, and SEM as our method. A brief description of the integrated framework proposed by Burger et al. (2015) is presented in Section The Integrated HH Energy Consumption Framework. It is followed by a mapping of the underlying relationships between the factors listed in the framework (Sections Implementation of the IHECF, Relationships of Social Opportunity Space and Individual Opportunity Space Factors, and Relationships of Decision-Making Factors and Choices/Routines). The paper then proceeds to lay out the data used for the empirical analysis (Section Method), followed by a presentation of the results (Section Results), discussion (Section Discussion) and conclusions (Section Conclusion).

BACKGROUND

Achieving low-carbon energy goals heavily depends on shifting demand (over time) to match supply (Shove, 2021). Many studies have pointed out behavioral barriers hampering policy interventions in reducing HH energy consumption. These obstacles range from undesirable consequences of public policies (e.g., Alberini et al., 2018) to a number of barriers operating at an individual and HH level (Cattaneo, 2019; del Mar Solà et al., 2021), such as rebound effects (De Borger et al., 2016; Stapleton et al., 2016), missing price incentives or imperfect knowledge (Allcott and Greenstone, 2012; Pothitou et al., 2016), overestimated technical promises (Fowlie et al., 2018) as well as fixed routines or habits determining daily life (Kurz et al., 2015; Kent, 2021). Most of these barriers operate through behavioral mechanisms, for example driven by cognitive heuristics (Kahneman, 2003), emotions (Brosch and Sander, 2014), values and norms (Ababio-Donkor et al., 2020; Bouman et al., 2021).

In this context, the availability of large amounts of information on HH behavior—energy consumption, in particular—could provide a window into the functioning of existing interventions and the potential of unexplored solutions. For instance, taking advantage of availability of hourly data on Swedish residential electricity usage, Brännlund and Vesterberg (2021) have explored whether there is a potential for shifting load between peak and off-peak hours. If possible, this load shifting would be a game changer as it would allow the covering of expected increases in demand without substantial infrastructure adjustments. Their analysis, while indicating a limited potential, does not, however, provide much insight about how such a shift could be achieved through interventions or policies. Indeed, many empirical analyses of rich datasets are unable to shed light into behavioral mechanisms needed to design policy interventions (e.g., Karimu et al., 2022). These studies often focus on tangible factors, and miss out modeling potentially important characteristics such as attitudes, emotions, and values. A structural model could be helpful to investigate the behavioral links between consumption-shifting with policy-relevant characteristics such as individual preferences and attitudes.

Notably, knowledge about mechanisms and barriers to behavior change is rooted in disciplinary frameworks which do not commonly consider the interplay of multiple factors determining energy demand. While interdisciplinary work is on the rise (e.g., De Witte et al., 2013; Van Acker et al., 2014; Stephenson et al., 2015; Stephenson, 2018; Koszowski et al., 2019), disciplinary divides, such as those between economics, psychology, sociology and geography, often prevent integrated analyses of determinants and the formulation of comprehensive and tailored intervention strategies (Burger et al., 2015; Hess et al., 2018).

In most empirical studies, the focus is either on data-driven predictive models or theory-driven single-equation regressions. However, neither of the two approaches is able to model multiple pathways of energy consumption behaviors. In this paper, we develop a comprehensive model of HH energy consumption and show how such models can be implemented with a SEM to provide structural framing of predictive analyses such as Moro and Holzer (2020). To this end, we build on Burger et al. (2015) who put forward an interdisciplinary model of HH energy consumption based on major empirical and disciplinary findings of research from the fields of psychology, sociology, geography, consumer behavior science, and economics. Reasons for choosing the Integrated HH Energy Consumption Framework (IHECF) by Burger et al. (2015) is precisely the fact that it is an interdisciplinary multi-theory-based [e.g., Rational Choice Theory (RCT), Theory of Planned Behavior (TPB), Value Belief Norm Theory (VBN), Norm Activation Model (NAM), Consumer Theory (CT), Behavioral Decision Theory (BDT), Social Practice Theory (SPT)] framework synthesizing the established main drivers of HH energy consumption. Although there are other multi-disciplinary frameworks (e.g., De Witte et al., 2013; Giulio et al., 2014; Götschi et al., 2017; Koszowski et al., 2019) and theories to the best of our knowledge there is none which is as comprehensive as the IHECF.

In the empirical analysis, we build a SEM explaining a major domain of HH energy consumption. Specifically, we focus on (self-reported) yearly distance traveled by Swiss HHs using their private car, a behavior that public policies can target to yield potentially large energy savings. In fact, the transport sector accounts for a third of total energy consumption in the European Union (Eurostat, 2021) and in Switzerland (SFOE, 2020). We develop the model in an illustrative manner, with the intention of highlighting challenges in translating such comprehensive models into empirically applicable models and outline data requirements.

The Integrated HH Energy Consumption Framework

The *IHECF* (**Figure 1**), proposed by Burger et al. (2015), distinguishes two types of energy consumption behaviors (ECBs): material-specific behaviors (e.g., buying a car), and action-specific behaviors (e.g., driving the car). These behaviors are influenced by a multitude of individual and socio-economic factors, embedding individuals with their choices and routines in a broad environment.

This broad environment is characterized by factors related to *social opportunity spaces (SOS)* and *individual opportunity spaces (IOS)*. IOS factors provide individual boundaries framed by SOS factors, that is, the external societal circumstances in which the individual is embedded. SOS factors include characteristics of available technology and facilities, economic factors (including prices), institutional norms and policies, geographic and climatic factors as well as demographic and cultural differences. IOS factors include aspects describing the individuals' social environment (i.e., social context, milieu and lifestyle), their socio-economic setup (i.e., personal appliances and facilities, place of dwelling, HH size) and their socio-demographics, such as income, age, gender and knowledge.

Individuals base their decisions on a complex interplay of internal *decision-making* factors and IOS/SOS factors. Internal *decision-making* factors are *attitudes, control, norms, values, heuristics* and *biases,* as well as *emotions*, which influence choices and routines, and ultimately ECB. Choices can either be habitual, i.e., embedded in routines (e.g., always driving to shops instead of taking other transport modes), or deliberate (e.g., purchasing an electric vehicle).

In addition to describing different types of ECB and their determinants, the IHECF makes suggestions regarding governance factors (i.e., instruments, institutions and actors) that could be activated for (re-)shaping ECBs. It is therefore designed to guide interdisciplinary research on energy consumption and offers a way of organizing findings and viewpoints from different disciplines. Due to its interdisciplinary nature and comprehensiveness, the IHECF is not based on a single theory, thus, causal claims must be established through empirical research.

Implementation of the IHECF

Utilizing the IHECF as foundation, we develop an empirically estimable model by linking the factors in the framework and drawing the relationships that delineate the underlying ECB

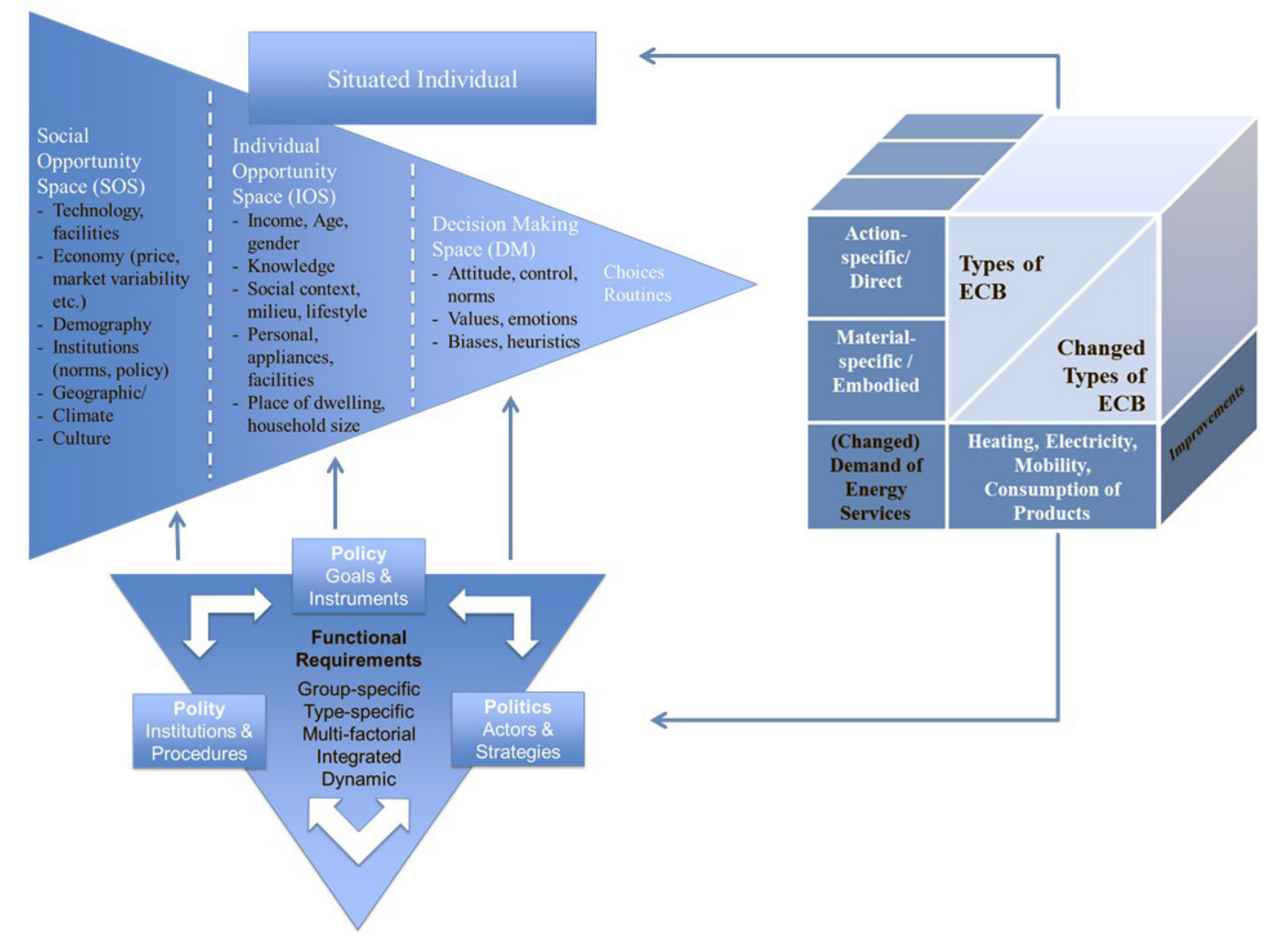


FIGURE 1 | The integrated HH energy consumption framework (IHECF, reproduced from Burger et al., 2015).

mechanisms. A complete overview of the relationships can be seen in **Figure 2**. Aiming at an empirically tractable model, we focus on forward flows, that is, relationships going from the broad end of the triangle (SOS) to the tip (choices and routines). The model does not, however, specify relationships in a purely sequential manner. Instead, while some factors influence other factors from an adjacent level, certain factors from all categories (SOS, IOS, decision-making and choices/routines) are modeled as direct predictors of behavior, as can be found in the literature on ECB. To this end, we draw upon empirical research and theories (e.g., TPB, VBN, CT, RCT, and BDT)[1] from psychology, sociology, geography, consumer-behavior science, and economics, focusing on the most robust findings. The generic model specified in **Figure 2** represents a tractable option of the IHECF. Without repeating the discussion in Burger et al. (2015), we present some major findings in the ECB literature in the following subsections to underpin the relevance of the four dimensions in our generic model.

Relationships of Social Opportunity Space and Individual Opportunity Space Factors

Starting with the SOS and IOS factors located at the base of the IHECF triangle, we outline the relationships of several determinants traditionally considered in economics. From an economic point of view, utility-maximizing individuals decide on their level of energy consumption while considering unit *price* of energy and available *income* (e.g., Borenstein, 2015). This has led to abundant literature on price- and income-elasticities of energy demand (e.g., Havranek and Kokes, 2015; Labandeira et al., 2017). As, however, energy is not consumed *per se*, but used as an input in the HH production function, the effect of energy prices should be mediated by the characteristics of *appliances* (e.g., vehicles or electronic devices), which in turn depend on the available *technology* and the *facilities* available to the HH.

Technology must be included in the model, as technological progress results in greater efficiency, hence lowering the price

[1]Because this is not a theoretical paper but the application of a validated interdisciplinary multi-theory framework, we refrain from diving deeper into the set of theoretical approaches which form its basis.

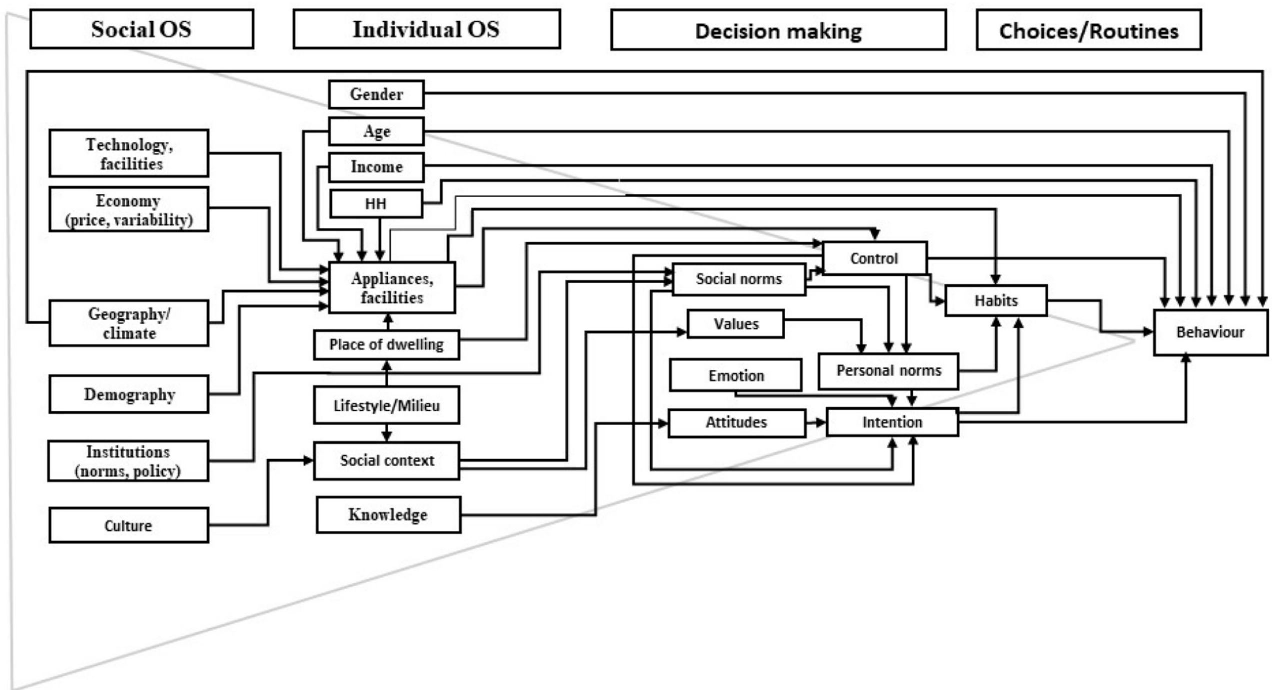


FIGURE 2 | Generic interdisciplinary model of ECB developed from the IHECF (Burger et al., 2015).

of energy service and stimulating demand (cf. also Fowlie et al., 2018 on relevance of available technology). For instance, when a HH purchases a more efficient car, driving becomes cheaper, possibly resulting in a greater usage. This "rebound" effect could offset part of the expected energy consumption reductions. While the existence of rebound effect is widely accepted, its magnitude remains contentious, with empirical estimates for private car travel ranging from negligible to almost 100% (e.g., Azevedo, 2014).

Other SOS factors include g*eography* and *climate*, taken in a generic way to refer to weather-related factors (e.g., temperature and precipitation), topographic characteristics of the HH's location (e.g., elevation and slope), and atmospheric conditions such as greenhouse gas concentration. All these conditions are also structural determinants of energy usage usually labeled as "demand shifters" (e.g., Kavousian et al., 2013; Winkler et al., 2014).

Demographic factors at the societal level, such as population size, age structure, urbanization, and population density, as well as at the HH level, such as *HH size*, *gender*, and *age,* have been found to significantly impact energy consumption and travel (Brounen et al., 2012; Liddle, 2014; Karatasou et al., 2018; Buylova, 2020). HH size is a structural determinant as, for instance, a 5-person HH naturally consumes more energy than a 2-person HH. However, economies of scale result in a less than proportional increase in energy usage for every additional member. Moreover, the composition of the HH (e.g., size, presence of children) and the type of accommodation increases the availability and use of appliances (e.g., number of cars and intensity of usage, cf. De Witte et al., 2013). Also, ECB appears to be associated with occupants' age, gender, education and ethnicity partly because of differences in activities (e.g., Brounen et al., 2012; McLoughlin et al., 2012; Tweed et al., 2015; Karatasou et al., 2018; Buylova, 2020).

In addition, and given existing evidence, socio-cultural characteristics and how they relate to potentially mediating factors such as values or norms should be integrated in the model. *Culture* has been defined as "the integrated pattern of meanings, beliefs, norms, symbols and values that individuals hold within a society, with values representing perhaps the most central cultural feature" (Oreg and Katz-Gerro, 2006, p. 466). Accordingly, culture provides the broader social context through which individuals learn what is valued and acceptable in their society. Moreover, *institutions* provide a broader context in which social norms are perceived (Allcott, 2011; Ostrom, 2014), leading to social norms as mediator between institutions and ECB. For example, Stephenson et al. (2015) and Stephenson (2018) developed a cultures framework to investigate energy cultures, whereas Stoppok et al. (2018) compared daily energy consumption practices of Kenyan, German, and Spanish households.

Between the broad socio-cultural context and the narrower context of social groups, *lifestyle*- and *milieu*-groups also play a role and have been observed to behave differently in terms of ECB (Spaargaren, 2003; Sütterlin et al., 2011; Schubert et al., 2020). Some groups, for example, perceive more social pressure through their social context to engage in ECB than others (Sütterlin et al., 2011; Schubert et al., 2021). A social milieu encompasses people, the physical and social conditions underlying traditions and values, which are relatively stable and resistant to societal changes (Mochmann and El-Menouar, 2005). Lifestyles, on the other hand, express a person's social position through their behavior and consumption patterns (Van Acker et al., 2014; Schubert et al., 2020). Whereas, social context mediates the influence of cultural and lifestyle/milieu on values and social norms, lifestyle and milieu, in turn, mediate other IOS factors such as the place of dwelling. Finally, environmental and energy-related *knowledge* can affect ECB via attitudes (e.g., Nayum and Klöckner, 2014; Pothitou et al., 2016).

Relationships of Decision-Making Factors and Choices/Routines

Psychological theories [e.g., TPB, NAM, VBN, and Social Cognitive Theory (SCT)] and empirical research (cf. meta-analysis by Klöckner, 2013, themselves based on main psychological theories) have identified *intention*, *control*, *habits/routines*, *heuristics/biases, attitudes, norms, values,* as well as *emotions*, as important predictors of environmental behavior, including ECB.

Intentions, *habits/routines*, and *control* are direct predictors of ECB. Even though intentions are not explicitly listed as a decision-making factor in the original IHECF (**Figure 1**), it constitutes a major predictor of ECB and is seen as a gauge of people's willingness to adopt environmentally friendly ECBs (e.g., Tan et al., 2017; Sun et al., 2018). We therefore include intention in our model (**Figure 2**), mediating the relationship between ECB and *attitudes*, *norms*, *emotions*, and *control* (Klöckner, 2013; Hiratsuka et al., 2018; Brosch, 2021).

Control constitutes a further major predictor of ECB, known as perceived behavioral control (PBC) in the TPB (Ajzen, 1991, 2006) and self-efficacy in the SCT (Bandura, 2001). Control explains how able—based on the circumstances and skills—a person feels to perform certain behaviors and is influenced by social norms (e.g., Klöckner, 2013; Fu, 2021).

A large amount of our daily behavior is deemed habitual with very little deliberation (Marien et al., 2018), also referred to as routines. Similar to habitual choices (i.e., using a car for commuting), one-off decisions (e.g., what car to buy) could have habitual aspects such as brand loyalty (Nayum and Klöckner, 2014). Habits are also related to *heuristics* and *biases* (Verplanken and Aarts, 1999; Klöckner, 2013) and can be considered as mechanisms for focusing on certain aspects of a complex decision while ignoring others (Tversky and Kahneman, 1974).

For simplicity, we assume that heuristics and biases are included in habits and do not consider separate relationships. Indeed, to indicate the overlap with routines and to highlight the habitual element in choices, we place habits in-between decision-making factors, choices, and routines. Habits have, in fact, been found to be a main predictor of behavior, mediating the relationship of ECB with intentions, personal norms and control (e.g., Nayum and Klöckner, 2014). Naturally, the available facilities and appliances, for instance access to a specific car, have an impact on the formation and persistence of habits (Klöckner and Matthies, 2009; Klöckner and Blöbaum, 2010; Hess and Schubert, 2019; Punzo et al., 2021). In addition, *personal norms*, people's personal and moral considerations, mediate the relationship of control, social norms and values with intentions and habits (Stern, 2000; Klöckner, 2013).

METHODS

To empirically illustrate our framework-based approach we carried out a SEM analysis. SEMs are a collection of different integrated analytical techniques. These include for example, path analysis (regression analysis) and factor analysis. Factor analysis can be utilized to estimate latent factors from observed variables. Path analysis, on the other hand, offers the opportunity to estimate the effect of one or more variables on others and hence allows the investigation of various hypotheses in a single model. SEM can fit data from experimental, non-experimental and observational studies. SEMs are able to simultaneously estimate multiple interrelated relationships of endogenous and exogenous variables, and account for measurement errors. In addition, SEMs provide fit statistics to evaluate the implications of theoretical assumptions or relationships (Bollen and Pearl, 2013).

The SEM developed in this paper is specified as displayed in **Figure 3**. The SEM is estimated using Stata 16, with all variables demeaned (i.e., the variable's mean is subtracted from all values so that the resulting variable is centered at zero), and covariations were only allowed between exogenous variables and not between measurement errors. All structural paths are grounded in theory as mentioned above and successfully tested in previous empirical studies (cf. Sections Relationships of Social Opportunity Space and Individual Opportunity Space Factors and Relationships of Decision-Making Factors and Choices/Routines).

For the estimation, we rely on Full Information Maximum Likelihood (FIML), which implies using all available data, even observations with missing values. For robustness check, we also applied the (standard) Maximum Likelihood (ML) approach on the subsample composed only of observations without missing values. We conclude that there are no systematic disparities between the FIML and ML sample[2].

For the estimated model we report both unstandardized and standardized coefficients (**Appendix Tables**). The two sets of coefficients are complementary: unstandardized coefficients provide quantitative impacts of the covariates on the endogenous variables, thus these path coefficients are in the same unit as the endogenous variable of that path. Standardized coefficients reveal the relative importance of each covariate, thus these coefficients are unit-free and therefore make it possible to compare variables of different magnitudes.

Data

Data analyzed in this paper was collected in April and May 2017, as part of the second wave of the Swiss Household

[2]For space reasons we only report FIML findings but ML findings are available upon request.

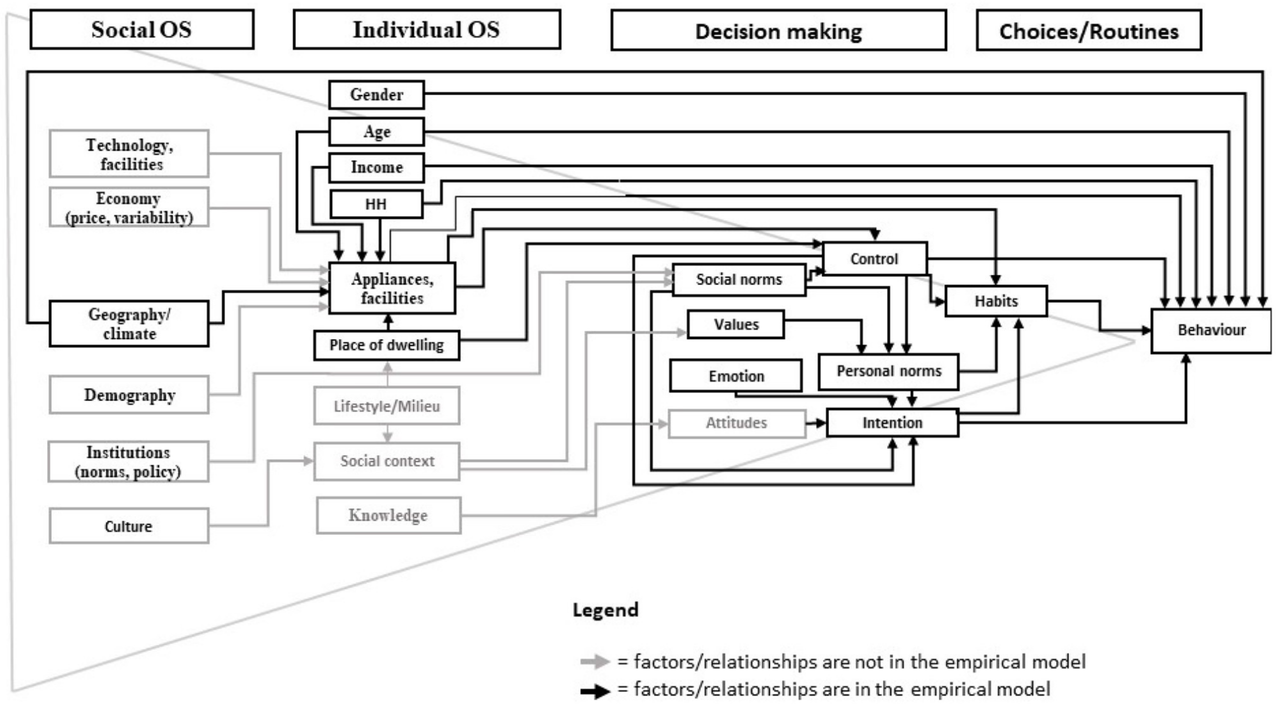


FIGURE 3 | Empirical implementation of the model.

Energy Demand Survey (SHEDS)[3]. SHEDS respondents are representative of the Swiss population[4] according to age, gender, region and home ownership. Respondents self-report their equipment and usage in several energy consumption domains (heating, electricity, and mobility), socio-demographic, psychological (e.g., environmental attitudes, values, etc.) and sociological characteristics (e.g., life events, etc.). Our illustrative empirical analysis (using FIML)[5] focuses on 3,362 car owners—a subsample of the entire SHEDS sample (see **Table A1**). This includes respondents who own a car running on gasoline or diesel, that is, about 73% of the entire sample of 5,015 HHs[6].

Table 1 provides an overview of all variables considered in our IHECF framework-informed model. Annual mileage is obtained as the answer to the question "On average, how many kilometers do you drive per year?" and is only asked to car owners. Further details of the psychological constructs are provided in the **Table A2**). In order to adapt to the available data, we exclude a number of variables from the final empirical SEM, as depicted in **Figure 3**[7].

Model Fit

Basing our decisions on Hu and Bentler's (1999) criteria, we find that our model[8] satisfies the suggested RMSEA[9] fit statistic (.04, cut off < 0.06). Other fit statistics are slightly outside recommended ranges, such as CFI (0.83, cut off 0.90 or higher) and χ^2/DF ratio ($\chi^2 = 7{,}394.15$, DF = 993, $\chi^2/DF = 7.45$, cut off < 2–5, $p < 0.001$), however research has shown that the optimal threshold depends on numerous features of the model, including estimation method, sample size, number of degrees of freedom, and the extent to which assumptions of multivariate normality are met (Hu and Bentler, 1999; Marsh et al., 2004; Tomarken and Waller, 2005)[10]. Noting that other large SEMs in the literature

[3]Weber et al. (2017) provide a detailed description of SHEDS, which was based on information of the IHECF.

[4]SHEDS collects data from all parts of Switzerland except Ticino, the Italian-speaking canton representing less than 5% of the Swiss population.

[5]The ML analysis includes 922 observations. Despite their important size difference, the two samples do not show statistical differences in main variables. An exception is the respondent age, which is on average lower in the ML sample.

[6]The focus is on cars running on gasoline or diesel. We excluded 152 observations corresponding to cars with other engine types (e.g., electric or hybrid cars) and 151 outlier observations with evident reporting errors in particular, suspiciously large reported values for fuel consumption.

[7]More precisely, the reasons for excluding a variable are as follows: (1) Data availability (i.e., attitudes, social context, institutional norms or policies) for 2017; (ii) Mediating factor missing (e.g., for knowledge); (iii) Multicollinearity and heavily unbalanced distribution of respondents across categories (for lifestyle categories).

[8]The final measurement model (of the latent psychological variables) shows appropriate fit statistics (FIML: $\chi 2 = 1{,}005.251$, DF = 202, $p < 0.001$, CFI = 0.969, RMSEA = 0.034). Drawing on modification indices, confirmatory factor analysis, composite reliability, discriminant and convergent validity as well as model fit statistics, we exclude five items showing low loadings on the latent constructs (details in **Tables A3, A4**).

[9]Often presented together with SRMR fit statistics, which were only available for the ML model and also acceptable [=0.053 (ML only), cut off < 0.08].

[10]As robustness checks, we estimated reduced versions of our model (available upon request), which show improved fit statistics.

TABLE 1 | Overview of the IHECF factors in the SEM.

Model factors	Variables included in SEM
Behavior	Kilometers driven per year by private car, scale: "up to 5,000 km"—"more than 50,000 km", steps of 5,000 km, with an "I don't know" option
SOS variables	
Geography/ climate	HH living in rural environment vs. city HH living in suburbs vs. city German-speaking parts of Switzerland vs. French-speaking part of Switzerland (Romandy) Commuting distance from home to work (km)
SOS factors excluded from SEM	Technology/facilities, Economy (e.g., market price, market variability), Institutions (norms, policy), Culture
IOS variables	
Gender	Gender
Age	Age categories: 18–34, 35–54, 55+
Income	HH income (log of mid-points)
Household (HH) composition	Number of HH members HH members younger than 6 years of age HH members older than 65 years of age
Appliances, Facilities	Number of general transport passes per HH member Number of regional transport passes per HH member New car (≤1 year) Fuel consumption in gasoline-equivalent (L/100 km) Car fuel type: diesel (vs. gasoline) Transmission type: automatic (vs. manual) More than one car in HH
Place of dwelling	Location index: Index for the distance from home to various places [Range 0–32: 8 places located from < 0.5 km (=1) to more than 2 km away (=4). Respondents who never go to a place are assigned 0.] Owned flat (vs. rented flat, combination of tenure and dwelling type) Rented house (vs. rented flat, combination of tenure and dwelling type) Owned house (vs. rented flat, combination of tenure and dwelling type) Away from home in second holiday residence (share of weeks per year)
IOS factors excluded from SEM	Social context, Lifestyle/Milieu, Knowledge
Decision-making (DM)	
Social norms	Combined descriptive norm (Perceptions of how others behave in relation to ECB) and injunctive norm (Perceptions of how others expect me to behave in relation to ECB)
Personal norms	Feelings of obligation to perform "sustainable" energy related behaviors
Values	Hedonic values: concerning personal pleasure and comfort Egoistic values: concerning personal gains Altruistic values: concerning the welfare of other people Biospheric values: concerning the quality of the natural environment
Control	Perceptions of control and ability to behave environmentally friendly
Emotions	Positive emotions toward own environmentally friendly behavior
Intention	Intention to reduce car usage within the next 12 months
Habits	Preferred modes of transport for commuting and leisure travel
DM factors excluded from SEM	Attitudes

*The decision-making variables are further described in the **Table A2**.*

show similar goodness-of-fit statistics (Bouscasse and Bonnel, 2016) we deem the fit statistics for our model acceptable. Overall, the model reports an explained variance (R^2) of 17% for the main endogenous variable (km driven/year) and the R^2 for the fifteen further endogenous variables range from 1 to 82% (see **Table A5**).

RESULTS

Our full SEM results are provided in the **Tables A5, A6**. Each factor in the model may directly or indirectly affect distance traveled (the final outcome in our model). Direct predictors (for instance gender and age) are connected to distance traveled without mediating factors, whereas indirect predictors (for instance social and personal norms) affect distance only through a mediating factor (habits in the case of social and personal norms). There can be several mediating factors between an indirect predictor and the final dependent variable. Predictors may also affect distance traveled both directly and indirectly, in which case the total effect is given by the sum of both. Thus, overall, there are a number of different possible pathways which can explain a given behavior that our analysis depicts. **Table A6** reports all direct, indirect and total effects.

To facilitate the interpretation of results, here we provide different figures summarizing the standardized coefficients of different investigated pathways.

Direct Pathways Explaining Km-Driven/Year

Figure 4 depicts the direct relationships or pathways between the main endogenous variable, annual km driven by private car, and different decision-making/routine, IOS and SOS factors. Findings show that among the direct decision-making/routine factors, only choices related to habitual transportation mode are significant. Unsurprisingly, people that routinely use public transport and soft modes (walking, cycling etc.) for commuting or leisure travel drive less km/year by car. Compared to the reference group (i.e., people using their private car as main travel modes), individuals that habitually use public transport or soft modes for commuting drive about 3,000 km less on average, and those using these modes of transport for leisure purposes drive about 2,000 km less per year on average (**Table A6**, total effects).

Several direct IOS-factors are significant. For instance, subscriptions to a general or a regional transport pass is associated with a lower usage of private car, about 5,000 and 3,000 km/year less, on average. On the other hand, people who own a diesel or automatic car drive about 3,000 and 700 km/year more on average, respectively, than those with gasoline engines or manual transmissions. Additionally, owners of new cars (defined as cars registered up to 1 year before implementation of our survey) drive around 800 km/year more on average than those with older cars. Looking at demographics we find that higher income is associated with greater car usage with 130 km/year increase for 10% increase in income on average; female respondents drive about 1,200 km/year less on average; and younger people (18–34 years of age) drive more with an average difference of 900 km/year with the middle-age (35–54 years) and 1,700 km/year with the old (55 years or more) respondents.

Several SOS factors, in particular geographical aspects, also explain annual mileage by private car. People living in the

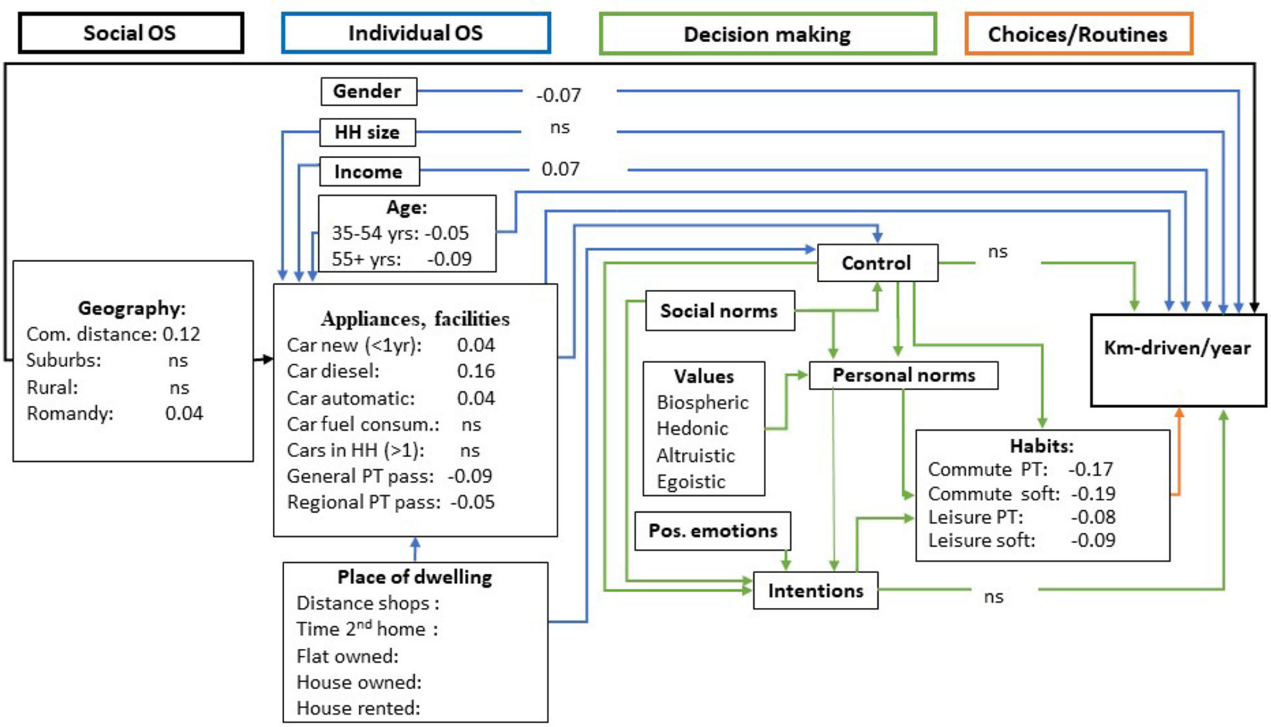


FIGURE 4 | Direct pathways explaining km-driven/year. The arrow colors depict the four different components of the empirical model. Reported are standardized coefficients of direct effects with $p < 5\%$ (complete results are provided in **Table A5**), ns = not significant at 5%; PT = public transport; soft = soft mobility; fuel consum. = fuel consumption.

French-speaking region of Switzerland (Romandy) drive about 700 km/year more on average, than those in the German-speaking region. In line with commuting transport habits, distance from home to work is also related to annual mileage: each increase by 1 km in the commuting distance is associated with a 40 km increase of the annual distance driven.

Some direct relationships of decision-making, IOS and SOS predictors of behavior are non-significant. These are relationships of behavior with control (DM), intention (DM), HH size and owning more than 1 car in the HH (IOS), rural and suburban dwelling (vs. city; SOS).

Overall, the strongest predictors are commuting habits, structural factors related to the type of engine (diesel vs. gasoline) and public transport passes (general and regional) (see **Table A6**, standardized total effects).

Indirect Pathways Explaining Km-Driven/Year *via* Habitual Transport Choices

In addition to distinguishing between direct and total effects for the main dependent variable (km driven/year), one strength of SEMs is that they can help identify different pathways of underlying mechanism for developing interventions. In **Figure 5** we display a number of such pathways focusing on the habitual transport choice for commuting and leisure, and their own explanatory variables.

People with a general or regional travel pass are more likely to use public transport for commuting and leisure. Likewise, higher personal norms, control and intentions to reduce car use/carbon footprint explain public transport choice for leisure.

Habitual use of soft modes (walking/cycling) for commuting and leisure transport are explained by a similar group of direct predictors, namely intentions, control and partly personal norms (only for commuting). Interestingly, we observe that intentions to reduce car use/carbon footprint are negatively related to habitual use of soft modes, whereas it is positively related to habitual use of public transport.

Exploring significant explanatory factors of habitual commuting and leisure transport choices further we find that higher intentions are related to higher control and positive emotions but lower social and personal norms.

Control, the feeling of being able to change one's behavior toward more environmentally friendly alternatives, is positively related to social norms (normative information from friends and family who behave and expect others to behave in a pro-environmental way), time spent at 2nd home and property ownership (vs. renting). There is a negative relationship between control and having a diesel car, fuel consumption and having more than one car in the HH. There is no significant relationship between control and having a new car or driving an automatic

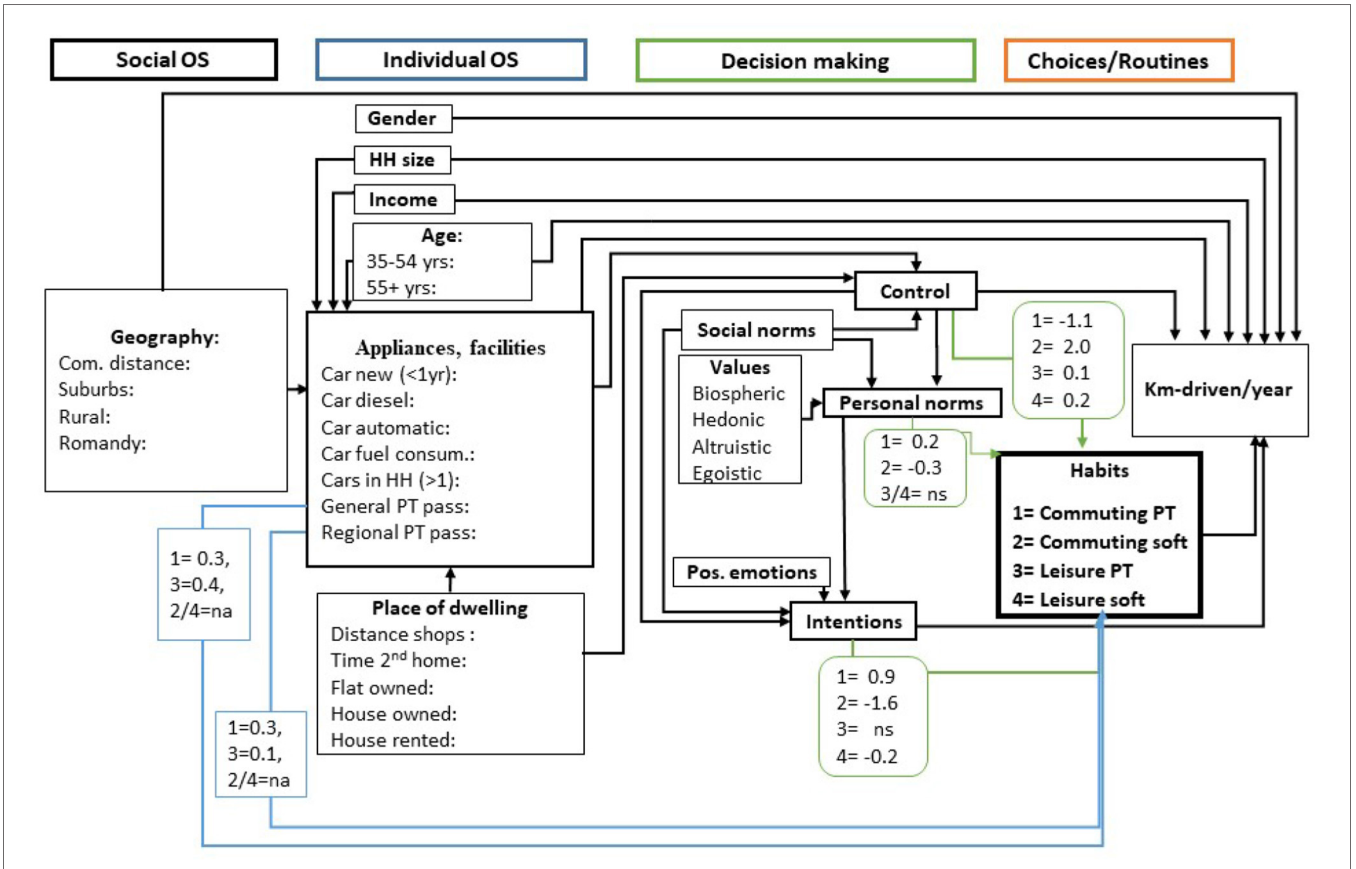


FIGURE 5 | Indirect pathways explaining km-driven/year via habits for commuting and leisure by public transport (PT) and soft mobility (e.g., bike, walking). Reported are standardized direct coefficients with $p < 5\%$ (complete results are provided in **Tables A5, A6**). Numbers 1–4 refer to the different types of habitual behavior: 1: habitual use of public transport (PT) for commuting, 2: habitual use of soft transport measures for commuting, 3: habitual use of PT for leisure purposes, 4: habitual use of soft mobility measures for leisure purposes. na = not applicable, meaning this path was not tested; ns = not significant at 5%.

car, ownership of general and regional transport passes, distance to amenities, or renting a house (vs. renting a flat).

High personal norms—moral personal standards to behave pro-environmentally—are related to a perception of control, positive social norms and positive biospheric (nature-focused) values. Personal norms are negatively related to egoistic (self-focused) and altruistic (other-focused) values. There is no significant relationship between personal norms and hedonic (pleasure-focused) values.

Indirect Pathways Explaining Km-Driven/Year *via* Relevant IOS and SOS Factors

Investigating the underlying mechanisms of IOS and SOS factors allows us to further understand possible pathways for behavior change (**Figure 6**). For example, we observe a higher incidence of diesel (vs. gasoline) engines in rural areas and in German-speaking regions. Larger commuting distances are also related to having diesel cars or newer cars, hence higher fuel efficiency. On the other hand, individuals with higher incomes are more likely to have cars with lower fuel efficiency. Living in rural and suburban areas is related to lower numbers of general and regional transport passes. Finally, having a higher commuting distance and living in the French-speaking region is related to larger numbers of general transport passes in the HH.

There are a number of non-significant relationships between the IOS and SOS variables, such as income and dwelling location (rural and suburb vs. city dwelling; German vs. French-speaking Swiss regions) that do not appear to be related to owning a new car. Furthermore, having a new car, the commuting distance or dwelling location do not seem to be related to fuel consumption.

DISCUSSION

In this paper, we illustrate how a SEM informed by a multidisciplinary framework can bridge the divide between predictive big data models and single-equation disciplinary models. Specifically, we estimate an interdisciplinary model of energy consumption behavior using data on annual mileage by private vehicle. The bridge can be seen to be established if we can reveal different underlying mechanisms and identify their relative importance, here to reduce private vehicle usage,

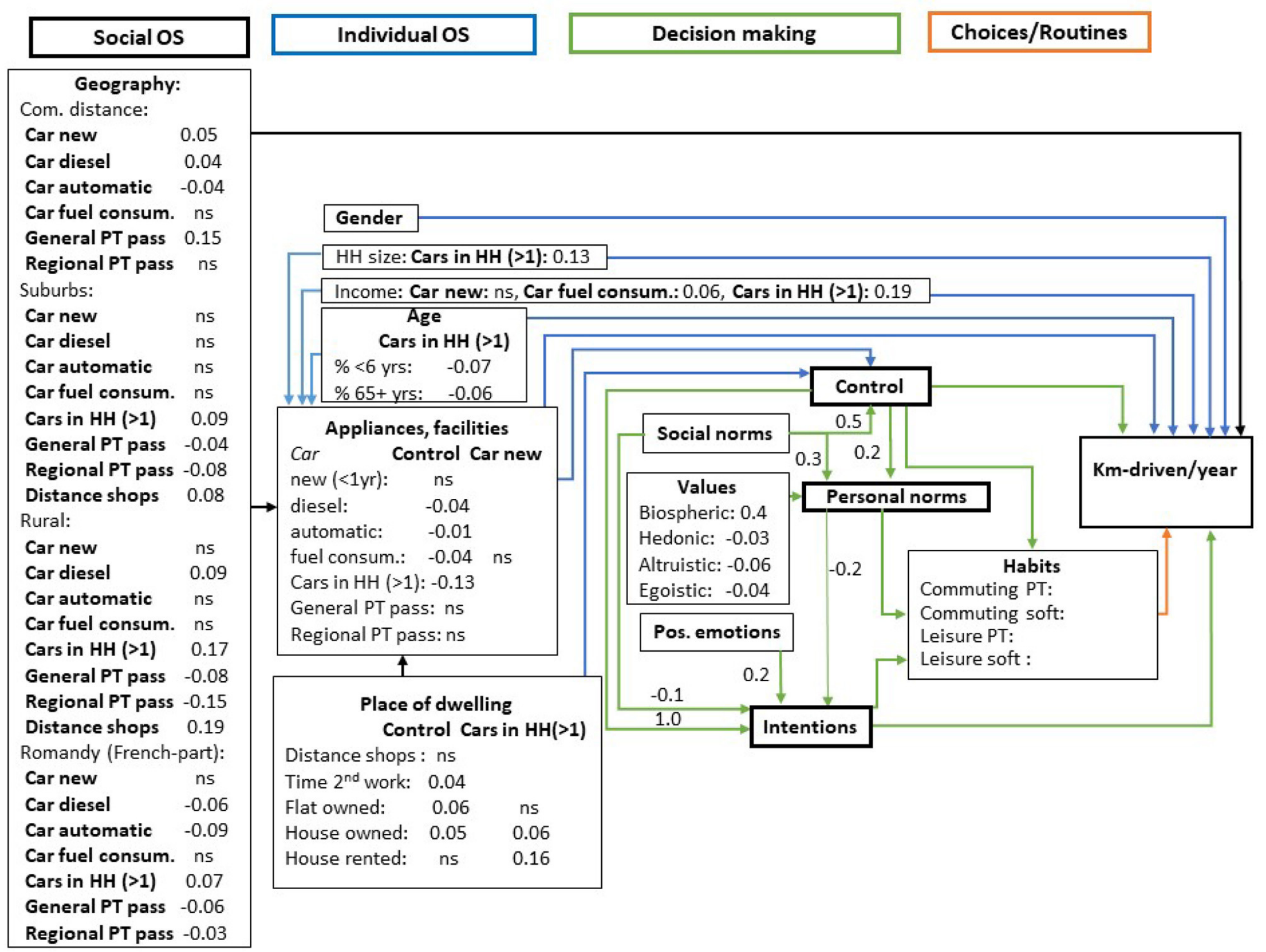


FIGURE 6 | Direct pathways explaining IOS and SOS factors. Reported are standardized direct coefficients with $p < 5\%$ (complete results are provided in **Table A5**), ns = not significant at 5%; bold variables = explained variable.

thus informing the specification of big data analyses on the one hand, and of disciplinary theory-based models on the other. We indeed demonstrate that framework-informed SEMs can draw out different suitable pathways to behavior change, and thus we illustrate a complementary way to direct the analysis of big data. In the following we discuss the empirical findings on annual mileage by private vehicles to point out the related achievements.

Understanding Different Direct and Indirect Pathways to Change Driving Behavior

A strength of using SEMs is that several possible intervention pathways can be simultaneously investigated. Our results show that lower annual mileage by private car is to a large extent explained by habitual use of alternative mode choices for commuting and leisure, such as taking public transport and walking/cycling. Additionally, higher annual mileage by car is related to owning a diesel, automatic or new car, and reduced annual mileage with transport passes. Our results point to a strong influence of habit and structural aspects on mode choices supporting previous empirical findings (e.g., Klöckner and Matthies, 2009; Hess and Schubert, 2019; Punzo et al., 2021). These findings suggest that interventions, focusing solely and separately on, for instance, taxes or information campaigns should focus on different mechanisms and be designed complementing each other, as suggested by others (Bornemann et al., 2018; Urbanek, 2021). For instance, the findings point to a promising hypothesis that could be tested regarding a combined intervention consisting of: (i) breaking unsustainable habits and forming new sustainable transport habits, (ii) measures to discourage car usage, and (iii) structural changes to facilitate commuting with other travel modes.

Furthermore, we find that habitual use of public transport is largely explained by subscription to the right travel means or “equipment”, general or regional passes as well as positive intentions to change behavior. Facilitating the purchase of alternative travel means or free public transport may therefore be a necessary step to increase usage of public transport (Dai et al., 2021). Possible interventions could look at factors related to differences between urban and suburban, rural living and possible region-specific cultural differences (Punzo et al., 2021). Cultural differences between French-speaking and German-speaking parts of Switzerland identified in our analysis are in line with previous findings and show that drivers in the French-speaking region drive longer distances and have a stronger preference for fuel efficient cars (Filippini and Wekhof, 2021). The uptake of regional passes is less well explained by the model but structural factors indicating a link to rural and suburban access are significant predictors. Regional differences are also observed regarding public transport passes, with lower subscription rates in the French-speaking region, in line with results from the Swiss Mobility and Transport Microcensus, which show that HHs in the French-speaking region own more cars, but fewer bikes and much fewer public transport passes than their German-speaking counterparts (SFSO, 2012, 2017). It therefore appears that HHs in the different linguistic regions behave differently regarding their transportation means and their mobility in general.

Factors related to structural preferences, such as owning a diesel car, should simultaneously be addressed in policy interventions. Diesel car ownership is related to structural factors (rural living) and regional differences (i.e., German- vs. French-speaking). However, the proposed model fell short in explaining the underlying determinants of diesel cars and further work is needed here.

Our results also indicate that some of the main disciplinary determinants of driving behavior, previously suggested as intervention or trigger points to change behavior, are either not significant or only indirectly related to private car usage. This finding suggests that some studies may overstate the relevance of the factors conventionally studied within each discipline. This could be the case for some psychological determinants of ECB such as intentions and control (e.g., Klöckner, 2013), albeit the later findings are in line with Fu's (2021) differentiated findings regarding control. Furthermore, this may also be the case for economic determinants such as fuel consumption (Linn, 2013), and HH size, previously documented as important in the transport literature (De Witte et al., 2013). Various reasons could explain the non-significance of a direct relationship between these factors and annual mileage by car. The non-significance of intentions to behave environmentally friendly may be due to a well-documented phenomenon referred to as the intention-behavior gap (e.g., Hassan et al., 2016; Zhang et al., 2019). Similar non-significance has been found in interdisciplinary research modeling car use (Klöckner and Friedrichsmeier, 2011). Ideally, intentions should be collected prior to the behavior, because if intentions (e.g., about reducing car usage) are collected at the same time as the behavior (car usage), as is done in SHEDS, the intentions may not have been implemented yet, unless they were formulated some time prior to the data collection. Longitudinal analysis could help overcome this limitation, and future research should investigate whether the intention-behavior gap remains, when estimating models on panel data.

While the positive effect of fuel efficiency on annual mileage by car (the so-called rebound effect) is a much-debated topic in the economics literature, our estimates yield no statistically significant evidence of such an effect. This can be explained by the strong heterogeneity of individuals in their rebound responses (as outlined in another context by Hediger et al., 2018) but also the theoretical structure imposed by the model. While recommending caution against interpreting this result as non-existence of rebound, we consider that the results point to the importance of moderating effects (here, psychological factors) in understanding rebound behavior. In fact, while many factors might influence both efficiency and the driven distance in a positive manner, factors such as social norms and personal values might lead to higher efficiency and lower usage at the same time.

While recognizing the data limitation on the respondent's knowledge/information, we observe that the model strongly favors policies targeting habit-formation mechanism through setting intentions, increasing control as well as social and personal norms. Furthermore, we find effects from IOS and SOS factors that differ from those previously reported in the transport literature. For example, unlike De Witte et al. (2013), we do not observe a relationship between HH size and distance traveled. Reasons for the non-significance of HH size may be due to the relative magnitude of the effect, meaning that the impact of this factor is relatively small, especially when compared to other factors.

We can summarize our main finding regarding different direct and indirect pathways in four points:

1) The habits/routine pathway shows the most significant impact and stands out as the main mechanism for almost all statistically significant IOS factors.
2) Diesel cars, a main IOS factor, also significantly relates to driving behavior.
3) The SOS factors mainly represent the lowest direct effects, apart from commuting distance, suggesting that mediating factors are important and could change (even reverse) the expected effects.
4) The intention pathway does not represent a significant importance, as shown by the lack of a direct effect of intention on behavior.

Overall, our findings highlight the usefulness of applying SEM to understand complex phenomena and to draw out which pathways would be most suitable for interventions. Our findings also show the importance of interdisciplinary models to provide a solid structure to analyse the complexity of factors (here in shaping driving behavior) and to shed light on the explanatory strengths of the factors and their interplay.

Data Limitations

The results from our illustrative case indicate that the application of SEM may help understand complex phenomena and bridge the gap between predictive big data models and less flexible regression models. While consistent with its theoretical model counterpart, the proposed empirical model is reduced, hence more tractable in certain dimensions. This has proved inevitable for models applied to survey data, mainly because of data availability that does not always match the model's requirements. We concede that the gap between the theoretical model and its empirical counterpart expresses the tension between an ideal model with maximum explanatory power and its empirical applicability.

Developing an *ideal* model, even if datasets can be expected to be suboptimal in most cases, is a relevant task, as it can also set data requirements for future research. In our context, the ideal dataset would include all factors on all levels (SOS, IOS, decision-making and routines/choices). Our current dataset misses some individual level factors, such as attitudes and social context, which could be collected through HH surveys. Importantly SOS data were also missing, such as higher order data on institutional policies and norms, weather and geographical information, technology and economy. Collecting SOS data could be time-consuming and the difficulty would lie in abstracting from the individual HH level. In order to merge individual factors (at the HH level) with social factors (available from various other sources), it is extremely important to collect information about the location of HHs and their work places. Finally, although theoretically possible, it might be practically impossible to gather an ideal dataset via surveys due to financial and time limitations, let alone participants' willingness to fill in very long and detailed surveys which might also lead to an increase in errors and answer biases. Thus, we suggest that a fruitful avenue of research is the exploration of strategies that rely on new communication technologies (i.e., apps, sensors, etc.) and can be linked to revealed preferences data.

CONCLUSION

Our empirical exercise highlights the usefulness of applying SEM to understand complex phenomena such as energy consumption behavior (or, more precisely, annual mileage by private vehicle in this case) and to identify suitable pathways to change behavior. It also further highlights the importance of conducting interdisciplinary research with models considering a broad range of potential predictors as opposed to models rooted in a single discipline. In our estimation, a number of otherwise significant factors have become non-significant, sometimes transgressing disciplinary context matters. The exercise of fitting data to such an interdisciplinary model focuses attention onto what the ideal data would look like, and on potential issues with collecting large sets of survey data. Nevertheless, despite data shortcomings and deviations from the ideal interdisciplinary model, our empirical model delivers relevant insights on determinants of annual mileage by Swiss HHs using private cars.

Our estimated model points to a number of mechanisms that can be targeted for reducing private car usage and increasing use of alternative modes of transport. Our policy relevant conclusions point to the importance of:

1. Promoting "habitual" alternative mode choice use for commuting and leisure.
2. Supporting suitable personal infrastructure changes such as public transport passes.
3. Discouraging the purchase of diesel cars.

Depending on the relative importance of each pathway, we can identify which mechanism should be prioritized for the greatest impact. This is an empirical question that can be addressed by holistic models such as this IHECF framework-informed SEM. The illustration presented in the paper shows that SEM can be used to effectively assess the relative importance of different direct and indirect pathways.

Findings in this study illustrate how SEM studies can be brought into a sustainametrics conversation. While this conversation tends to focus on how increasingly available data can support transitions to sustainable societies, our study directs attention to limitations inherent to predictive models based on big data and their role in supporting sustainability transitions. Furthermore, our study illustrates how big data analysis can be complemented by SEM analysis on available data—not ideal data but available ones. In particular, we argue that SEM analyses can support fine-tuning of policy interventions informed by predictive analysis relying on big data. While predictive methods relying on big data can be used to estimate the impact of interventions and identify the most reactive segments of population—"low-hanging fruits"—, SEM analysis can inform the design of intervention policies by focusing on specific and multiple mechanisms. For instance, certain machine learning frameworks can be used to predict sustainability-relevant individual behaviors based on readily available HH characteristics thus identifying relevant target groups for policy interventions. However, they cannot help defining a mechanism to prioritize various alternatives. A SEM analysis such as the one proposed in this paper can provide information about the relative importance of specific pathways based on multidisciplinary models. These pathways can be used to identify the mechanisms that should be targeted by policy interventions and to design targeted policy interventions on specific responses in relevant segments of population. In a certain sense, SEM stand in between regression and predictive analyses and could hence be used to bridge the gap between the two types of analyses.

DATA AVAILABILITY STATEMENT

The SHEDS data is available to academic researchers after one year of the launch of the respective wave. Researchers have to sign a confidentiality agreement and submit a short research proposal to indicate the intended use of the data. Data use for commercial purposes is not allowed. For more information on the SHEDS survey see https://www.sccer-crest.ch/research/swiss-household-energy-demand-survey-sheds/.

ETHICS STATEMENT

Ethical review and approval was not required for this type of study with human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors were jointly responsible for the conceptualization and the writing of the paper, with IS taking the lead. IS and SW were jointly responsible for the empirical analysis, which they interpreted with the collaboration of ALMC and MF. PB, MF, SW, and IS are responsible for the implementation of the Swiss Household Energy Demand Survey (SHEDS), with SW taking the lead. All authors contributed to the article and approved the submitted version.

FUNDING

This research project was part of the Swiss Competence Center for Energy Research SCCER CREST which was financially supported by the Swiss Innovation Agency Innosuisse under Grant No. KTI. 1155000154.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsus.2022.837427/full#supplementary-material

REFERENCES

Ababio-Donkor, A., Saleh, W., and Fonzone, A. (2020). The role of personal norms in the choice of mode for commuting. *Res. Transport. Econ.* 83, 100966. doi: 10.1016/j.retrec.2020.100966

Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-T

Ajzen, I. (2006). *Constructing a TPB Questionnaire: Conceptual and Methodological Considerations*. Available online at: https://www.researchgate.net/publication/235913732_Constructing_a_Theory_of_Planned_Behavior_Questionnaire/link/56f00f4508aeae9f93e804b6/download (accessed January 29, 2022).

Alberini, A., Bareit, M., Filippini, M., and Martinez-Cruz, A. L. (2018). The impact of emissions-based taxes on the retirement of used and inefficient vehicles: the case of Switzerland. *J. Environ. Econ. Manage.* 88, 234–258. doi: 10.1016/j.jeem.2017.12.004

Allcott, H. (2011). Social norms and energy conservation. *J. Public Econ.* 95, 1082–1095. doi: 10.1016/j.jpubeco.2011.03.003

Allcott, H., and Greenstone, M. (2012). Is there an energy efficiency gap? *J. Econ. Perspect.* 26, 3–28. doi: 10.1257/jep.26.1.3

Allcott, H., and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation. *Am. Econ. Rev.* 104, 3003–3037. doi: 10.1257/aer.104.10.3003

Azevedo, I. M. (2014). Consumer end-use energy efficiency and rebound effects. *Annu. Rev. Environ. Resour.* 39, 393–418. doi: 10.1146/annurev-environ-021913-153558

Bandura, A. (2001). Social cognitive theory: an agentic perspective. *Annu. Rev. Psychol.* 52, 1–26. doi: 10.1146/annurev.psych.52.1.1

Bollen, K. A., and Pearl, J. (2013). "Eight myths about causality and structural equation models," in *Handbook of Causal Analysis for Social Research*, editor S. L. Morgan (Dordrecht: Springer), 301–328.

Borenstein, S. (2015). A microeconomic framework for evaluating energy efficiency rebound and some implications. *Energy J.* 36, 1–21. doi: 10.5547/01956574.36.1.1

Bornemann, B., Sohre, A., and Burger, P. (2018). Future governance of individual energy consumption behavior change – A framework for reflexive designs. *Energy Res. Soc. Sci.* 35, 140–151. doi: 10.1016/j.erss.2017.10.040

Bouman, T., van der Werff, E., Perlaviciute, G., and Steg, L. (2021). Environmental values and identities at the personal and group level. *Curr. Opin. Behav. Sci.* 42, 47–53. doi: 10.1016/j.cobeha.2021.02.022

Bouscasse, H., and Bonnel, P. (2016). "Socio-psychological determinants of mode choice habits" in *Working Papers 2016-05*. Grenoble Applied Economics Laboratory (GAEL). Available online at: https://www.researchgate.net/profile/Helene-Bouscasse/publication/312371964_Socio-psychological_Determinants_of_Mode_Choice_Habits/links/587bbf9708ae9a860fe9b8e0/Socio-psychological-Determinants-of-Mode-Choice-Habits.pdf (accessed January 29, 2022),

Brännlund, R., and Vesterberg, M. (2021). Peak and off-peak demand for electricity: is there a potential for load shifting? *Energy Econ.* 102, 105466. doi: 10.1016/j.eneco.2021.105466

Brosch, T. (2021). Affect and emotions as drivers of climate change perception and action: a review. *Curr. Opin. Behav. Sci.* 42, 15–21. doi: 10.1016/j.cobeha.2021.02.001

Brosch, T., and Sander, D. (2014). Appraising value: the role of universal core values and emotions in decision-making. *Cortex* 59, 203–205. doi: 10.1016/j.cortex.2014.03.012

Brounen, D., Kok, N., and Quigley, J. M. (2012). Residential energy use and conservation: economics and demographics. *Eur. Econ. Rev.* 56, 931–945. doi: 10.1016/j.euroecorev.2012.02.007

Burger, P., Bezençon, V., Bornemann, B., Brosch, T., Carabias-Hütter, V., Farsi, M., et al. (2015). Advances in understanding energy consumption behavior and the governance of its change–outline of an integrated framework. *Front. Energy Res.* 3, 29. doi: 10.3389/fenrg.2015.00029

Buylova, A. (2020). Spotlight on energy efficiency in Oregon: investigating dynamics between energy use and socio-demographic characteristics in spatial modeling of residential energy consumption. *Energy Policy* 140, 111439. doi: 10.1016/j.enpol.2020.111439

Cattaneo, C. (2019). Internal and external barriers to energy efficiency: which role for policy interventions? *Energy Efficiency* 12, 1–19. doi: 10.1007/s12053-019-09775-1

Dai, J., Liu, Z., and Li, R. (2021). Improving the subway attraction for the post-COVID-19 era: the role of fare-free public transport policy. *Transport Policy* 103, 21–30. doi: 10.1016/j.tranpol.2021.01.007

De Borger, B., Mulalic, I., and Rouwendal, J. (2016). Measuring the rebound effect with micro data: a first difference approach. *J. Environ. Econ. Manage.* 79, 1–17, doi: 10.1016/j.jeem.2016.04.002

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., and Macharis, C. (2013). Linking modal choice to motility: a comprehensive review. *Transport. Res. Part A Policy Pract.* 49, 329–341. doi: 10.1016/j.tra.2013.01.009

del Mar Solà, M., de Ayala, A., Galarraga, I., and Escapa, M. (2021). Promoting energy efficiency at household level: a literature review. *Energy Efficiency* 14, 1–22. doi: 10.1007/s12053-020-09918-9

Eurostat (2021). *Energy Statistics – An Overview*. Available online at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_statistics_-_an_overview#Energy_intensity (accessed January 29, 2022).

Filippini, M., and Wekhof, T. (2021). The effect of culture on energy efficient vehicle ownership. *J. Environ. Econ. Manage.* 105, 102400. doi: 10.1016/j.jeem.2020.102400

Fowlie, M., Greenstone, M., and Wolfram, C. (2018). Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *Q. J. Econ.* 133, 1597–1644. doi: 10.1093/qje/qjy005

Fu, X. (2021). A novel perspective to enhance the role of TPB in predicting green travel: the moderation of affective-cognitive congruence of attitudes. *Transportation* 48, 3013–3035. doi: 10.1007/s11116-020-10153-5

Giulio, A. D., Fischer, D., Schäfer, M., and Blättel-Mink, B. (2014). Conceptualizing sustainable consumption: toward an integrative framework. *Sustain. Sci. Pract. Policy* 10, 45–61. doi: 10.1080/15487733.2014.11908124

Götschi, T., de Nazelle, A., Brand, C., and Gerike, R. (2017). Towards a comprehensive conceptual framework of active travel behavior: a review and synthesis of published frameworks. *Curr. Environ. Health Rep.* 4, 286–295. doi: 10.1007/s40572-017-0149-9

Hassan, L. M., Shiu, E., and Shaw, D. (2016). Who says there is an intention–behaviour gap? Assessing the empirical evidence of an intention–behaviour gap in ethical consumption. *J. Bus. Ethics* 136, 219–236. doi: 10.1007/s10551-014-2440-0

Havranek, T., and Kokes, O. (2015). Income elasticity of gasoline demand: a meta-analysis. *Energy Econ.* 47, 77–86. doi: 10.1016/j.eneco.2014.11.004

Hediger, C., Farsi, M., and Weber, S. (2018). Turn it up and open the window: on the rebound effects in residential heating. *Ecol. Econ.* 149, 21–39. doi: 10.1016/j.ecolecon.2018.02.006

Hess, A. K., Samuel, R., and Burger, P. (2018). Informing a social practice theory framework with social-psychological factors for analyzing routinized energy consumption: a multivariate analysis of three practices. *Energy Res. Soc. Sci.* 46, 183–193. doi: 10.1016/j.erss.2018.06.012

Hess, A. K., and Schubert, I. (2019). Functional perceptions, barriers, and demographics concerning e-cargo bike sharing in Switzerland. *Transport. Res. Part D Transport Environ.* 71, 153–168. doi: 10.1016/j.trd.2018.12.013

Hiratsuka, J., Perlaviciute, G., and Steg, L. (2018). Testing VBN theory in Japan: relationships between values, beliefs, norms, and acceptability and expected effects of a car pricing policy. *Transport. Res. Part F Traffic Psychol. Behav.* 53, 74–83. doi: 10.1016/j.trf.2017.12.015

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Eq. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *Am. Econ. Rev.* 93, 1449–1475. doi: 10.1257/000282803322655392

Karatasou, S., Laskari, M., and Santamouris, M. (2018). Determinants of high electricity use and high energy consumption for space and water heating in European social housing: socio-demographics and building characteristics. *Energy Build.* 170, 107–114. doi: 10.1016/j.enbuild.2018.04.019

Karimu, A., Krishnamurthy, C. K. B., and Vesterberg, M. (2022). Understanding hourly electricity demand: implications for load, welfare and emissions. *Energy J.* 43, 161–189. doi: 10.5547/01956574.43.1.akar

Kavousian, A., Rajagopal, R., and Fischer, M. (2013). Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy* 55, 184–194. doi: 10.1016/j.energy.2013.03.086

Kent, J. L. (2021). The use of practice theory in transport research. *Transport Rev.* 1–23. doi: 10.1080/01441647.2021.1961918

Klöckner, C. A. (2013). A comprehensive model of the psychology of environmental behaviour-A meta-analysis. *Glob. Environ. Change* 23, 1028–1038. doi: 10.1016/j.gloenvcha.2013.05.014

Klöckner, C. A., and Blöbaum, A. (2010). A comprehensive action determination model: toward a broader understanding of ecological behaviour using the example of travel mode choice. *J. Environ. Psychol.* 30, 574–586. doi: 10.1016/j.jenvp.2010.03.001

Klöckner, C. A., and Friedrichsmeier, T. (2011). A multi-level approach to travel mode choice–How person characteristics and situation specific aspects determine car use in a student sample. *Transport. Res. Part F Traffic Psychol. Behav.* 14, 261–277. doi: 10.1016/j.trf.2011.01.006

Klöckner, C. A., and Matthies, E. (2009). Structural modeling of car use on the way to the university in different settings: interplay of norms, habits, situational restraints, and perceived behavioral control 1. *J. Appl. Soc. Psychol.* 39, 1807–1834. doi: 10.1111/j.1559-1816.2009.00505.x

Koszowski, C., Gerike, R., Hubrich, S., Götschi, T., Pohle, M., and Wittwer, R. (2019). "Active mobility: bringing together transport planning, urban planning, and public health," in *Towards User-Centric Transport in Europe*, eds B. Müller and G. Meyer (Cham: Springer). p. 149–171. doi: 10.1007/978-3-319-99756-8_11

Kurz, T., Gardner, B., Verplanken, B., and Abraham, C. (2015). Habitual behaviors or patterns of practice? Explaining and changing repetitive climate-relevant actions. *Clim. Change* 6, 113–128. doi: 10.1002/wcc.327

Labandeira, X., Labeaga, J. M., and López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy* 102, 549–568. doi: 10.1016/j.enpol.2017.01.002

Liddle, B. (2014). Impact of population, age structure, and urbanization on carbon emissions/energy consumption: evidence from macro-level, cross-country analyses. *Popul. Environ.* 35, 286–304. doi: 10.1007/s11111-013-0198-4

Linn, J. (2013). "The rebound effect for passenger vehicles," in *Discussion Paper 13-19-REV, Resources for the Future*. Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2292007 (accessed January 29, 2022).

Marien, H., Custers, R., and Aarts, H. (2018). "Understanding the formation of human habits: an analysis of mechanisms of habitual behaviour," in *The Psychology of Habit*, editor B. Verplanken (Cham: Springer).

Marsh, H., Hau, K., and Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999). Findings. *Struct. Eq. Model. Multidiscip. J.* 11, 320–341. doi: 10.1207/s15328007sem1103_2

McLoughlin, F., Duffy, A., and Conlon, M. (2012). Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. *Energy Build.* 48, 240–248. doi: 10.1016/j.enbuild.2012.01.037

Mochmann, I. C., and El-Menouar, Y. (2005). Lifestyle groups, social milieus and party preference in Eastern and Western Germany: theoretical considerations and empirical results. *Ger. Polit.* 14, 417–437. doi: 10.1080/09644000500340594

Moro, A., and Holzer, A. (2020). A framework to predict consumption sustainability levels of individuals. *Sustainability* 12, 1–26. doi: 10.3390/su12041423

Nayum, A., and Klöckner, C. A. (2014). A comprehensive socio-psychological approach to car type choice. *J. Environ. Psychol.* 40, 401–411. doi: 10.1016/j.jenvp.2014.10.001

Oreg, S., and Katz-Gerro, T. (2006). Predicting proenvironmental behavior cross-nationally: values, the theory of planned behavior, and value-belief-norm theory. *Environ. Behav.* 38, 462–483. doi: 10.1177/0013916505286012

Ostrom, E. (2014). Collective action and the evolution of social norms. *J. Natural Resources Policy Res.* 6, 235–252. doi: 10.1080/19390459.2014.935173

Pearl, J. (2012). "The causal foundations of structural equation modeling," in *Handbook of Structural Equation Modeling*, Chapter 5, editor R. H. Hoyle (Guilford press), 68–91.

Pothitou, M., Hanna, R. F., and Chalvatzis, K. J. (2016). Environmental knowledge, pro-environmental behaviour and energy savings in households: an empirical study. *Appl. Energy* 184, 1217–1229. doi: 10.1016/j.apenergy.2016.06.017

Punzo, G., Panarello, D., and Castellano, R. (2021). Sustainable urban mobility: Evidence from three developed European countries. *Qual. Quant.* doi: 10.1007/s11135-021-01253-0

Schubert, I., de Groot, J. I., and Newton, A. C. (2021). Challenging the Status Quo through social influence: changes in sustainable consumption through the influence of social networks. *Sustainability* 13, 5513. doi: 10.3390/su13105513

Schubert, I., Sohre, A., and Ströbel, M. (2020). The role of lifestyle, quality of life preferences and geographical context in personal air travel. *J. Sustain. Tour.* 28, 1519–1550. doi: 10.1080/09669582.2020.1745214

SFOE (2020). *Schweizerische Gesamtenergiestatistik 2017*. Swiss Federal Office of Energy, Bern. Available online at: https://www.bfe.admin.ch/bfe/en/home/versorgung/statistik-und-geodaten/energiestatistiken/gesamtenergiestatistik.exturl.html/aHR0cHM6Ly9wdWJkYi5iZmUuYWRtaW4uY2gvZGUvcHVibGljYX/Rpb24vZG93bmxvYWQvMTA1Mzc=.html (accessed January 29, 2022).

SFSO (2012). *La mobilité en Suisse: Résultats du microrecensement mobilité et transports 2010*. Available online at: https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/publications.assetdetail.348722.html (accessed January 29, 2022).

SFSO (2017). *Comportement de la population en matière de transports 2015: Principaux résultats du microrecensement mobilité et transports*. Available online at: https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/publications.assetdetail.1840605.html (accessed January 29, 2022).

Shove, E. (2021). Time to rethink energy research. *Nat. Energy* 6, 118–120. doi: 10.1038/s41560-020-00739-9

Spaargaren, G. (2003). Sustainable consumption: a theoretical and environmental policy perspective. *Soc. Natural Resources* 16, 687–701. doi: 10.1080/08941920309192

Stapleton, L., Sorrell, S.t., and Schwanen, T. (2016). Estimating direct rebound effects for personal automotive travel in Great Britain. *Energy Econ.* 54, 313–325. doi: 10.1016/j.eneco.2015.12.012

Stephenson, J. (2018). Sustainability cultures and energy research: an actor-centred interpretation of cultural theory. *Energy Res. Soc. Sci.* 44, 242–249. doi: 10.1016/j.erss.2018.05.034

Stephenson, J., Barton, B., Carrington, G., Doering, A., Ford, R., Hopkins, D., et al. (2015). The energy cultures framework: exploring the role of norms, practices and material culture in shaping energy behaviour in New Zealand. *Energy Res. Soc. Sci.* 7, 117–123. doi: 10.1016/j.erss.2015.03.005

Stern, P. C. (2000). New environmental theories: toward a coherent theory of environmentally significant behavior. *J. Soc. Issues* 56, 407–424. doi: 10.1111/0022-4537.00175

Stoppok, M., Jess, A., Freitag, R., and Alber, E. (2018). Culture, consumption and cost: a comparative analysis of household energy consumption in Kenya, Germany and Spain. *Energy Res. Soc. Sci.* 40, 127–139. doi: 10.1016/j.erss.2017.12.004

Sun, P. C., Wang, H. M., Huang, H. L., and Ho, C. W. (2018). Consumer attitude and purchase intention toward rooftop photovoltaic installation: the roles of personal trait, psychological benefit, and government incentives. *Energy Environ*. 31:0958305X1775427. doi: 10.1177/0958305X17754278

Sütterlin, B., Brunner, T. A., and Siegrist, M. (2011). Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics. *Energy Policy* 39, 8137–8152. doi: 10.1016/j.enpol.2011.10.008

Tan, C. S., Ooi, H. Y., and Goh, Y. N. (2017). A moral extension of the theory of planned behavior to predict consumers' purchase intention for energy-efficient household appliances in Malaysia. *Energy Policy* 107, 459–471. doi: 10.1016/j.enpol.2017.05.027

Tomarken, A., and Waller, N. (2005). Structural equation modeling: strengths, limitations, and misconceptions. *Annu. Rev. Clin. Psychol.* 1, 31–65. doi: 10.1146/annurev.clinpsy.1.102803.144239

Tversky, A., and Kahneman, D. (1974). Judgements under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124

Tweed, C., Humes, N., and Zapata-Lancaster, G. (2015). The changing landscape of thermal experience and warmth in older people's dwellings. *Energy Policy* 84, 223–232. doi: 10.1016/j.enpol.2015.03.011

Urbanek, A. (2021). Potential of modal shift from private cars to public transport: A survey on the commuters' attitudes and willingness to switch-A case study of Silesia Province, Poland. *Res. Transport Econom*. 85, 101008. doi: 10.1016/j.retrec.2020.101008

Van Acker, V., Mokhtarian, P. L., and Witlox, F. (2014). Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy* 35, 88–99. doi: 10.1016/j.tranpol.2014.006

Verplanken, B., and Aarts, H. (1999). Habit, attitude, and planned behaviour: is habit an empty construct or an interesting case of goal-directed automaticity? *Eur. Rev. Soc. Psychol.* 10, 101–134. doi: 10.1080/14792779943000035

Weber, S., Burger, P., Farsi, M., Martinez-Cruz, A. L., Puntiroli, M., Schubert, I., et al. (2017). "Swiss Household Energy Demand Survey (SHEDS): Goals, design, and implementation," in *SCCER CREST Working Paper WP2 - 2017/04*. Available online at: https://www.sccer-crest.ch/fileadmin/FILES/Datenbank_Personen_Projekte_Publikationen/Publications/Working_Papers/Work_Package_2/Weber_Burger_et_al_2017_SHEDS_Official_description.pdf (accessed January 29, 2022).

Winkler, R., Almer, C., Bader, C. C., Gonseth, C., Laurent-Lucchetti, J., Thalmann, P., et al. (2014). "Energy consumption of buildings – Direct impacts of a warming climate and rebound effects," in *CH2014-Impacts, Toward Quantitative Scenarios of Climate Change Impacts in Switzerland*. OCCR, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim, Bern, Switzerland, 99–105.

Zhang, B., Lai, K. H., Wang, B., and Wang, Z. (2019). From intention to action: how do personal attitudes, facilities accessibility, and government stimulus matter for household waste sorting? *J. Environ. Manage.* 233, 447–458. doi: 10.1016/j.jenvman.2018.12.059

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Sustainability

ORIGINAL RESEARCH
published: 25 March 2022
doi: 10.3389/frsus.2022.834849

ChemESI (Chemical Environmental Sustainability Index)—A KPI (Key Performance Indicator) for Standardizing Environmental Metrics for Chemical Sustainability

*Hans Plugge**

Safer Chemical Analytics LLC, Columbia, MD, United States

As part of the ESG (Environmental, Social, and Governance) dynamics, chemical environmental sustainability, i.e., the impact from chemicals and how to measure it across hundreds of products, is a daunting, but necessary task. Although methods are available to measure the effects and impacts of a single chemical, most enterprises do not focus on a single chemical, let alone produce "pure" chemicals for sale. Nearly all chemicals in commerce are chemical products, i.e., mixtures, while assessment methods for mixtures are few and far between. What is needed is a metric that tracks the potential risk of an enterprise's total product inventory while monitoring its improvements as it greens in coming years. The Chemical Environmental Sustainability Index (ChemESI) metric measures both risk (as the product of exposure and hazard) and hazard across numerous chemicals as a single metric/KPI. The ChemESI's for chemicals, products and facilities are expressed such that they can be summed across facilities to wrap up into a single corporate ChemESI KPI for either hazard or risk. But what about growth—if growth occurs using greener chemical inventory, a company can both grow and improves its ChemESI KPI. Given the lead-time for developing more, true green alternatives to existing chemicals, intermediate substitution of less hazardous "analogs" may drive initial ChemESI reduction. To achieve a representative risk estimate, a primary data need is chemical characterization data for products. SDS's (Safety Data Sheets) unfortunately make poor substitutes for true constituent analyses. A definite need exists for better, more detailed chemical characterization data for both mixtures and individual chemicals, as the latter most often are not 100 % "pure." However, these SDS's are available universally, across the globe, and provide GHS (UN Globally Harmonized System) classifications for single endpoint chemical hazard assessment. GHS classifications are near universal and here used to derive chemical hazard scores over multiple endpoints for each chemical. A ChemESI Risk metric can be derived by multiplying the hazard score by the exposure, here using inventory as a surrogate. The ChemESI Hazard metric is then derived as an inventory weighted hazard score.

Keywords: chemical sustainability, sustainability metrics, sustainability index, chemical products, KPI, GHS, chemical inventory, SDS

OPEN ACCESS

Edited by:
Shutaro Takeda,
Kyoto University, Japan

Reviewed by:
Li Li,
University of Nevada, Reno, United States
Daikichi Seki,
Kyoto University, Japan

Correspondence:
Hans Plugge
hans.plugge@saferchemicalanalyticsllc.com

Specialty section:
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

Received: *13 December 2021*
Accepted: *21 February 2022*
Published: *25 March 2022*

Citation:
Plugge H (2022) ChemESI (Chemical Environmental Sustainability Index)—A KPI (Key Performance Indicator) for Standardizing Environmental Metrics for Chemical Sustainability. Front. Sustain. 3:834849. doi: 10.3389/frsus.2022.834849

INTRODUCTION

Environmental sustainability of chemical inventory is an often-overlooked aspect of ESG, where environmental impacts often reference greenhouse gases or decarbonization (Investopedia, 2021). Enterprise chemical sustainability gets short thrift in nearly all assessments systems, except perhaps TRACI (Tool for Reduction and Assessment of Chemicals and Other Environmental Impacts, e.g., Bare, 2002). Recently systems have been developed to assess sustainability in procurement of health products (UNDP, 2020; SHiPP, 2021). Given that most companies maintain an inventory of chemicals whether it be in large or small quantities, often in an astounding variety, environmental sustainability metrics are needed. Here we propose one ESG metric, the Chemical Environmental Sustainability Index or ChemESI for short. ChemESI aims to measure the inherent risk of chemical inventory across all locations within an enterprise, allowing for identification of locations with increased hazard and risk within the enterprise. Note that it produces both hazard and risk estimates, the latter incorporating the exposure to allow one to distinguish between a test tube and a tanker trailer of a chemical.

Environmental and health effects of chemicals within a value chain are managed through Product Stewardship (Hart, 2018). Product stewardship of location inventory often focuses on the inherent hazard of chemicals classified using the eighth or ninth edition of GHS (Globally Harmonized System) information within SDS's (Safety Data Sheets) (UN, 2021). GHS thus perpetuates the focus on hazard only. Several chemical regulations, including REACH (Registration, Evaluation, Authorization and Restriction of Chemicals, ECHA, 2017), have tried to address this *via* individual exposure limits based on acceptable risk. The definition of acceptable risk there is fuzzied once again by the use of ballpark uncertainty factors (USEPA, 1993; ECHA, 2017) as compared to data analytics-based factors. These so-called safety/uncertainty factors (better referred to as assessment factors), often of multiple orders of magnitude, are used to account for perceived uncertainties in regulatory risk assessments. Such factors have been shown to be overestimations in those cases where they have been calculated from data analytics (Escher et al., 2020) rather than taken off the proverbial back of an envelope. Health and environmental hazard data also have their own inherent uncertainties based on a comparison of tests for the same chemical (Pham et al., 2019; Kostal et al., 2020; Plugge et al., 2021), often again in the (multiple) order of magnitude range. In-depth, regulatory chemical-specific risk assessments are thus inherently uncertain and available for only a limited number of chemicals. They are also costly: regulatory, single chemical risk assessments without data acquisition often exceed $1,000,000 (Maertens and Plugge, 2018).

While GHS provides a good classification method for chemicals and their associated hazard, it does not account for exposure, especially from a corporate risk inventory/insurance point of view (Chemsec, 2021). Risk is what determines an enterprise's liability/insurance "exposure" from its chemical inventory, hazard is just one of the factors. Risk and exposure are used here as commonly used within the environmental health community, which is incongruent with definitions in other communities. In the insurance community; Risk = Uncertainty arising from the possible occurrence of given events or the actual property/insured (IRMI, 2022a). Similarly: Exposure = The state of being subject to loss because of some hazard or contingency, also used as a measure of the rating units or the premium base of a risk (IRMI, 2022b). The environmental health definition of Risk= hazard * exposure is thus not congruent with insurance usage. Similarly for the financial definition of risk (Investopedia, 2022).

Hazard assessment now nearly universally starts with the (ninth edition of the) GHS (UN, 2021). REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) regulations (ECHA, 2017) offer a similar but not quite congruent system of classification. Not always acknowledged, GHS is also neither completely global nor harmonized. Although negotiated at the UN level, most countries provide their own (small) spin on, or lag in adoption of the latest edition. Nonetheless, these GHS hazard classes are more universal than any other system and chemical specific, single endpoint classifications are widely available commercially (Scivera, 2021; Toxnot, 2021; Verisk 3E, 2021). GHS assessments however still have drawbacks most of which were designed into the system:

- GHS classification is always based on the most sensitive, i.e., hazardous datapoint available, with, in practice, little consideration of data quality and variability; REACH (ECHA, 2017) being a partial exception
- GHS classification bins into rather broad categories, especially for acute toxicity, e.g., a datapoint of 49 vs. 51 can result in classification in different toxicity categories (Kostal et al., 2020)
- GHS only notes absence of data and does not penalize for missing toxicity data
- Classification use/interpolation of data from and/or incorporation of NAMs (New Approach Methodologies, USEPA, 2021) often lags
- GHS compliant SDS's (Safety Data Sheet) often lack a full chemical characterization, i.e., the ingredient percentages do not add up to (or exceed) 100%, resulting in a less detailed GHS classification of a product.

GHS-based SDS's were developed for the purpose of occupational hazard communication for transport of hazardous chemicals, which accounts for some of these limitations. Notwithstanding these weaknesses, GHS is accepted in nearly all countries around the world as "the" hazard assessment/classification system allowing the preparation of globally accepted SDS's with mostly universal classifications, as evidenced by its proliferation in EHS (Environmental Health and Safety) systems (e.g., Verisk 3E, 2021). GHS provides single endpoint scores which are not summable into one score for a chemical, let alone products, e.g., one cannot compare or weight highly acutely toxic *via* oral route with skin sensitization. ChemESI transforms the GHS classification scores into an all endpoints combined total

Chemical Hazard Score (CHS) for each chemical. The hazard portion of the ChemESI is thus based on GHS classifications, while noting its drawbacks, i.e., generally resulting in an overestimation of hazard, although in a non-quantifiable manner.

Risk, as described above, is a more accurate measure of chemical sustainability: after all it accounts for the exposure in addition to the inherent hazard of chemicals. Risk can then be quantified, as always [in (environmental) toxicology/risk assessment], as the product of exposure and hazard: risk = exposure * hazard. GHS classifications act as surrogate hazard quantifiers *via* the CHS. Here exposure estimates for a facility are approximated as quantities held within a facility, i.e., using inventory as a surrogate for exposure. A similar approach was used to prioritize chemical risk for reproductive effects within the EU: quantity produced times REACH based DNEL (Derived No-Effect Level) estimates, prioritized chemicals for assessment (Risk Policy Analysts, 2019). ChemESI Risk is then calculated as the product of the two, i.e., risk = exposure * hazard summed over all chemicals or products. ChemESI Hazard is calculated as total risk divided by total exposure, i.e., a weighted Hazard Score across a facility, while maintaining individual chemical/product hazard and risk scores.

Exposure models exist for all kinds of media and exposure types. Often chemicals exert effects through a variety of routes, i.e., air, skin, oral, and for biota effects through environmental effects mostly in water. Persistence also plays a large role. Note that all of these routes require their own modeling "software" with specific data requirements and calculations to produce quantitative estimates with some certainty. Although feasible for a few chemicals, this process becomes unmanageable for a large variety of chemicals as well as products. High throughput modeling can overcome this singular exception but also adds high variability (Li et al., 2021). One can model single chemical concentrations reasonably well, but model outputs suffer from fairly high uncertainty intervals which, when modeled in detail for multichemical assessments, exceed 4 orders of magnitude (Li et al., 2021). Mixture interactions are rarely described and then generally only for pair-wise comparisons.

The ChemESI Risk model used here assesses impacts from multiple chemicals in multiple media. As mentioned above exposure modeling for such a scenario would be a nightmare. Other risk assessment methodologies have tried doing this: predict environmental concentrations and calculate risk (e.g., Arnot and Mackay, 2008; Arnot et al., 2012). These are not traditional exposure/risk assessments—they employ high-throughput exposure scenarios to derive risks. Such schemes employ numerous assumptions of toxicity and persistence to derive environmental concentrations coupled with risk estimates. Initially performed for environmental pollutants, risk was used a function of quantity, toxicity, bioaccumulation, and persistence in a model to look at a continuous distribution to prioritize chemicals for regulation instead of the rigid yes/no cutoff model. Exposure estimates include an estimated emission rate based on quantity produced. Here risk needed to be calibrated by actual (rather than estimated) toxicity values and calibrated by actual exposure measurements. Further modeling of environmental risk resulted in a more than 4 orders of magnitude 95% confidence interval, leaving the authors to conclude that quantity may need to be the regulatory driver at this time (Li et al., 2021). Correlating known emission rates with environmental monitoring data demonstrated that existing exposure models could be off by an order of magnitude even when calibrated with actual release factors (Spaniol et al., 2021). Review of risk assessment applications used simplified exposure models in a risk assessment of transformation products in an aquatic environment (Escher and Fenner, 2011). Risk assessment of all transformation products was found to be out of reach. Tiered approaches were recommended based off parent compounds' risk assessments using advanced simulation methods which are now (2022) becoming more available. Specific modes of action would need to be identified for such risk assessments to proceed.

Exposure modeling always starts off with total quantity followed by a mathematical approach using various surrogate/estimated parameters. Given the observed uncertainties in the final exposure estimate, it was acknowledged that in a regulatory approach, quantity may be the best approximation for effect concentration (Li et al., 2021). Regulation of chemicals will thus remain focused on the total quantity, until the modeling becomes more available for multiple chemicals and routes, with a narrower range of uncertainty. Hence, total quantity present can be used as a surrogate for exposure measures.

Some might feel apprehensive at the coarseness of the estimates used in risk estimation here. To put this into perspective: regulatory risk assessment methodology often imputes high levels of data uncertainty (e.g., uncertainty factors), as high as 1-4 orders of magnitude range (e.g., ECHA, 2017; Kostal et al., 2020). GHS ranking is definitively conservative in nature, both by its broad use of binning and more importantly the use of the lowest effect or result, often regardless of data quality and variability, although such data restrictions are sometimes incorporated into REACH based risk estimates and classifications (ECHA, 2017). Risk estimates such as DNEL/PNECs (Derived No-Effect Level/Predicted No-Effect Concentration) often incorporate assessment factors ranging between 100 and 10,000 (ECHA, 2017). Modeled exposure estimates (Li et al., 2021) can give uncertainties in excess of four orders of magnitude for a 95% confidence interval.

ChemESI uses discrete judgment (i.e., classification categories) and transforms them into semicontinuous functions. As our goal is to provide a simple metric allowing for broad chemical coverage, we do not need the same granularity as a regulatory (single-chemical) decision tool which indeed would require much more and detailed data. ChemESI as a metric using hazard derived from GHS classifications and exposure as inventory is indeed coarse but most likely with no worse uncertainty than existing exposure/risk models. ChemESI has also another, great advantage: both hazard and exposure datasets are easily obtainable and most likely are already available within an enterprise.

Most of the chemicals in commerce do not exists as 100% pure chemicals, but as products. Products are defined as mixtures of various chemicals, although the actual composition can be hard to define based on SDS's. As with most other assessment systems, GHS classifications for mixtures are hazard based and classified based on weighted contribution from individual chemicals. No consideration of effect interaction between chemicals, be it negative or positive, is currently incorporated in GHS, nor is the collection of endpoints within GHS comprehensive, e.g., the lack of Endocrine Disruptor considerations. It should be noted here that even most individual chemicals are not 100% pure but, depending on quality grade, can contain up to 5–10% "impurities/contaminants" for, e.g., technical grade chemicals. Risk assessment needs to account for these impurities, which is often problematic due to lack of chemical characterization as well as hazard/effect data. Another, more granular, endpoint-specific, hazard screening program, GreenScreen™, requires (confidential) disclosure of impurities above 100 ppm (Clean Production Action, 2021).

The major advantage of GHS has always been its simplicity—the ease of communication across languages and cultures. A GHS compliant SDS looks the same the world over and has a rather narrow range of possible classifications for the various endpoints (UN, 2021). Product Stewards within the chemical community are very familiar with SDS's and associated data. The ChemESI aims to further condense the amount of available information into a facility/enterprise wide KPI (Key Performance Indicator) metric, easily transmittable to the C-suite. As the concept of ChemESI is developed below, the influence of the robustness of the GHS based Hazard Score as well as the influence of inventory fluctuation, organic growth and introduction of new products on the ChemESI scores will be described.

METHODS

Scores were assigned to each of the classification categories for a particular chemical based on classifications following the ninth edition of the GHS (UN, 2021). These scores were then agglomerated for each chemical into a chemical specific score, the actual Chemical Hazard Score, CHS. These scores or weighting factors were derived based on the perception of endpoints: chronic endpoints such as carcinogenicity, reproductive and mutagenic effects are generally given more weight than acute toxicity (e.g., Swanson et al., 1997). These weighting factors are somewhat arbitrary, and users may want to assign their own weighting factors. If done consistently, the rank order of chemical hazards would not be affected, just the magnitude of the hazard score.[1] **Table 1** describes all the GHS categories for environmental health classifications and their corresponding score. The maximum hazard score for a chemical would be capped at 100 but is unlikely to be achieved by many chemicals. Benzene, for instance, would be scored as a carcinogen but does not have a high acute toxicity, which results, as described in **Table 2**, in a weighted derivation of the CHS for benzene of 78, based on a publicly available SDS (Airgas, 2020). For very toxic compounds, e.g., chlorinated dioxins such as TCDD, CHS would be capped at 100 but such compounds are unlikely to be present at high concentrations, i.e., more than 0.1 or 1%.

The lowest score would be 1, e.g., water which is also the default score in the presence of all negative data, i.e., no effects found with supporting data. Missing data are treated differently; unlike in the GHS based SDS, missing data automatically contribute an additional score of 1 for each endpoint with missing data. Most chemicals for instance are missing actual genotoxicity or carcinogenicity information. A score of 1 would thus be hard to achieve without full effect characterization, which may be obtained in future using *in silico* or *in vitro* data aka NAMs (New Approach Methodologies, USEPA, 2021). Missing and/or incomplete data are a continuous problem within GHS classifications and especially those reported on SDS's (ECHA, 2019). Classifications may thus need to be checked with "complete" classification databases such as provided commercially (see below).

As shown in **Tables 1**, **2**, individual-endpoint GHS classifications are easily converted into combined Chemical Hazard Scores for each individual chemical. Such conversions are even easier when using some type of online database, e.g., the publicly available CAt databases (CAt, 2021) or commercial equivalents (e.g., Scivera, 2021; Toxnot, 2021; Verisk 3E, 2021). As always classifications are supposed to be "universal," but small differences based on the regulatory approaches followed, e.g., GHS vs. REACH/CLP may result in different classifications and hence scores. Following the GHS approach, the lowest classification, i.e., highest Hazard Score would be the one used here.

Note the difference between Chemical Hazard Scores (CHS) and GHS: GHS provides classifications for 17 different endpoints, all of which can be blank or have data, with no way of "summing" these effects for each individual chemical. Although theoretically it is possible to calculate weighted GHS scores for each facility, this would still result in 17 parameters that are not comparable. Chemical Hazard Scores on the other hand allow for unified expression of hazard in a single number which is comparable between chemicals. In addition, Chemical Hazard Scores also account to a limited extent for missing data—rare would be the chemical where 3–5 parameters were not identified as no data. Even a major industrial chemical such as benzene (**Table 2**) only has data for 8 out of 15 classification categories. One thus expects a "minimum" CHS to be in the 4–6 range, except for water which is assigned a hazard score of 1. These minimum scores could drop in the future as more economical and ethical NAMs for, e.g., carcinogenic effects, become available: only then can we, with data demonstrating the absence of effects, truly assign an overall CHS of 1 to a given chemical.

ChemESI Risk and Hazard calculations' workflow is shown in **Figure 1**. An inventory database provides an inventory of products for which SDS's are obtained. Based on the SDS's and their composition information both a list of products' compositions and an inventory of individual chemicals are derived. The SDS (or external data) provides the GHS classification information for each individual endpoint for each chemical. Using the scoring system from **Table 1**, a Chemical Hazard Score is derived for each chemical, which can also inform the

[1] For Benzene for instance, assigning maximum weighting factors of 10 for each endpoint would result in a hazard score of 58 rather than 78 with a maximum score still capped at 100. The maximum score of 10 would only effect chronic effects such as carcinogenicity and mutagenicity, for which data are only available for select chemicals.

TABLE 1 | Hazard Scores for individual GHS endpoints and categories.

Hazard scores by endpoints								
Endpoint	**Classification**						**Unknown**	**Not classifiable[a]**
	Category 1	**Category 2**	**Category 3**	**Category 4**	**Category 5**			
Acute toxicity-oral	10	7	5	3	1		1	0
Acute toxicity-dermal	10	7	5	3	1		1	0
Acute toxicity-inhalation	10	7	5	3	1		1	0
STOT[b]-acute	10	7	5				1	0
STOT[b]-repeated	10	7					1	0
	Category 1	**Category 1A**	**Category 1B**	**Category 1C**	**Category 2**	**Category 3**		
Skin corrosion/irritation	10	12	10	7	4	1	1	0
	Category 1	**Category 2**	**Category 2A**	**Category 2B**				
Eye corrosion/irritation	10	6	6	3			1	0
	Category 1	**Category 1A**	**Category 1B**					
Sensitizers—respiratory	15	10	5				1	0
Sensitizers—skin	15	10	5				1	0
	Category 1	**Category 1A**	**Category 1B**	**Category 2**				
Mutagen-germ cell	20	25	15	10			1	0
Carcinogen	20	25	15	10			1	0
Reproductive toxicant	20	25	15	10			1	0
Effects on/*via* lactation	10						1	0
	Category 1	**Category 2**	**Category 3**					
Aquatic hazard-acute	10	6	3				1	0
Aquatic hazard chronic non-degradable[c]	10	6						0
Aquatic hazard chronic degradable[d]	10	6	3					0
Aquatic hazard chronic no data[e]	10	6	3				1	0

[a]*Data available but no classification warranted.*
[b]*Specific Target Organ Toxicity.*
[c]*Non-rapidly biodegradable substances.*
[d]*Rapidly biodegradable substances.*
[e]*No chronic toxicity data available.*

product hazard score (based on product composition). Using inventory quantities, Risk Scores can then be derived for each chemical and product which when summed result in a ChemESI Risk Score. Using the total weight (kkg) of inventory data, a ChemESI Hazard Score can be derived (risk/exposure = hazard).

Note that Chemical Hazard Scores are for one individual chemical which can then be converted into Chemical Hazard product scores based on composition weighting factors. Conversely one can split each product up into individual chemicals and calculate the total amount of chemical X within an enterprise. There are advantages to both approaches; the product version lets one assess whether products have exceeded a hazard cutoff beyond which they are no longer "acceptable" in this enterprises' value chain. The splitting approach identifies how much a particular chemical is present across all products and, coupled with quantity data, how much these chemicals contribute to the overall risk level within an enterprise. The facility-wide ChemESI Hazard score then provides an indication of how "green" the overall enterprise is. As shown above, a facility-wide ChemESI Risk score identifies the total risk inherent to a facility/enterprise. Thus, ChemESI informs enterprise sustainability through both hazard and risk estimates.

The methodology as described here has the advantage of being relatively easy to implement, at the expense of losing detailed information. Drill-down into the data by experts is a definite possibility, but the ChemESI is meant as a KPI, a management (and investor) metric. Uncertainties vary: product quantities can be rather precise, but chemical characterization data for most products are limited and/or plain sloppy, e.g., the SDS component characterization data exceed 110%. The

TABLE 2 | Chemical Hazard Score for benzene based on sample SDS (Airgas, 2020) and **Table 1**.

Benzene example								
Endpoint	**Classification**						**Unknown**	**Not classifiable**[a]
	Category 1	**Category 2**	**Category 3**	**Category 4**	**Category 5**			
Acute toxicity-oral				3				3
Acute toxicity-dermal							1	1
Acute toxicity-inhalation				3				3
STOT[b]-acute							0	0
STOT[b]-repeated	10							10
	Category 1	**Category 1A**	**Category 1B**	**Category 1C**	**Category 2**	**Category 3**		
Skin corrosion/irritation					4			4
	Category 1	**Category 2**	**Category 2A**	**Category 2B**				
Eye corrosion/irritation			6					6
	Category 1	**Category 1A**	**Category 1B**					
Sensitizers-respiratory							1	1
Sensitizers-skin							1	1
	Category 1	**Category 1A**	**Category 1B**	**Category 2**				
Mutagen-germ cell	20							20
Carcinogen	20							20
Reproductive toxicant							1	1
Effects on/*via* lactation							1	1
	Category 1	**Category 2**	**Category 3**					
Aquatic hazard-acute		6						6
Aquatic hazard chronic non-degradable[c]							0	0
Aquatic hazard chronic degradable[d]							0	0
Aquatic hazard chronic no data[e]							1	1
							Chemical Hazard Score	78

[a] *Data available but no classification warranted.*
[b] *Specific Target Organ Toxicity.*
[c] *Non-rapidly biodegradable substances.*
[d] *Rapidly biodegradable substances.*
[e] *No chronic toxicity data available.*

methodology generally trends conservative, i.e., will produce a higher risk estimate due to GHS rules and the scoring of so-called data gaps. The latter can be diminished by the use of less expensive, more ethical non-animal NAMs (New Approach Methodologies) (USEPA, 2021) to fill in some of the data gaps. Unfortunately, GHS assessment of hazardous chemicals at this time only considers a circumscribed number of endpoints; consideration of, e.g., Endocrine Disruptors is still lacking.

Exposure is defined here as the amount of chemical or product in inventory with chemical characterization defined initially by the SDS, hopefully improved by further (confidential) input data from within the supply chain. Ideally such characterization should occur at 100 ppm for each constituent but given the current data quality, that will not be achieved rapidly. A resolution of 1% would seem to be in reach. Exposure estimates are used expressed here in kkg (metric ton) of inventory. For cross value chain comparisons, we recommend that an inventory

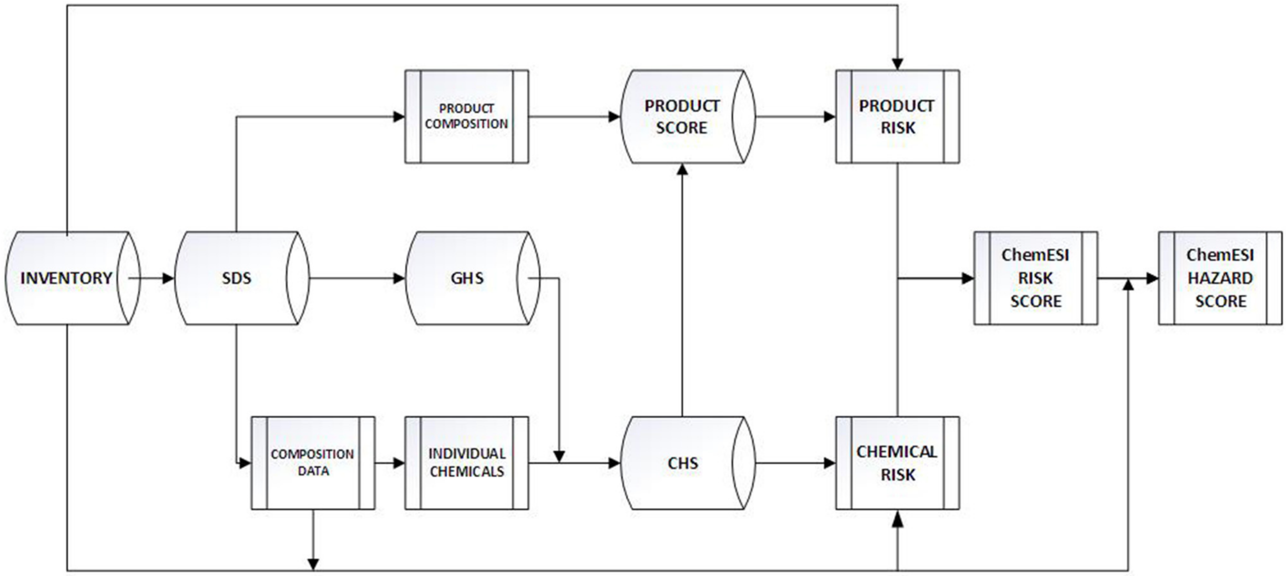


FIGURE 1 | ChemESI workflow.

unit of kkg (metric ton) be adopted, in order to maintain consistent ChemESI Risk Scores. No units need to be assigned or considered as part of the ChemESI Hazard score.

RESULTS

The scoring system in ChemESI was applied to a fictional facility/enterprise that uses and stores a limited array of three products with known composition of 7 constituent chemicals at various concentrations/percentages (**Table 3**). Both ChemESI Risk and Hazard scores were calculated for this enterprise. Results are reported without the use of decimals since the original input data do not employ decimals, i.e., a consistent number of significant digits. Sample Chemical Hazard Scores were ascribed to seven constituent chemicals. Here Chemical Hazard Scores were calculated for each mixture/product with low composition granularity (as is common in SDS's), which unfortunately ignores or minimizes the effect of minor, possibly more toxic constituents. Chemical Risk scores were then derived as the product of exposure (i.e., quantity in inventory) and Chemical Hazard Scores. Note that the product B with the highest product Hazard Score is the least risky and vice versa for product A, quantity and hence exposure matters. Summing of the individual product Risk Scores then produces an enterprise ChemESI Risk Score of 2114 which divided by the total inventory results in an inventory weighted enterprise ChemESI Hazard Score of 14.

In **Table 4** we use the chemical composition data from **Table 3** to derive a facility wide inventory of chemicals 1–7. A chemical-specific risk was then derived using individual chemical inventory estimates (across all products) and the same individual Chemical Hazard Scores as in **Table 3**. Note that here the highest risk contribution comes from chemical 3 which has an average

TABLE 3 | Product-based ChemESI calculation.

Product	Product Quantity		%	Chemical hazard Score	Product hazard Score	Risk
A	100	Chemical 1	20	4		
		Chemical 2	25	6		
		Chemical 3	40	14		
		Chemical 4	15	20	11	1,090
B	10	Chemical 1	35	4		
		Chemical 4	25	20		
		Chemical 5	20	30		
		Chemical 6	20	50	22	224
C	40	Chemical 1	45	4		
		Chemical 3	30	14		
		Chemical 6	20	50		
		Chemical 7	5	80	20	800
	150					
				ChemESI	14	2,114

Chemical Hazard Score. Not surprisingly, the facility ChemESI Hazard and Risk scores are identical between **Tables 3**, **4**—just two different approaches, chemical or product based. For both of these approaches one can make a policy decision to weed out all chemicals/products with a Hazard Score of above, e.g., 50 or focus on the most "risky" chemicals/products. Obviously, a hybrid approach for chemicals/products, e.g., Chemical Hazard Scores over 50 up for elimination followed by a focus on the most risky chemical/product, would work even better. A restricted chemicals list should already be in use within an enterprise,

TABLE 4 | Chemical-based calculation of ChemESI.

	Product Quantity		Hazard Score	Risk
Chemical 1	42		4	166
Chemical 2	25		6	150
Chemical 3	52		14	728
Chemical 4	18		20	350
Chemical 5	2		30	60
Chemical 6	10		50	500
Chemical 7	2		80	160
	150			
		ChemESI	14	2,114

thereby peremptorily eliminating the chemicals with very high hazard scores (e.g., ECHA, 2021).

In the following tables, we focused our assessment on individual chemicals. Given the proliferation of products containing multiple ingredients, an ingredient, i.e., individual chemical-based approach to greening inventory would appear to be more logical as replacement of a single ingredient should result in a greener ingredient and hence multiple greener products. The following analyses thus focus on individual chemical-based ChemESI calculations.

Introduction of a new chemical 8 at the enterprise resulted in a ChemESI Hazard Score of 18 (and ChemESI Risk Score of 3394) as shown in **Table 5**. Although Product Stewardship should have been involved in new product development throughout the entire process (Hart, 2018), there appears to have been a disconnect. Marketing/R&D introduced a single chemical 8 with a high hazard score of 32 and a 25+% increase in overall Product Quantity, resulting in an undesirable increase in ChemESI Risk by more than 50% to 3,394 and a corresponding increase in ChemESI Hazard to 18 (from 14). This metric was flagged in a C-suite meeting resulting in a directive to green the new product, while maintaining functionality (Principle 4 of Green Chemistry in Anastas and Wagner, 1998). The product was revised in two stages: a temporary direct drop-in replacement with chemical X and a final replacement of chemical 8 with chemical 9 as shown in **Table 6**. The result is a 25+% increase in product sales while dropping the ChemESI Hazard Score from 14 to 13, with only a 15% increase in ChemESI Risk Score. Note that we have a decrease in ChemESI Hazard Score with a 25% growth in inventory.

Here, two pathways appear for further reduction of ChemESI Risk and Hazard. One approach would be to replace chemicals with a hazard score of 50 and above, here chemicals 6 and 7. The other approach would be to reduce and/or substitute for chemical with a high risk, here chemicals 3 and 6. Since chemical 6 ends up in both categories, it appears to be a good place to start. Introduction of greener chemical substitutes, e.g., replacing chemical 6 with a less hazardous chemical would result in further reduction of both the ChemESI Hazard and Risk, while simultaneously providing greener products. Continuous improvement should obviously be a part of chemical manufacturing/product formulation going forward—this will result in a continual drop in ChemESI Hazard (and often Risk) as greener products are slated to be developed (Golden et al., 2021). The lag from R&D to commercialization of new greener alternative chemicals however approaches a decade, taking into account all processes including market penetration. Intermediate substitution with slightly greener alternatives may thus be important while working toward the ultimate green substitute.

TABLE 5 | Chemical-based ChemESI following introduction of chemical 8.

	Product Quantity		Hazard Score	Risk
Chemical 1	42		4	166
Chemical 2	25		6	150
Chemical 3	52		14	728
Chemical 4	18		20	350
Chemical 5	2		30	60
Chemical 6	10		50	500
Chemical 7	2		80	160
Chemical 8	40		32	1280
	190			
		ChemESI	18	3,394

TABLE 6 | Chemical-based ChemESI following substitution of chemical 8 with chemical 9.

	Product Quantity		Hazard Score	Risk
Chemical 1	41.5		4	166
Chemical 2	25		6	150
Chemical 3	52		14	728
Chemical 4	17.5		20	350
Chemical 5	2		30	60
Chemical 6	10		50	500
Chemical 7	2		80	160
Chemical 8	0		32	0
Chemical 9	40		8	320
	190			
		ChemESI	13	2,434

As described here most chemical enterprises use multiple chemicals in a variety of recipes to produce their formulated products aka mixtures. In addition, depending on the (technical) grade used for formulation, more or less hazardous, often ill described, impurities may enter the supply chain. These are often not accounted for due to the lack of detailed composition information on the SDS and/or lack of hazard data. This is an information gap that has many implications and, especially for larger quantities of chemicals (missing information data for 1% of a 100 kkg chemical results in 1 kkg of undefined chemicals), could significantly affect hazard/risk assessment. As mentioned above other, more granular systems require characterization down to 100 ppm (Clean Production Action, 2021). Similarly, one assumes that all chemicals listed on an SDS have a unique identifier as well as at least some hazard information. This assumption is often refuted: chemicals end up being described as complex mixtures without a clear component characterization (think hydrocarbon fractions), and hazard data and information are often lacking, resulting in a non-uniform assignment of GHS classification. These data gaps need to be addressed through expert or automated evaluation *via* NAMs which adds complexity. Often, however, GHS classifications for these "exceptions" can be found on SDS's, but with varying classifications between manufacturers.

The ideal ChemESI would of course be 1. It is unlikely that any enterprise will get an ChemESI of 1 in the foreseeable future, primarily as the result of a lack of chemical characterization and full hazard data. As mentioned above, even for a commercially important chemical such as benzene only 8 out of 15 classification categories are available (Airgas, 2020). Another incentive here lies in full characterization of effects from each chemical, using for instance New Approach Methodologies, so-called NAMs (USEPA, 2021), without which a ChemESI hazard score of 1 is not attainable.

The ChemESI Hazard/Risk within the supply chain on a facility basis, can be combined into a division/enterprise level ChemESI. Even though Product Stewardship should already be limiting very hazardous chemicals from entering a supply chain, the ChemESI process will also identify remaining very hazardous chemicals with a Hazard Score beyond "acceptable." Acceptable of course has to be defined on an enterprise-wide basis and may be accompanied by an exposure limit, e.g., no more than X kkg of Y can be within the facility at all times. Such exceptions may prove especially useful when technical grade (i.e., impure) chemicals are in the supply chain; such chemicals often contain more hazardous, contaminant constituents.

Exposure can also vary depending on the amount of inventory, e.g., running up inventory in anticipation of supply chain disruptions or taking advantage of discounted pricing on chemicals in bulk. Such inventory changes will most likely increase the ChemESI Risk. Risk minimization, i.e., fast inventory turnover will limit the risk incurred during daily operations, and in optimized cases even reduce ChemESI Risk while maintaining ChemESI Hazard. One should still perform an annual product inventory quantity survey to determine the total amount of "risk" incurred within the enterprise from an ESG point of view, although a running average would provide a better immediate KPI.

The ChemESI KPI indicates a general level of chemical environmental sustainability, i.e., the progress made in providing greener alternatives to society within the existing supply chain. Such sustainability indices become increasingly important as greenness become a major economic driver within (the chemical) industry. Much progress has been made—a lot of effort (Zimmerman et al., 2020) remains to be expended to truly bring Green Chemistry to the forefront as evidenced by the eventual possibility of an ChemESI near 1. In the meantime, growing industrial output with greener chemicals will slowly reduce the ChemESI KPI, while allowing for substantial growth.

DISCUSSION

ChemESI Risk and Hazard are important KPIs especially for the chemical industry including formulators. These KPIs use exposure and hazard information readily available from multiple sources: in-house, or from free and commercially available databases. Inventory data here is used as a surrogate for exposure. The Chemical Hazard Score system described here in **Table 1** is based on the widely accepted and adopted GHS classification approach—one can always dispute the weighting and prioritization of certain endpoints, but such changes would only have a slight impact on the overall ChemESI. Better to adopt a universal metric than to customize each enterprise.

Given the inherent uncertainties and biases built into the GHS classifications [e.g., a direct result of the derivation methodology for hazard classes using lowest effect concentration (ECHA, 2017; UN, 2021)], the ChemESI will tend toward overestimating hazard and hence risk. One of the inherent drawbacks in the GHS is that lack of data simply results in no classification—addressed here by adding additional scores for each hazard endpoint for which no data are available—one would hope that, especially for the higher volume chemicals, such full information would become available, most likely based on NAMs. Another major drawback is the GHS's dependency on the lowest available hazard information, i.e., the most toxic number is always used with little consideration of data quality and variability. Given that where there are repeat data, values often range across more than an order of magnitude (e.g., Plugge et al., 2021), this also adds a severe bias toward a more hazardous classification, especially in the absence of data quality considerations.

The lack of information data for risk assessment from GHS classifications and in SDS's is to some extent not surprising. SDS's were developed as hazard communication tools in occupational health and hazardous material transport. The lack of granularity, i.e., very few components characterized at concentrations below 1%, however is not inherent to the design of SDS's and probably remains the least reliable aspect of an SDS.

Some might question the coarseness of the approach. As addressed above the simplifying assumptions here allow for an uncertainty in the same order(s) of magnitude as the original data. Hazard estimates can have uncertainty/variability well in excess of an order(s) of magnitude. Similarly, modeled exposure often has orders of magnitude of uncertainty/variability. Use of the GHS data does allow for a simpler format, inherently familiar to the Product Steward, thereby enhancing the understanding and hence communication of the ChemESI to the C-suite and investors. Single number KPI's are always attractive while simultaneously allowing drill down into the data to assess which chemicals are contributing most to the overall risk and hazard. Often the chemicals with medium hazard are the biggest drivers of ChemESI risk parameters, simply because they often account for the largest tonnage. The ChemESI data informs both hazard and risk metrics that can be used to address the overall hazard and risk inherent to the inventory. Growth will decrease the ChemESI Hazard Score where growth is derived from greener products. If the Chemical Hazard Scores are low enough even ChemESI Risk Score would drop, coupled with substantial growth in greener products. Supply chain effects on retention/storage of both raw and finished products will most likely increase the ChemESI Risk Score, but have a lesser effect on the ChemESI Hazard Score. Dilution of more hazardous chemicals would decrease the ChemESI Hazard score but not the Chem ESI Risk score. Incorporation of LCA (Life Cycle Analysis) components and additional non-GHS endpoints could be the next logical set of enhancements to the ChemESI, although obtaining (semi-)curated data for all chemicals may prove problematic.

ESG (Environmental, Social, and Governance) discussions often lack (quantitative) data on the environmental impact of the chemicals within an enterprise, beyond greenhouse gases and decarbonization (Investopedia, 2021). The ChemESI described here is a quantitative risk and hazard metric combining chemical exposure and environmental health hazard within the supply chain. It is based on a scoring of readily available GHS hazard/classification data, coupled with proprietary inventory data. The ChemESI will provide a facility/enterprise based KPI metric of risk and hazard, easily communicable to the C-suite, investors, and others within the ESG sphere.

Although achievement of an ChemESI Hazard of 1 is theoretically achievable, it will be most likely a decade before green chemistry-based alternatives will be able to achieve such scores, although penetration of green chemistry derived products is rapidly increasing (Golden et al., 2021). An ChemESI Hazard of 20 or below should be achievable now, with of course continuous improvement expected. Organic growth will not affect a ChemESI *per se*: most likely a combination of a reduction in ChemESI Hazard Score and increased growth will occur through continued introduction of greener products, especially with the phasing out of less greener products. Adoption of ChemESI will go a long way toward quantifying the environmental sustainability metric in ESG.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

Airgas (2020). *Safety Data Sheet – Benzene*. Available online at: https://www.airgas.com/msds/001062.pdf (accessed December 6, 2021).

Anastas, P., and Wagner, J. C. (1998). *Green Chemistry: Theory and Practice*. Oxford: Oxford University Press.

Arnot, J. A., Brown, T. N., Wania, F., Breivik, K., and McLachlan, M. S. (2012). Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. *Environ. Health Perspect.* 120, 1565–1570. doi: 10.1289/ehp.1205355

Arnot, J. A., and Mackay, D. (2008). Policies for chemical hazard and risk priority setting: can persistence, bioaccumulation, toxicity and quantity information be combined? *Environ. Sci. Technol.* 42, 4648–4654. doi: 10.1021/es800106g

Bare, J. C. (2002). TRACI: The tool for the reduction and assessment of chemical and other environmental impacts. *J. Indus. Ecol.* 6, 49–78. doi: 10.1162/108819802766269539

CAt (2021). Available online at: https://www.chemicalawarenesstoolkit.com/home (accessed December 8, 2021).

Chemsec (2021). *Hazard vs. Risk – What Is Best Practice When Assessing Chemicals?* Available online at: https://chemsec.org/policy-and-positions/hazard-risk/?utm_source=ChemSec$+$newsletter&utm_campaign=7a51b01ea8-Hazard_risk&utm_medium=email&utm_term=0_03a9107408-7a51b01ea8-483751655 (accessed December 8, 2021).

Clean Production Action (2021). Available online at: https://www.cleanproduction.org/programs/greenscreen (accessed December 8, 2021).

ECHA (2017). *Guidance on the Application of the CLP Criteria*. Available online at: https://echa.europa.eu/documents/10162/23036412/clp_en.pdf/58b5dc6d-ac2a-4910-9702-e9e1f5051cc5 (accessed December 1, 2021).

ECHA (2019). *FORUM Report on Improvement of Quality of SDS WG "Joint initiative ECHA Forum – ECHA ASOs on Improvement of the quality of SDS"*. Available online at: https://echa.europa.eu/documents/10162/22749747/echa_sds_report_en.pdf/0ddb021f-bd1a-09e6-29f3-c52ca476e6aa (accessed December 8, 2021).

ECHA (2021). *Substances Restricted Under REACH*. Available online at: https://echa.europa.eu/substances-restricted-under-reach (accessed December 8, 2021).

Escher, B. I., and Fenner, K. (2011). Recent advances in environmental risk assessment of transformation products. *Environ. Sci. Technol.* 45, 3835–3847. doi: 10.1021/es1030799

Escher, S. E., Mangelsdorf, I., Hoffmann, D. S., Partosch, F., Karwath, A., Schroeder, K., et al. (2020). Time extrapolation in regulatory risk assessment: the impact of study differences on the extrapolation factors. *Regul. Toxicol. Pharmacol.* 112, 104584. doi: 10.1016/j.yrtph.2020.104584

Golden, J., Handfield, R., Daystar, J., Kronthal-Sacco, R., and Tickner, J. (2021). *Green Chemistry: A Strong Driver of Innovation, Growth, and Business Opportunity*. Green Chemistry and Commerce Council. Available online at: https://greenchemistryandcommerce.org/documents/GC3GreenChemReport-ES-Nov2021.pdf?utm_source=BZ (accessed December 1, 2021).

Hart, G. (ed.). (2018). *Professional Practices of Product Stewardship*. Falls Church, VA: AIHA Product Stewardship Society.

Investopedia (2021). Available online at: https://www.investopedia.com/terms/e/environmental-social-and-governance-esg-criteria.asp (accessed December 6, 2021).

Investopedia (2022). Available online at: https://www.investopedia.com/terms/r/risk.asp (accessed February 4, 2022).

IRMI (2022a). Available online at: https://www.irmi.com/term/insurance-definitions/risk#:$\sim$:text=Risk%20%E2%80%94%20(1)%20Uncertainty%20arising,which%20an%20insurance%20policy%20relates (accessed February 4, 2022).

IRMI (2022b). Available online at: https://www.irmi.com/term/insurance-definitions/exposure (accessed February 4, 2022).

Kostal, J., Plugge, H., and Raderman, W. (2020). Quantifying uncertainty in ecotoxicological risk assessment: MUST, a modular uncertainty scoring tool. *Environ. Sci. Technol.* 54, 12262–12270. doi: 10.1021/acs.est.0c02224

Li, L., Sangion, A., Wania, F., Armitage, J. A., Toose, L., Hughes, L., and Arnot, J. A. (2021). Development and Evaluation of a Holistic and Mechanistic Modeling Framework for Chemical Emissions, Fate, Exposure, and Risk. *Env. Health Persp*. 129:127006. doi: 10.1289/EHP9372

Maertens, A., and Plugge, H. (2018). Better metrics for "sustainable by design": toward an *in silico* green toxicology for Green(er) Chemistry. *ACS Sustain. Chem. Eng.* 6, 1999–2003. doi: 10.1021/acssuschemeng.7b03393

Pham, L., Sheffield, T. Y., Pradeep, P., Brown, J., Haggard, D. E., Wambaugh, J., et al. (2019). Estimating uncertainty in the context of new approach methodologies for potential use in chemical safety evaluation. *J. Compu Tox.* 15, 40–47. doi: 10.1016/j.cotox.2019.04.001

Plugge, H., Das, N., and Kostal, J. (2021). Towards a universal acute fish threshold of toxicological concern (TTC). *Environ. Toxicol. Chem.* 40, 1740–1749. doi: 10.1002/etc.4991

Risk and Policy Analysts (2019). *Study to Collect Recent Information Relevant to Modernising EU Occupational Safety and Health Chemicals Legislation With a Particular Emphasis on Reprotoxic Chemicals With the View to Analyse the Health, Socio-Economic and Environmental Impacts in Connection With Possible Amendments of Directive 2004/37/EC and Directive 98/24/EC.* Prepared for European Commission, DG Employment, Social Affairs and Inclusion. Available online at: https://op.europa.eu/en/publication-detail/-/publication/094387fb-da9a-11e9-9c4e-01aa75ed71a1; https://www.researchgate.net/publication/334260797_Study_to_collect_recent_information_relevant_to_modernising_EU_Occupational_Safety_and_Health_chemicals_legislation_with_a_particular_emphasis_on_reprotoxic_chemicals_with_the_view_to_analyse_the_heal

Scivera (2021). Available online at: https://www.scivera.com/ (accessed December 13, 2021).

SHiPP (2021). *Sustainable Health in Procurement Project*. Available online at: https://noharm-global.org/issues/global/sustainable-health-procurement-project (accessed December 8, 2021).

Spaniol, O., Bergheim, M., Dawick, J., Kotter, D., McDonough, K., Schowanek, D., et al. (2021). Comparing the European Union System for the Evaluation of Substances (EUSES) environmental exposuremcalculations with monitoring data for alkyl sulphate surfactants. *Environ. Sci. Eur.* 33, 3. doi: 10.1186/s12302-020-00435-1

Swanson, M. B., Davis, G. A., Kincaid, L. E., Schultz, T. W., Bartmess, J. E., Jones, S. L., et al. (1997). A screening method for ranking and scoring chemicals by potential human health and environmental impacts. *Environ. Toxicol. Chem.* 16, 372–383. doi: 10.1002/etc.5620160237

Toxnot (2021). Available online at: https://toxnot.com/ (accessed December 13, 2021).

UNDP (United Nations Development Programme) (2020). *Sustainable Health Procurement Guidance Note.* Available online at: https://www.undp.org/sites/g/files/zskgke326/files/publications/undp-SPHS-bpps-health_Sustainable_Health_Procurement_Guidance_Note.pdf (accessed December 8, 2021).

United Nations (UN) (2021). *Globally Harmonized System of Classification and Labelling of Chemicals (GHS Rev. 9, 2021).* Available online at: https://unece.org/transport/standards/transport/dangerous-goods/ghs-rev9-2021 (accessed December 1, 2021).

USEPA (1993). *Reference Dose (RfD): Description and Use in Health Risk Assessments.* Background Document 1A. Available online at: https://www.epa.gov/iris/reference-dose-rfd-description-and-use-health-risk-assessments (accessed December 8, 2021).

USEPA (2021). *Alternative Test Methods and Strategies to Reduce Vertebrate Animal Testing.* Available online at: (accessed December 6, 2021).

Verisk 3E (2021). Available online at: https://www.verisk3e.com/ (accessed December 13, 2021).

Zimmerman, J. B., Anastas, P. T., Erythropel, H. C., and Leitner, W. (2020). Designing for a green chemistry future. *Science* 367, 397–400. doi: 10.1126/science.aay3060

Conflict of Interest: HP is employed by Safer Chemical Analytics LLC.

frontiers | Frontiers in Sustainability

ORIGINAL RESEARCH
published: 08 July 2022
doi: 10.3389/frsus.2022.909239

The Ultimate Owner of Environmental, Social, and Governance Investment

Alexander R. Keeley[1,2]*, *Chao Li*[1], *Shutaro Takeda*[2], *Tom Gloria*[3] and *Shunsuke Managi*[1,2]

[1] Department of Civil Engineering, Kyushu University, Fukuoka, Japan, [2] Urban Institute, Kyushu University, Fukuoka, Japan, [3] Industrial Ecology Consultants, Newton, MA, United States

The coronavirus (COVID-19) pandemic has affected society in immeasurable ways, including investment. As the pandemic has impacted society's values, it has proven to be a major turning point for environmental, social, and governance (ESG) investment. This investment approach, which evaluates a company's ESG ratings alongside traditional financial metrics, was already "coming off a banner year," and its reach continues to expand. Although numerous studies have investigated the impact of ESG scores on financial returns and the trend in ESG investment strategies, only a limited number of studies have attempted to capture the key players in ESG investment. Therefore, to determine the most influential investors in the ESG investment field, the cumulative impacts are calculated based on the ESG scores of invested companies, the total market price of invested companies, and the investor history portfolio report. We perform an iteration of calculation to convey the impacts that the invested companies have on the ultimate investors, and we identify the major players in the field and differences in the trend by type of investor and country.

Keywords: environmental, social, governance, ultimate ownership structure, network analysis, socially responsible investing, sustainable investment

OPEN ACCESS

Edited by:
Syed Abdul Rehman Khan, Tsinghua University, China

Reviewed by:
Muhammad Tanveer, Prince Sultan University, Saudi Arabia
Usama Awan, Lappeenranta University of Technology, Finland

***Correspondence:**
Alexander R. Keeley
keeley.ryota.alexander.416@m.kyushu-u.ac.jp

Specialty section:
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

Received: *31 March 2022*
Accepted: *03 June 2022*
Published: *08 July 2022*

Citation:
Keeley AR, Li C, Takeda S, Gloria T and Managi S (2022) The Ultimate Owner of Environmental, Social, and Governance Investment. Front. Sustain. 3:909239. doi: 10.3389/frsus.2022.909239

INTRODUCTION

Through more than 20-years history of sustainable finance, there has been significant surge in ESG investment, ESG metrics, and related studies in the last few years. The 2020 *Global Sustainable Investment Review* (GSIR) published by the Global Sustainable Investment Alliance (GSIA) reports that sustainable investments across five markets, including the United States (the US) and the European Union, have reached US Dollar (USD) 35.3 trillion in assets under management, which is equivalent to 36% of all professionally managed assets in these markets. The US market alone grew 42% from 2018 to 2020, reaching USD 17.1 trillion (Global Sustainable Investment Alliance, 2021). Furthermore, the GSIA concluded that the most common sustainable investment strategy is "ESG integration, followed by negative screening, corporate engagement and shareholder action, norms-based screening and sustainability-themed investment" (Global Sustainable Investment Alliance, 2021). The coronavirus (COVID-19) pandemic has severely affected companies in the global supply chain, and studies have found that firms' governance and supply chain management ability are important in responding to adverse events (Khan et al., 2021a,b). Moreover, the ESG performances of firms have been considered a key to the resiliency of firms' financial performance (Yoo et al., 2021). As the pandemic has impacted society's values, it has proven to be a major turning point for ESG investment. In June 2020, the Morningstar report found that there were approximately 3,432 sustainable funds on the market, comprising a mix of newly formed and revamped or rebranded funds (MorningStar, 2022). This represents an increase of 50% in the total number of sustainable funds over the past three years (Bryan et al., 2020). Furthermore, studies assessing ESG ratings are increasing. For instance, if the term "ESG ratings" is searched on Google Scholar, it produces 1,500 results from 2021. The increased interest in ESG has also resulted in an increased influence on financial decisions, asset prices, and corporate policies, with potentially far-reaching effects on political and economic events, such as the Russia–Ukraine war (Funk and Hento, 2022).

Over the past decade, numerous studies have focused on the relation between ESG investment and financial performance, demonstrating that ESG portfolio tilting and integration of ESG

TABLE 1 | Summary of the ESG score multiplied with the total market price value.

Country/region	Count	Mean ESG·price	Country/region	Count	Mean ESG·price
United States	2,616	17,290,890,769	Colombia	22	2,576,361,779
China	885	9,259,978,013	Poland	21	1,632,832,335
United Kingdom	560	5,545,439,379	Qatar	18	3,096,385,653
Canada	507	10,210,697,484	United Arab Emirates	17	6,765,372,320
Japan	426	9,284,369,449	Philippines	16	4,771,530,768
Australia	298	3,878,531,833	Cyprus	10	1,149,318,279
Sweden	288	3,760,015,807	Jersey	10	938,270,975
Germany	242	8,922,779,941	Portugal	10	3,164,730,769
Switzerland	176	15,585,900,319	Cayman Islands	9	1,177,230,585
Brazil	173	5,052,692,113	Egypt	7	704,923,782
France	166	17,116,931,160	Iceland	7	945,416,432
India	153	9,839,711,767	Kuwait	6	1,654,790,323
South Korea	152	10,134,724,108	Oman	6	523,785,241
Taiwan	138	15,178,094,606	Monaco	5	202,451,681
Hong Kong	116	7,792,178,910	Puerto Rico	5	1,171,903,482
Italy	102	4,506,868,253	Czech Republic	4	5,002,480,431
Thailand	100	3,292,839,345	Isle of Man	4	2,514,196,007
South Africa	95	2,531,222,955	Macau	4	9,158,967,375
Netherlands	73	21,126,002,256	Malta	4	551,009,988
Russia	67	12,756,897,908	Pakistan	4	392,707,082
Singapore	67	4,596,479,516	Panama	4	302,139,638
Argentina	66	1,813,294,133	Bahrain	3	709,414,006
Finland	66	4,413,335,979	Kazakhstan	3	7,381,260,564
Spain	66	10,778,734,679	Romania	3	2,825,694,139
Denmark	63	12,341,266,284	Vietnam	3	3,887,567,877
Norway	63	12,561,551,031	Hungary	2	6,428,356,124
Malaysia	62	2,723,950,440	Jordan	2	1,675,108,638
Mexico	56	6,275,043,837	Liechtenstein	2	639,502,060
Bermuda	53	2,223,524,645	Ukraine	2	564,153,156
Turkey	53	1,108,180,405	Uruguay	2	634,325,149
Chile	51	2,405,046,793	Bahamas	1	317,208,058
Ireland	47	14,277,302,284	British Virgin Islands	1	13,259,754
New Zealand	45	1,290,262,405	Cambodia	1	2,699,443,454
Belgium	44	6,344,100,616	Costa Rica	1	760,185,572
Luxembourg	40	3,680,306,579	Faroe Islands	1	2,516,543,634
Indonesia	36	5,527,856,129	Gibraltar	1	776,433,042
Israel	29	3,443,914,513	Kenya	1	6,111,597,436
Peru	29	914,026,365	Morocco	1	6,346,036,210
Greece	25	915,702,038	Papua New Guinea	1	3,808,978,938
Guernsey	25	738,509,213	Slovenia	1	2,559,272,403
Saudi Arabia	25	8,298,867,053	Missing	70	2,215,114,276
Austria	23	2,020,398,644			

factors can impact portfolio and corporate financial performance in various ways (OECD, 2021). Whereas, some studies have demonstrated that integrating ESG criteria has no significant impact on portfolio return (Bauer et al., 2007; Cortez et al., 2009); many studies have found that it has a positive effect on returns by comparing equity portfolios with high and low ESG scores (Widyawati, 2020). Widyawati (2020) conducted a systematic review on socially responsive investment (SRI) and ESG metrics and asserted that ESG metrics play a role of an enabler in the SRI market.

In addition, there has been growing interest in the study of the relation between the ESG strategy and actual ESG portfolio allocation. Brandon et al. (2021) calculated the direct ESG ownership scores of institutional investors and investigated whether institutional investors' public commitments to responsible investing and higher reported levels of ESG

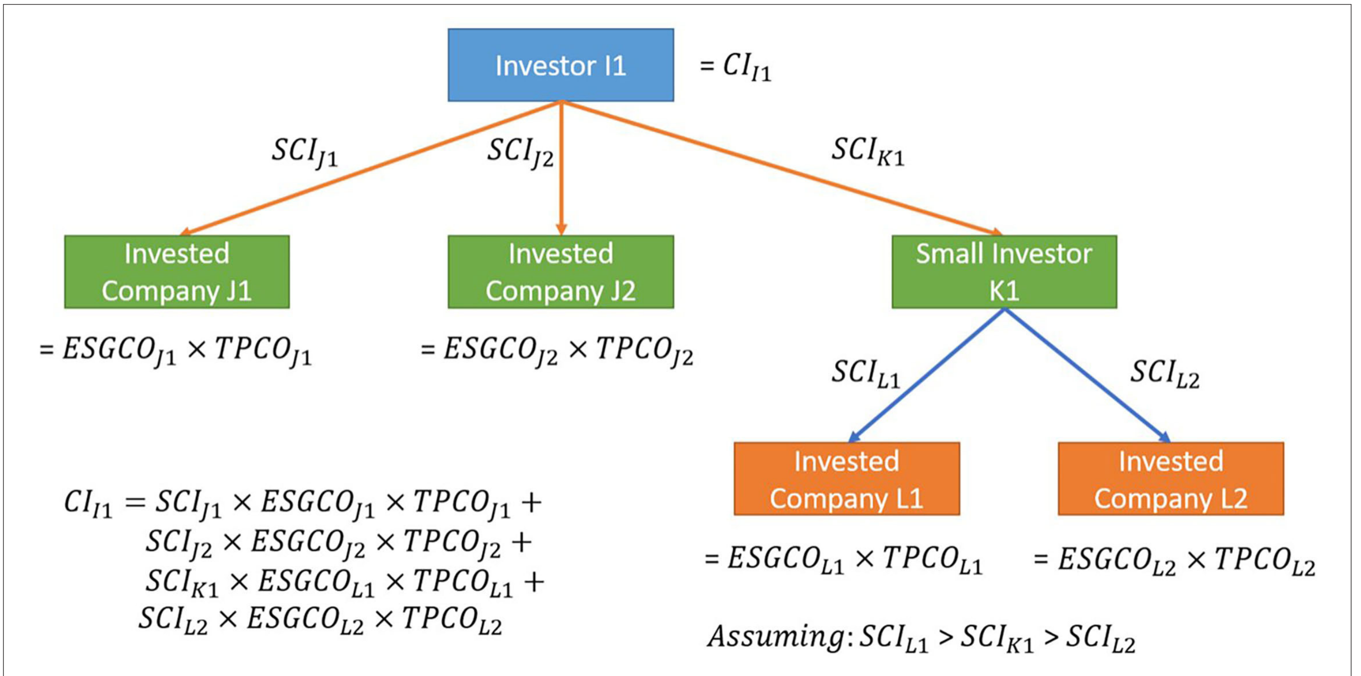


FIGURE 1 | Example to illustrate the calculation process.

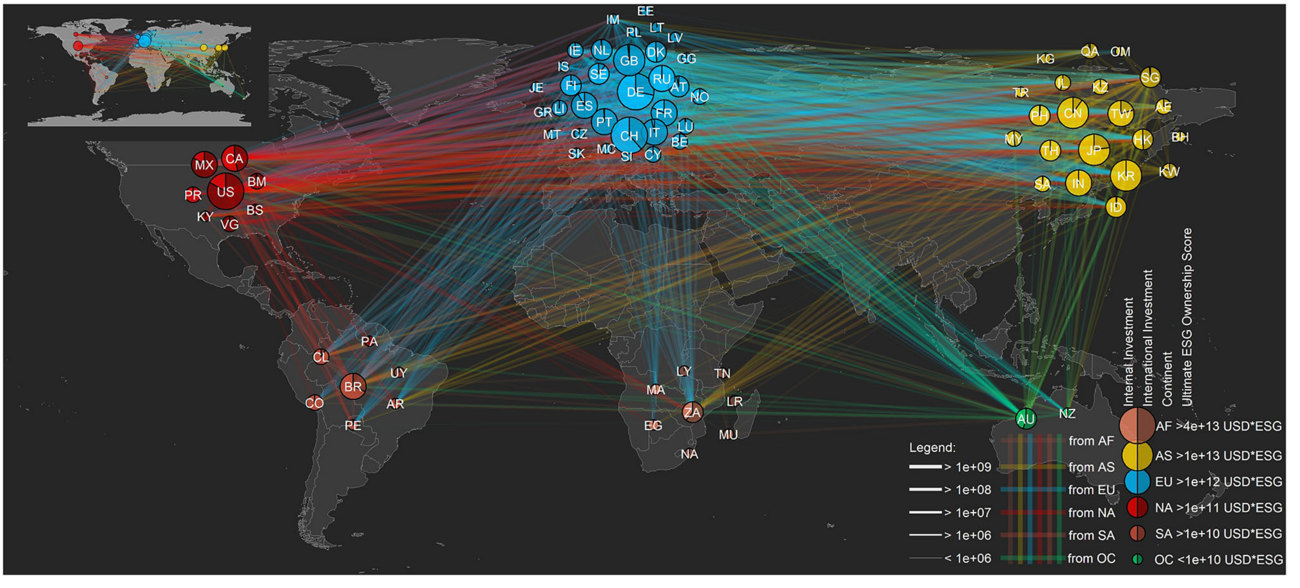


FIGURE 2 | ESG ownership country-level network map.

incorporation result in more sustainable portfolio allocations. The findings revealed that the Principles for Responsible Investment (PRI) signatory institutions have better ESG scores in general but not in the US. To examine the ownership and control structure of large shareholders, relevant studies have examined the ultimate ownership of companies by mostly tracking portfolio and total market price data. For example, La Porta et al. (1999) investigated the control chains of a sample of 30 firms in 27 countries and identified the ultimate controlling owners. Similarly, Claessens et al. (2000) examined 2,980 listed companies in nine East Asian countries, revealing that a large percentage of stock market capitalization in the investigated countries is owned by a small number of investors. Faccio and Lang (2002) analyzed the ownership structure using the ultimate

TABLE 2 | Total ultimate ESG ownership score by country.

Country	Total ultimate ESG ownership score	Number of investor	Country	Total ultimate ESG ownership score	Number of investor
Switzerland	8.27014E + 13	175	Cyprus	25,675,250,559	20
Germany	5.72266E + 13	231	Ireland	17,977,734,143	20
United States	4.65301E + 13	2,606	Colombia	16,937,905,030	23
China (Mainland)	2.27566E + 13	573	Liechtenstein	14,871,610,630	28
Japan	2.00865E + 13	1,379	Bermuda	12,970,382,575	9
South Korea	2.00643E + 13	85	Kuwait	12,541,070,793	3
United Kingdom	1.03016E +13	458	Kazakhstan	10,920,018,272	1
France	7.94945E + 12	97	Cayman Islands	7,924,821,736	13
Canada	6.95242E + 12	186	Poland	5,921,687,310	21
Spain	5.64705E + 12	84	Turkey	4,611,654,837	2
Italy	4.9574E + 12	82	Peru	3,942,964,465	11
Brazil	2.86728E + 12	236	Morocco	3,493,924,925	9
India	2.18313E + 12	57	Monaco	3,319,268,299	2
Portugal	1.60532E + 12	4	New Zealand	2,561,005,844	16
Mexico	1.49459E + 12	29	Jersey	1,734,744,467	14
Russia	1.42537E + 12	17	Bahamas	1,539,192,807	8
Taiwan	1.02775E + 12	124	Guernsey	1,409,583,433	9
Sweden	9.38617E + 11	137	Czech Republic	1,277,659,891	5
Finland	4.36005E + 11	63	Argentina	1,054,893,765	20
Philippines	3.84733E + 11	26	Bahrain	913,981,720	1
Austria	3.33984E + 11	30	Tunisia	545,118,593	2
Netherlands	3.28674E + 11	67	Malta	534,298,615	7
South Africa	2.8943E + 11	80	Isle of Man	327,054,416	11
Singapore	2.60384E + 11	98	NA	296,935,231	1
Thailand	1.89867E + 11	18	Greece	262,443,065	1
Australia	1.71408E + 11	181	Kyrgyzstan	261,321,470	1
Hong Kong	1.2845E + 11	79	Oman	228,403,657	3
Denmark	1.20054E + 11	39	Mauritius	182,277,069	6
Indonesia	1.03904E + 11	45	Iceland	175,605,083	1
Luxembourg	96618161373	86	Panama	163,443,293	1
Norway	86334851352	65	Liberia	147,353,870	2
Qatar	70311100528	7	Libya	139,581,027	1
Saudi Arabia	64429600766	14	Egypt	122,983,423	2
Chile	55231057120	52	Gibraltar	16,330,931	3
United Arab Emirates	55085170802	21	Latvia	16,259,175	1
Malaysia	48030188239	58	Uruguay	10,020,831	1
Israel	38441061679	23	Slovakia	8,139,424	1
British Virgin Islands	32456747890	38	Lithuania	3,969,290	1
Belgium	31782228767	21	Namibia	1,592,377	1
Puerto Rico	30891546768	2	Estonia	1,374,334	1
Cyprus	25675250559	20	Slovenia	65,803	1

ownership data of 5,232 listed Western European corporations. Their research found that financial and large firms were more likely to be widely held, whereas nonfinancial and small firms were more likely to be privately owned. These aforementioned studies on ultimate ownership provide insight into patterns by the type of investor and nation–state political boundaries in the context of direct and indirect ownership chains. In the case of ESG ownership, previous studies have mostly used direct ownership scores of ESG at the portfolio level using the ESG portfolio data, total market price data, and ESG scores provided by leading ESG rating agencies. To capture the ownership structure of ESG investment, it is crucial to consider both direct and indirect ownership. Notably, while most studies on ESG ownership emphasize institutional investor, intercorporate investment also plays a critical role in the ESG investment field as investments in financial assets and associates and business combinations (Silver et al., 2021). Intercorporate investment can occur when a company invests in another company. These types of investments can be considered in different ways depending on the investment. In general, the archetypes of ownership stake by percentage fall into three categories—minority passive (<20% ownership), minority active (20–50% ownership), and controlling interest (over 50% ownership) (Kenton et al., 2021).

This study addresses these research gaps and investigates the ultimate ESG ownership, considering both direct and indirect ownership of investors, including institutional investors and corporations, following the ultimate ownership analysis framework employed by Faccio and Lang (2002). We identify the major players in the field in terms of the ultimate ESG ownership score and provide insight into patterns by the type of investor and nation–state political boundaries in the context of direct and indirect ownership chains. By expanding the study by Brandon et al. (2021), we also re-examine whether the PRI signatory institutions and corporations demonstrate better ESG ownership scores and confirm that the PRI signatory investors have better ESG ownership score in general. Furthermore, our study contributes to extant literature examining fund flow reactions to ESG ratings (Hartzmark and Sussman, 2019) by providing the evidence that it is difficult to target the actual flow and ownership structure of ESG investment through ESG ratings of the investors themselves. In addition, this study reports that the top 10% investors own approximately 98% of the total ultimate ESG ownership score, showing that ESG investment is led by relatively small number of investors. Furthermore, the results of the study confirm the important role of corporations in ESG investment, which owns the most global shares (85.77%) measured by the ESG ownership score considering both direct and indirect ownership chains.

The remainder of this paper is structured as follows. Section Data and Methods outlines the data and analytical methods employed in this study. Section Results discusses the results of

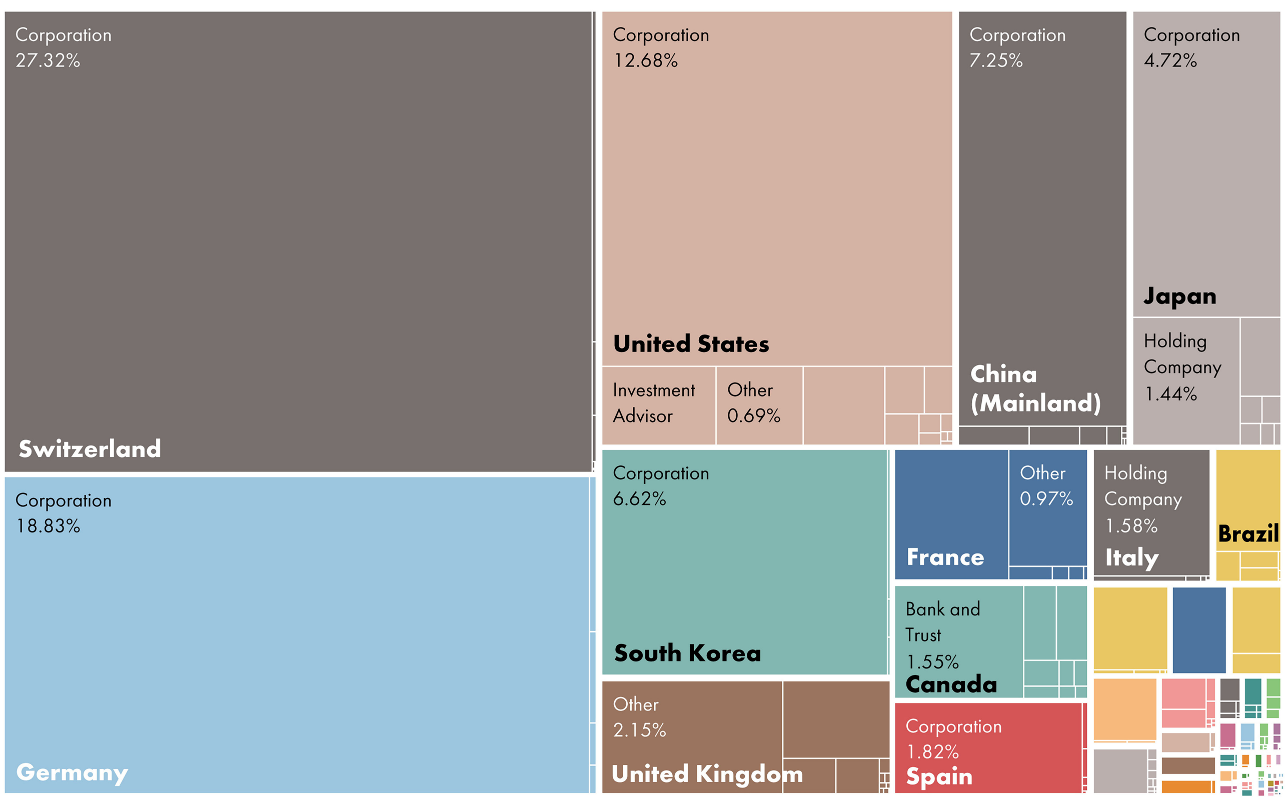


FIGURE 3 | Total ultimate ESG ownership by country.

the ultimate ESG ownership analysis, and Section Discussion and Conclusion presents the discussion and conclusion.

DATA AND METHODS

Data

ESG Score

To obtain the ESG scores of important companies across the world, we use the data management system of Refinitiv Eikon to download the records of 30,857 companies. As Refintiv Eikon does not gather all the data of these companies, we acquire the ESG combined scores of 8,662 companies. Companies that can be used in the analysis have to provide their ESG combined scores and the total market price. In this study, we employ the latest ESG combined scores in 2020. The ESG combined scores range from 0 to 1. Additionally, although the ESG combined scores of several companies are not listed on the summary page, they present their ESG reports, comprising the rankings of their ESG combined scores. The rankings are from "A+" to "D-," that is, 12 levels. According to the official statements from Refintiv, the raw ESG scores are converted into a numeric format, from 1 (12/12) to 0.0833 (1/12). The companies with ESG scores are from 83 countries or regions. Most companies in our dataset are based in the US, China (Mainland), and the UK, having 2,616, 885, and 560 companies, respectively (detailed information is presented in **Appendix Table A**).

Investor History Portfolio Reports

The investors' data, which are the investor history portfolio reports of each investor company, are also from Refinitiv Eikon. These reports capture the shares of the investors in the invested company. We assume that the shares represent the voice of the ESG strategies of the investors in the invested companies. The most recent data, generally for the first quarter of 2021, are employed in this study. To cover the invested companies with ESG scores, 43,482 investors, including investment companies and individual investors, are investigated and 12,258 investor history portfolio reports are downloaded. The investment share of a specific investor is recorded in the investor history portfolio. For example, the investor history portfolio of the investor company Storebrand Kapitalforvaltning AS is 0.04% of Apple Inc., 0.04% of Microsoft Corp., 0.03% of Amazon.com Inc., and others. In this case, 0.04% of the ESG investment results of Apple Inc. should be attributed to Storebrand Kapitalforvaltning AS. The investor history portfolios are the critical data set to link the invested and investing companies.

Total Market Price of Companies With ESG Reports

Furthermore, the total market prices of the aforementioned invested companies are from Refinitiv Eikon. The total market price varies in real time. We secured all the available data within a week—the last week in December 2021. After several downloading waves, we retrieved the total market prices of 8,662

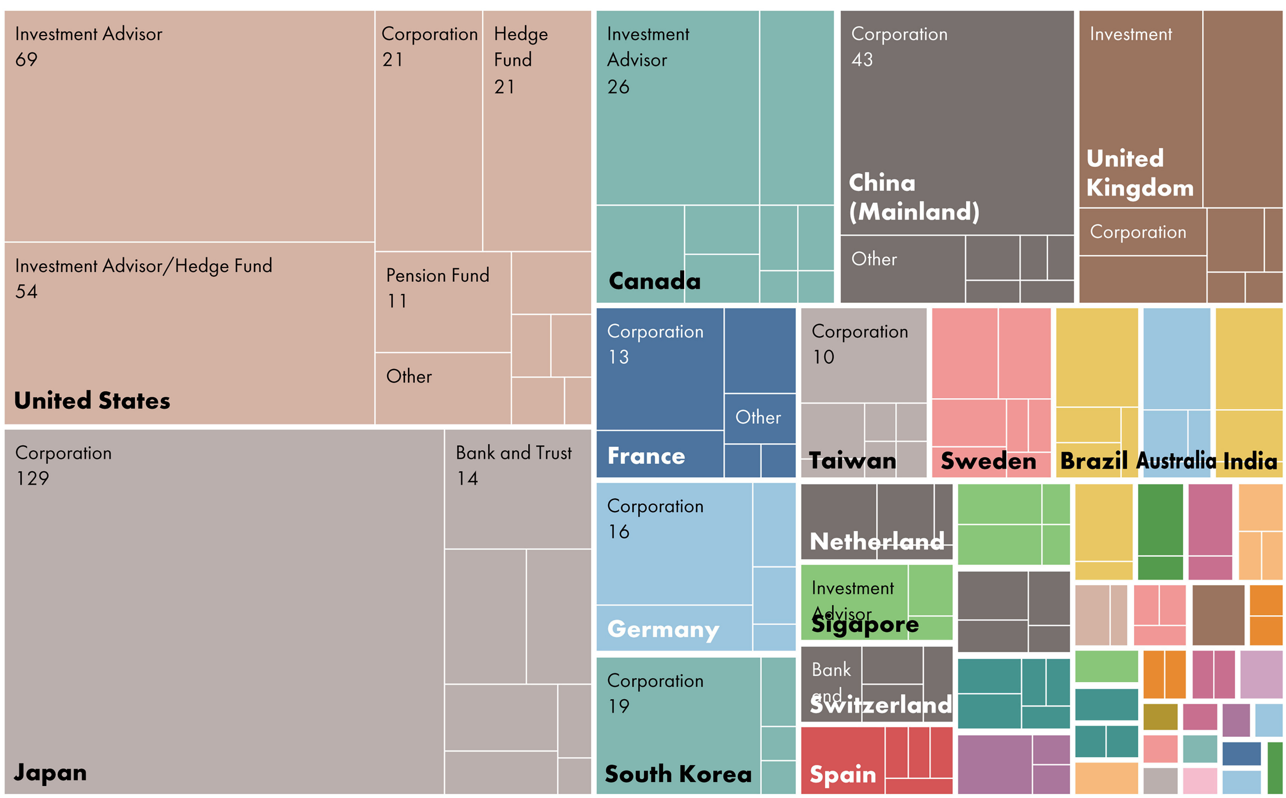


FIGURE 4 | Number of investors in the top 10% by country.

companies with the ESG combined scores. The currency unit of the total market price is the USD, which is converted by the Refinitiv Eikon system directly. The total market price represents the value of the invested company. A summary of the total market values of the companies is reported in **Table 1**.

METHODS

To detect the most influential investors in the ESG investment field, the cumulative impacts are calculated based on the ESG scores of invested companies, total market prices of invested companies, and investor history portfolio reports. The basic logic of calculating the cumulative impacts is to accumulate the products of the ESG scores and total market prices of the invested companies. However, in the real world, some large investors also invest in small investors. Therefore, we iterate the calculation to convey the impacts of invested companies on the ultimate investors. The cumulative impacts are calculated as follows:

$$CI_i = \sum_{j=1}^{m} \left(SCI_j \times ESGCO_j \times TPCO_j\right) + \sum_{k=1}^{n} \{ \sum_{l=1}^{p} [\min \left(SCI_k,\ SCI_l\right) \times \left(ESGCO_l \times TPCO_l\right)]\} \tag{1}$$

where CI_i denotes the cumulative impacts of investor i; SCI_j indicates the shares of investor i in invested company j; $ESGCO_j$

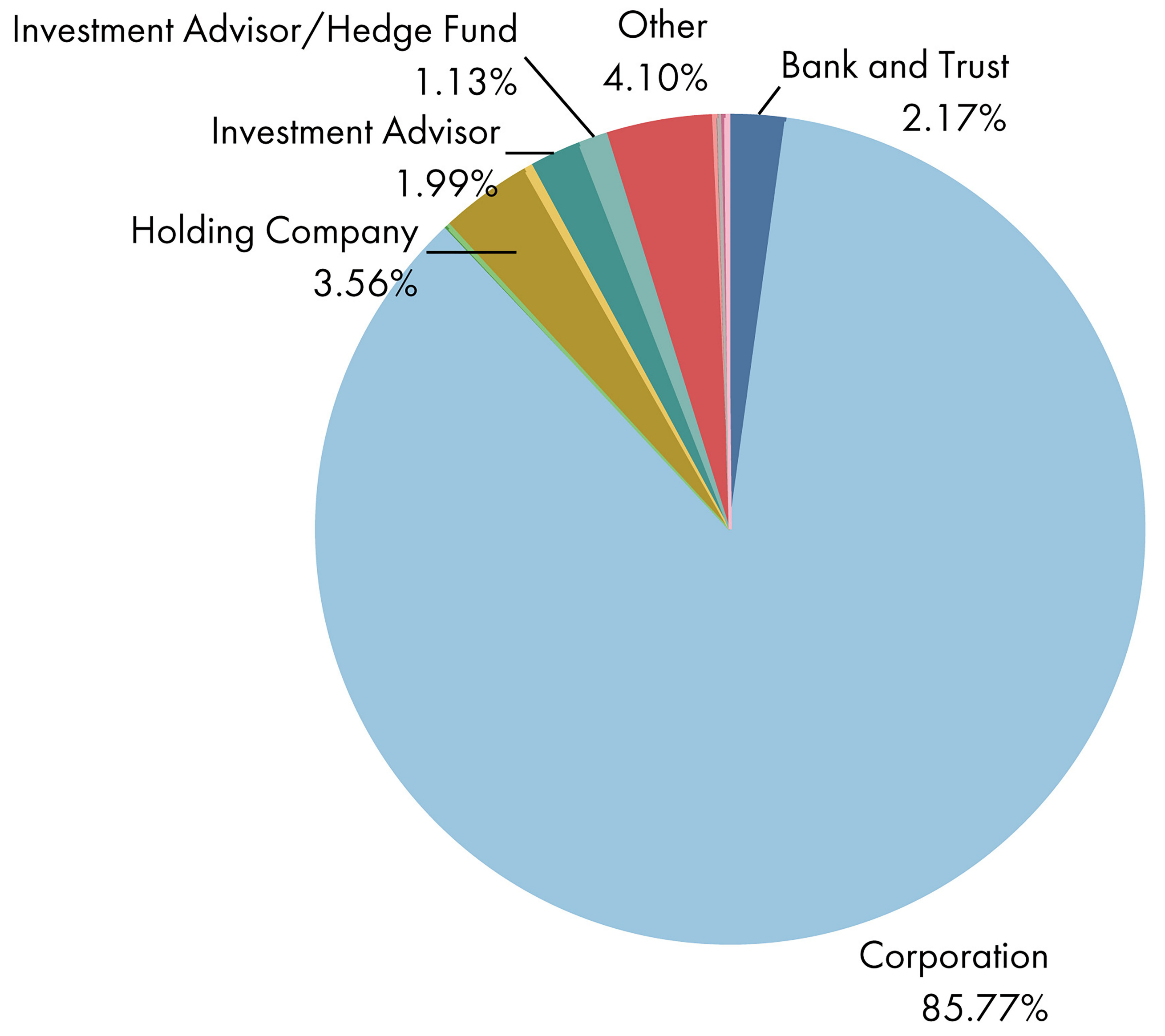


FIGURE 5 | Share of the ultimate ESG ownership by the investor type.

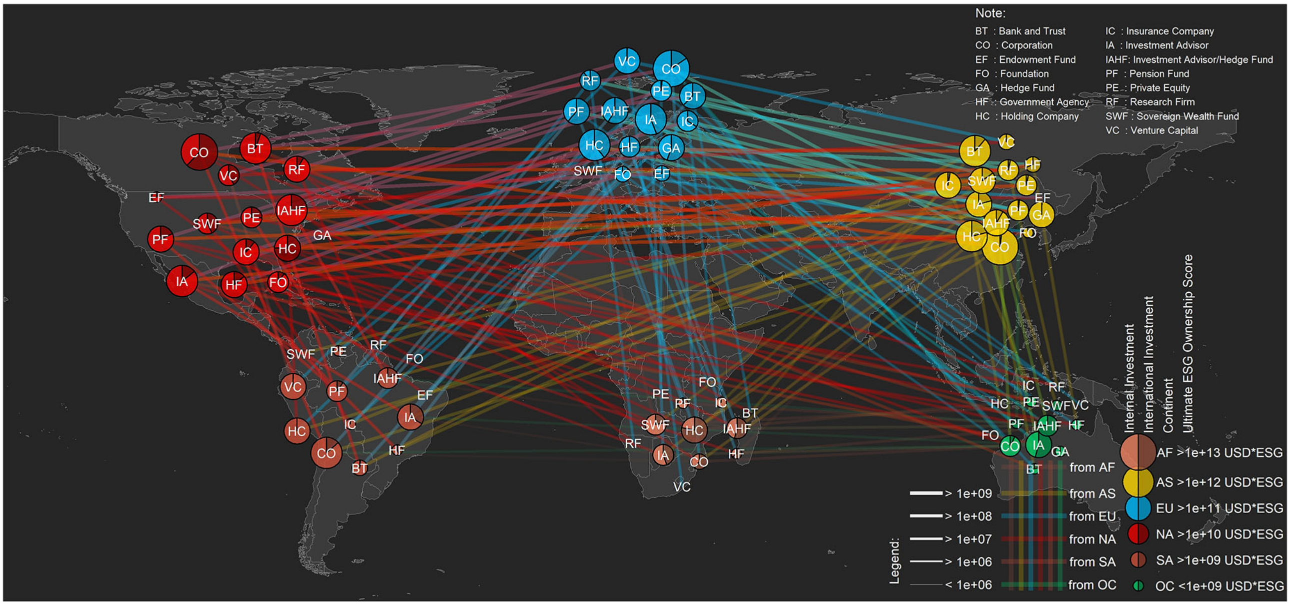


FIGURE 6 | ESG ownership network map by the investor type.

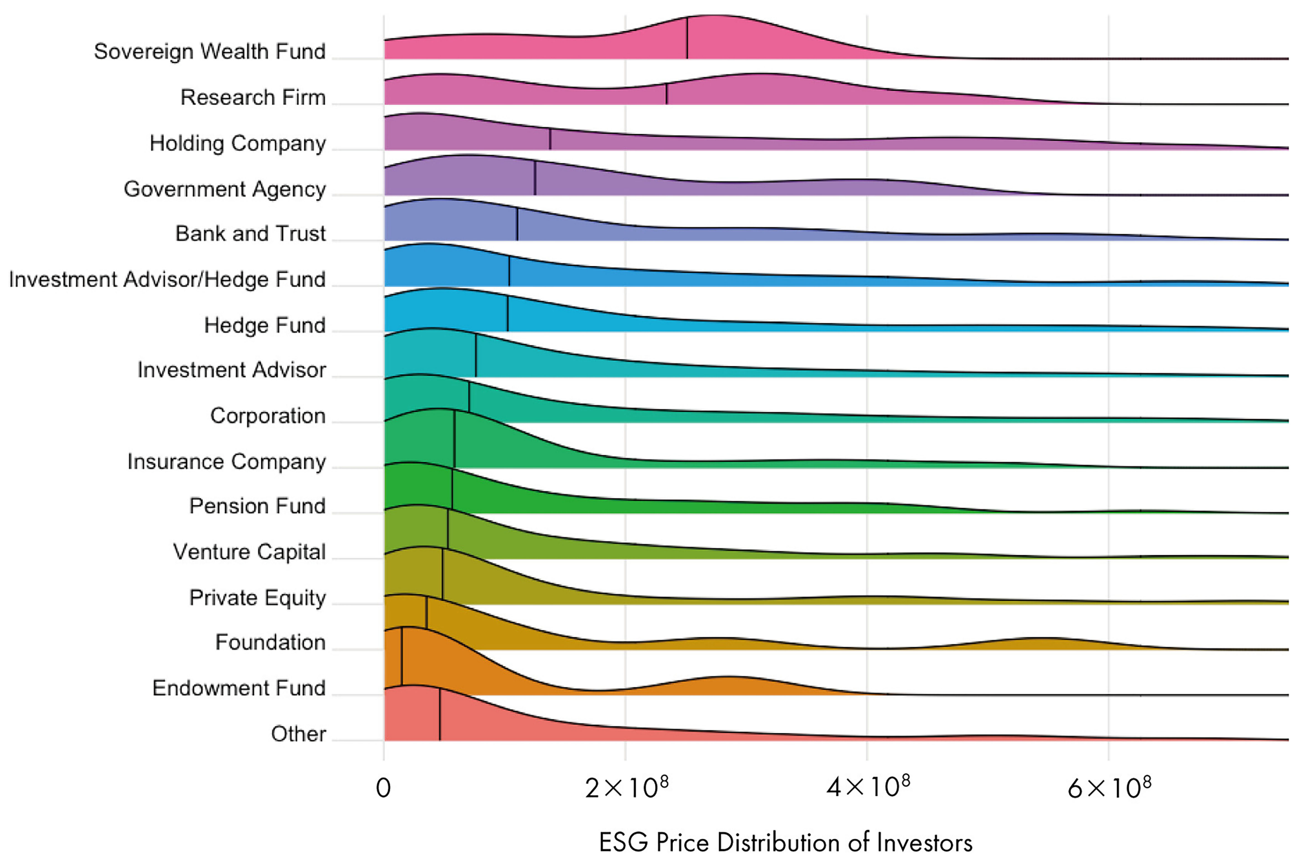


FIGURE 7 | Density ridgeline plots by the investor type.

represents the ESG scores of invested company *j*; $TPCO_j$ denotes the total market price of invested company *j*; *m* indicates the number of companies investor *i* has invested in; SCI_k is the shares of investor *i* (a large investor) in small investor *k*; SCI_l is the shares of small investor *k* in invested company *l*; $ESGCO_l$ denotes the ESG scores of invested company *l*; $TPCO_l$ represents the total market price of invested company *l*; *n* denotes the number of small investors that investor *i* has; and *p* indicates the number of companies that small investor *k* has invested in. We provide a simple example of the calculation process in **Figure 1**.

RESULTS

Based on the evaluations of the ultimate ESG ownership scores of investors, including institutional investors and corporations, this section presents the results by first revealing the patterns of the ESG ownership structure by the type of investor and nation–state political boundaries, followed by the results of the relation between the calculated ESG ownership scores and investors' ESG performances (e.g., PRI signatory status and investors' ESG ratings).

Ultimate ESG Owner: Trend by Country Level

Based on the cumulative ESG ownership calculated using the methods discussed in Section Data and Methods, **Table 2** reports the results of the total ultimate ESG ownership score of 7,957 investors at the country level (the investors' country distribution map is illustrated in Figure A of the **Appendix**). The investors' country-level ESG ownership network map is presented in **Figure 2**. **Figure 2** presents the share of the internal ESG investment and international ESG investment volumes in bubbles and denotes outflow investments through lines colored by the continent level. The size of the bubble represents the volume of the cumulative ultimate ESG ownership score.

As illustrated in **Figure 2**, the international ESG investment volume is much larger than the internal ESG investment volume in most North American countries, whereas the internal ESG investment volume has the largest share in Asia. The network figure demonstrates that the link between North American and Asian countries is especially strong. The strong link between European and Asian countries, with a stronger link from European to Asian countries, is noteworthy. The links from Asian to European countries are stronger than the links from Asian to other countries.

Figure 3 presents the total ultimate ESG ownership by country, with the size of the box representing the volume of the share of the total ultimate ESG ownership score. According to the results, Switzerland, Germany, and the US are the leading countries in terms of the volume of the share of the total ultimate ESG ownership score. In addition, the results reveal that the

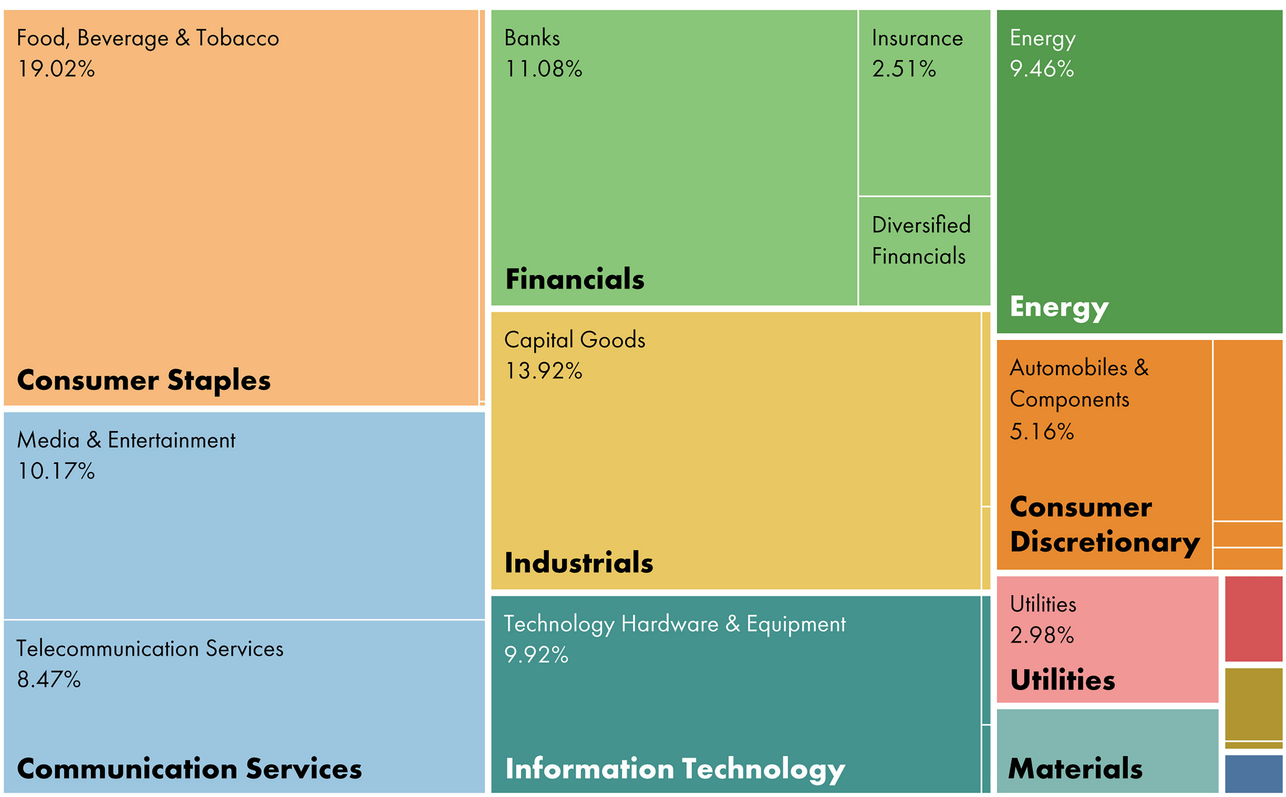


FIGURE 8 | Total ESG ownership by the industry group.

top 10% investors own approximately 98% of the total ultimate ESG ownership score. The number of investors in the top 10% by country is another important perspective to interpret the result, which is presented in **Figure 4**. Regarding the number of investors in the top 10%, the US, Japan, and China are the leading countries. Among the leading countries, there are differences in the types of investors leading (non-financial firms, such as corporations, and financial firms, including investment advisors, insurance companies, as well as banks and trusts). In the US, financial firms are the major players, whereas non-financial firms are the major players in Japan and China. Switzerland is not among the top countries in terms of the number of investors in the top 10% as the two leading investors in Switzerland (Novartis AG and Nestle SA) have a 27.3% share of the total ultimate ESG ownership score.

Ultimate ESG Owner: Trend by Investor Type and Industry

Based on the Refinitiv Business Classification, investors are classified into one of the following: corporation (CO), holding company (HC), insurance company, investment advisor (IA), bank and trust (BT), endowment fund, government agency (GA), foundation (FO), hedge fund, sovereign wealth fund (SWF), investment advisor/hedge fund (IAHF), venture capital (VC),

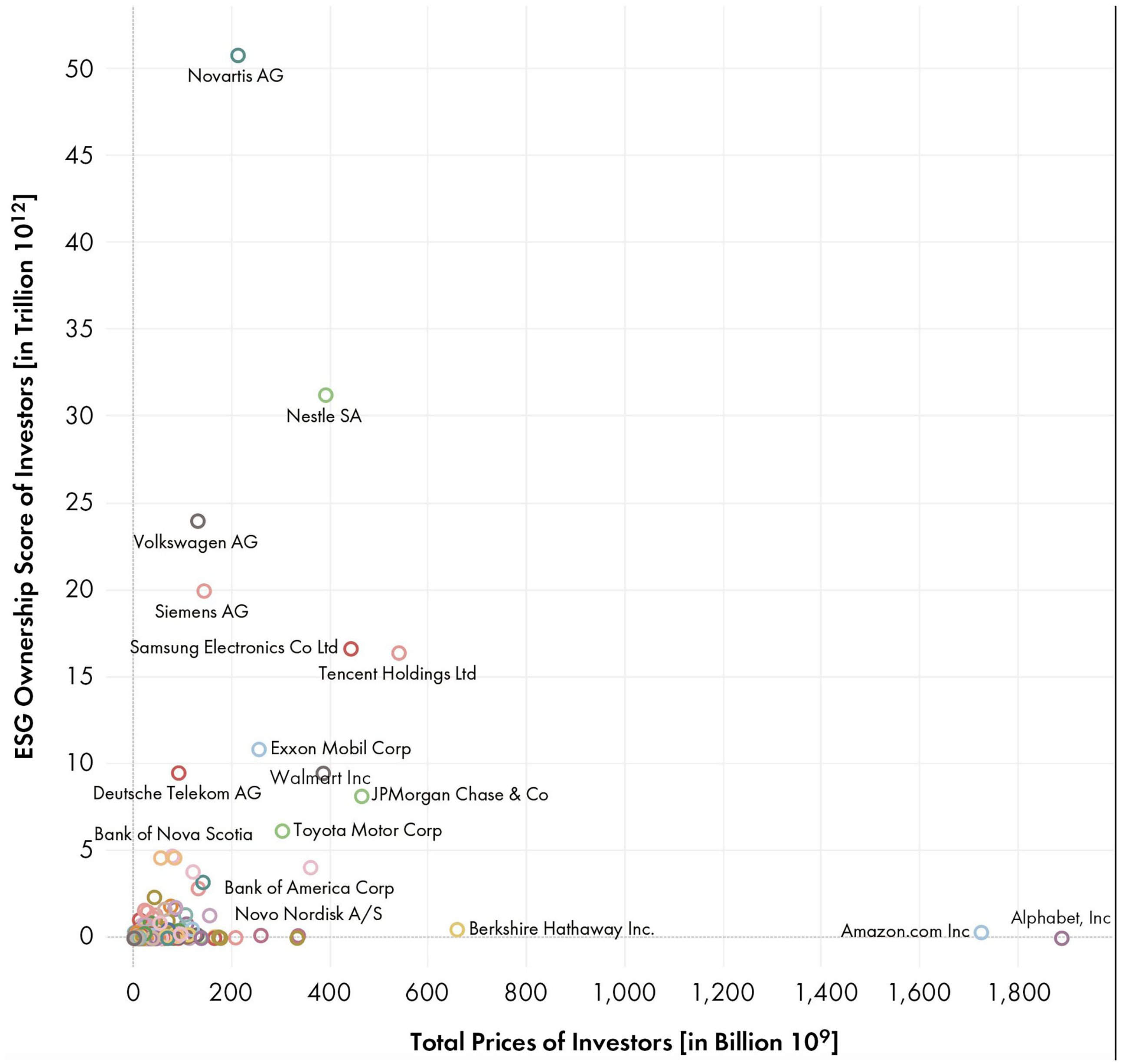


FIGURE 9 | Relation between investors' total market price and ESG ownership (n = 1,716).

research firm (RF), pension fund (PF), and private equity. **Figure 5** reports the share of the total ultimate ESG ownership by investor type. CO, HC, and IA are the three top investor types, but as depicted in the figure, CO has the most global shares (85.77%).

We observe different trends among the continents in terms of the cross-continent and internal ESG investment volumes, as presented in **Figure 6**. In Asian countries, the internal ESG investment comprises the most ESG investment volume, especially for CO, GA, and VC. However, in North America, the cross-continent ESG investment volume is larger than the internal volume for CO; in Europe, GA and PF actively conduct cross-continent investment, whereas CO has more internal ESG investment volume. Similar to the country-level network map presented in **Figure 2**, we observe strong links from North American investors to Asian companies, whereas between Europe and Asia, the links are more bidirectional. The links from European investors are relatively more diversified than those from other regions, with slightly stronger links to Asia.

Figure 7 illustrates the density ridgeline plots by the investor type, revealing that, on average, SWF and RF are performing better.

We provide a closer look of the leading industries in **Figure 8**. Each investor is classified into an industry group based on the Global Industry Classification Standard. As **Figure 8** demonstrates, consumer staples, financials, and industrials are

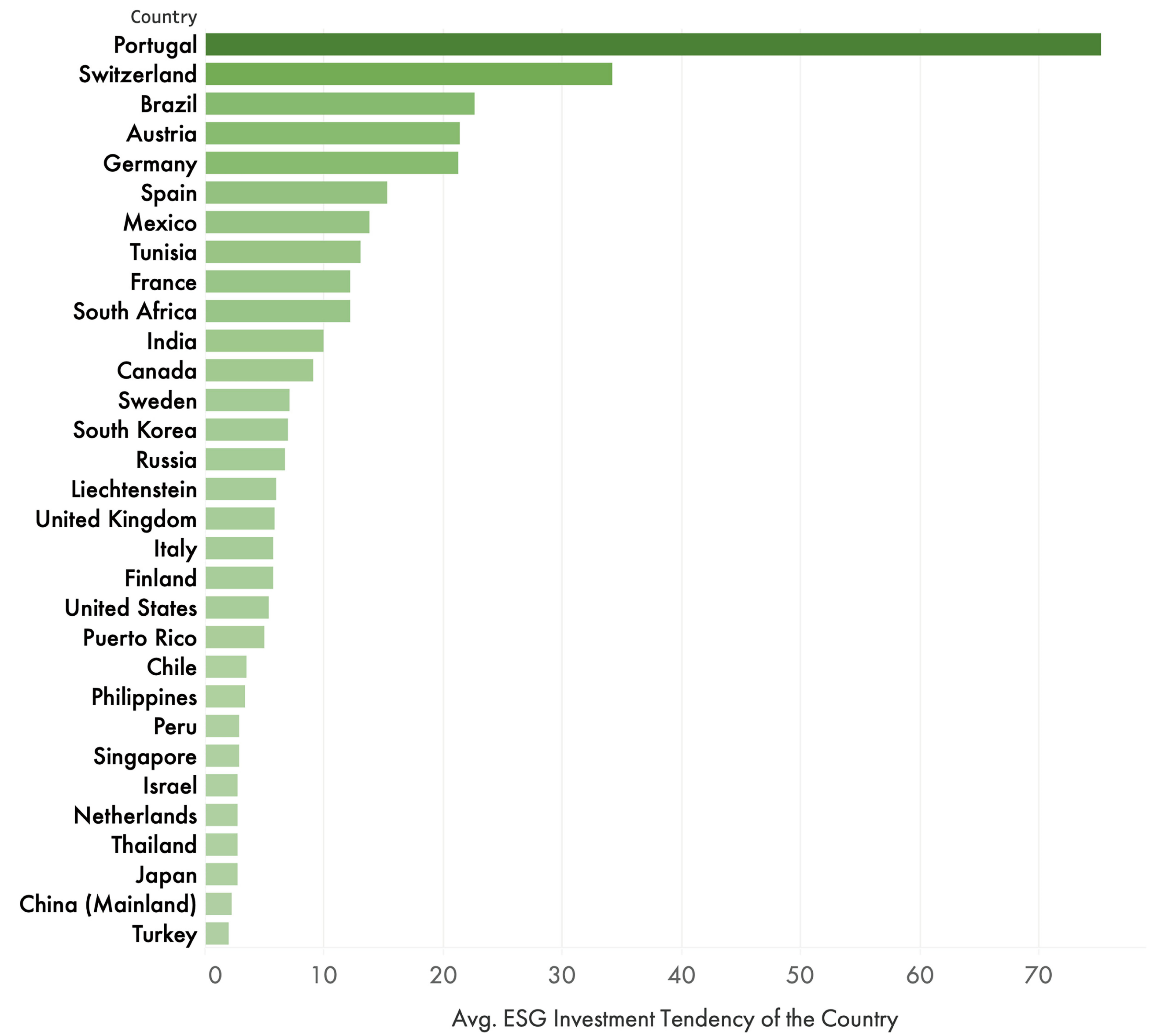


FIGURE 10 | Average ESG tendency by country.

the leading industry groups in terms of the volume of the share of total ultimate ESG ownership.

ESG Tendency: Total Market Price Adjusted ESG Ownership

As an investor with a larger total market price can make more investments than smaller investors, its ultimate ESG ownership score can be affected by investors' total market prices. However, as shown in **Figure 9**, the total market prices of the investors and their ultimate ESG ownership scores have a very weak correlation. The Pearson correlation is performed to analyze the relationship, and the correlation strength is defined as a correlation coefficient (R-value) of 0.27 ($p \leq 0.001$).

Although the correlation between the total market prices of investors and their ultimate ESG ownership scores is very weak, to reduce the potential bias owing to the size of the investor, we calculate the ESG tendency score by dividing the ultimate ESG ownership score by the total market price of each investor.

Figures 10, **11** report the results of the ESG tendency by country and investor type based on 1,716 investors. Here, two outliers (Aditya Birla Sun Life AMC Limited [2048] and Schweizerische Nationalbank [588]) are removed. By country, Portugal, Switzerland, and Brazil are the top countries, whereas in terms of the accumulated total ultimate ESG ownership score of the 7,957 investors, several leading countries, such as the US, Japan, and China, have low average ESG tendency. **Figure 11** demonstrates an interesting trend among the investor type—the average ESG tendency is high in VC, IAHF, and BT, whereas it is low in CO, which owns the most global shares.

Using the calculated ESG tendency score, **Figure 12** illustrates the relation between investors' ESG tendency and their own ESG score rated by Refinitiv. The Pearson correlation is performed to analyze the relation between investors' ESG tendency and their own ESG score, and the correlation strength is defined as a correlation coefficient (R-value) of −0.05 ($p \leq 0.155$). **Figure 12** shows that some investors, such as Siemens AG, Deutsche Telecom AG, Novartis AG, and Volkswagen AG, have relatively high ESG scores and a high ESG tendency. However, overall, the figure and results reveal that investors with high ESG ratings do not necessarily have higher incorporation of ESG investment strategies.

ESG Ownership and PRI Signatory

To examine the relation between the PRI signatory status and ESG ownership, we classify the investors into PRI and non-PRI signatory investors based on the PRI (2022). **Figure 13** reveals that the PRI signatories have higher ESG ownership with higher mean value, and **Figure 14** reveals that, on average, the PRI signatories' average ESG tendency, that is, the total market price adjusted ESG ownership, is higher than that of the non-PRI signatories.

In addition, we perform the Welch two-sample t-tests to determine if the PRI signatory investors' ultimate ESG ownership scores differ significantly from those of the non-PRI signatory investors. The results of the two-sample t-test reveal statistically significant difference in the mean ultimate ESG ownership score of the PRI and non-PRI signatory investors, with $t(\text{df} = 7609) = -3.23$, $p = 0.0012$, 95% CI as the difference in means [−5.24e10, −1.28e10]. The figures and result imply that investors who commit to invest responsibly certainly do so in practice. This result is mostly consistent with the recent study of Brandon et al. (2021), which revealed that, overall, signatory investors have better portfolio-level ESG scores, barring the US.

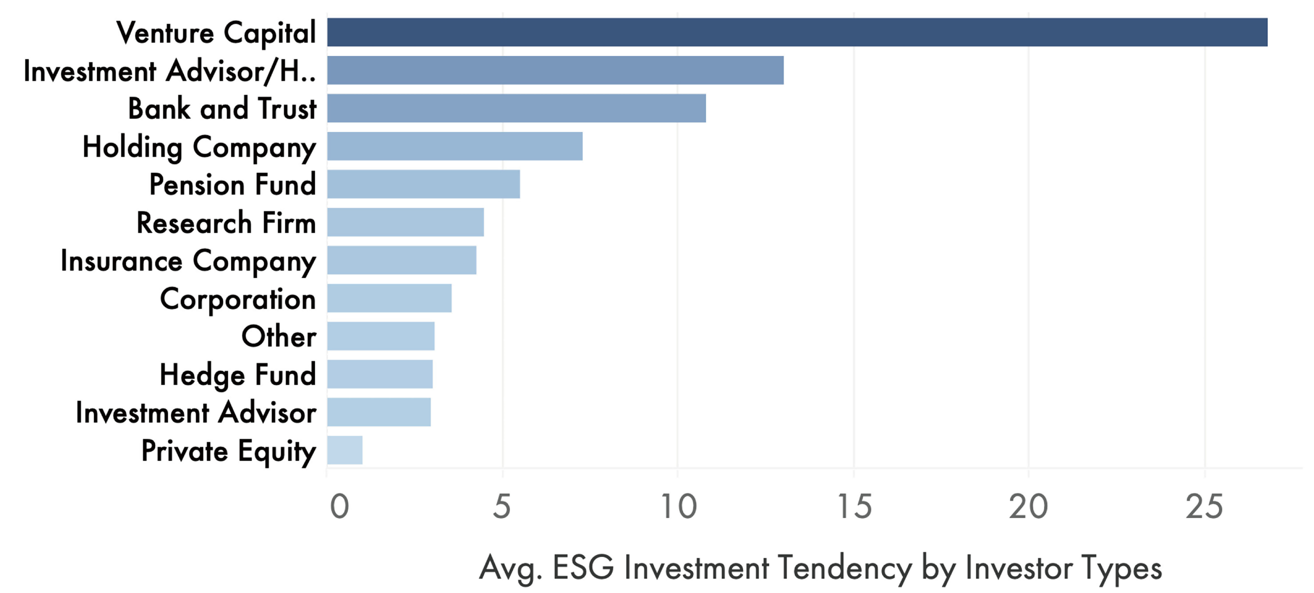


FIGURE 11 | Average ESG tendency by the investor type.

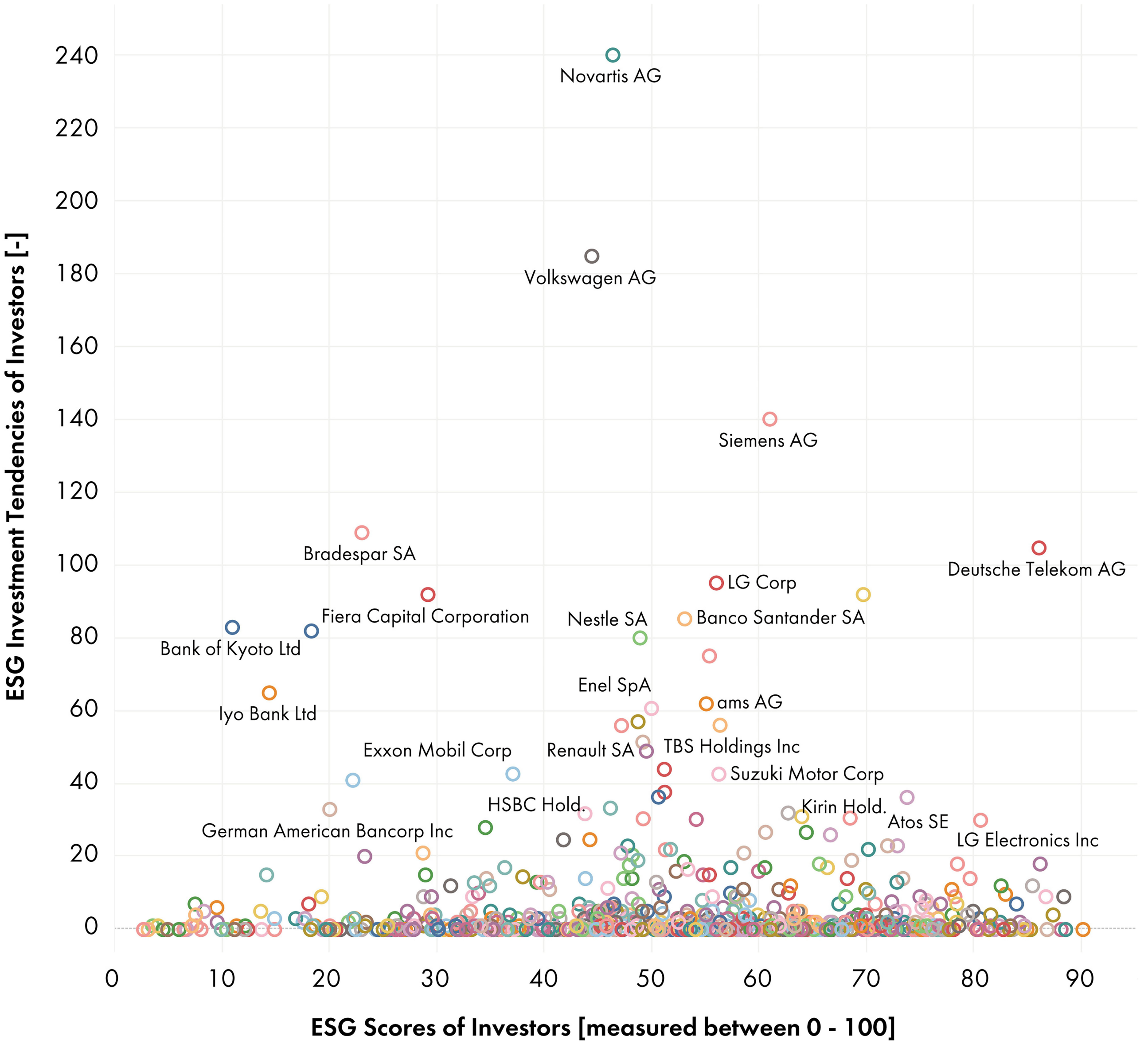


FIGURE 12 | Relation between investors' ESG scores and ESG tendency.

DISCUSSION AND CONCLUSION

Businesses are increasingly expected to understand and manage their exposure to ESG risks associated with their investments in financial assets, strategic partners, and business portfolio interests. Although numerous studies have investigated the impact of the ESG scores of companies on their financial returns and the trend in the ESG investment strategies, studies that identify the major players and regional trend in ESG investment by considering the direct and indirect ownership structure are scarce. Faccio and Lang's (2002) framework considers the cumulative impacts of ownership using the total market price of invested companies and investor history portfolio report by iterating the calculation to convey the impacts of invested companies to ultimate investors. The study expands this comprehensive approach to the field of ESG investment to address the research gap and detects the most influential investors in the ESG investment field and the ownership structure by considering the direct and indirect ESG ownership.

Based on the results, the patterns by the type of investor and nation–state political boundaries in the context of direct and indirect ownership chains are examined. We identify strong links between several countries, such as a strong link between North America and Asian countries, and differences in the pattern of ESG investment in terms of regional inflow and outflow (i.e., the international ESG investment volume is much larger than

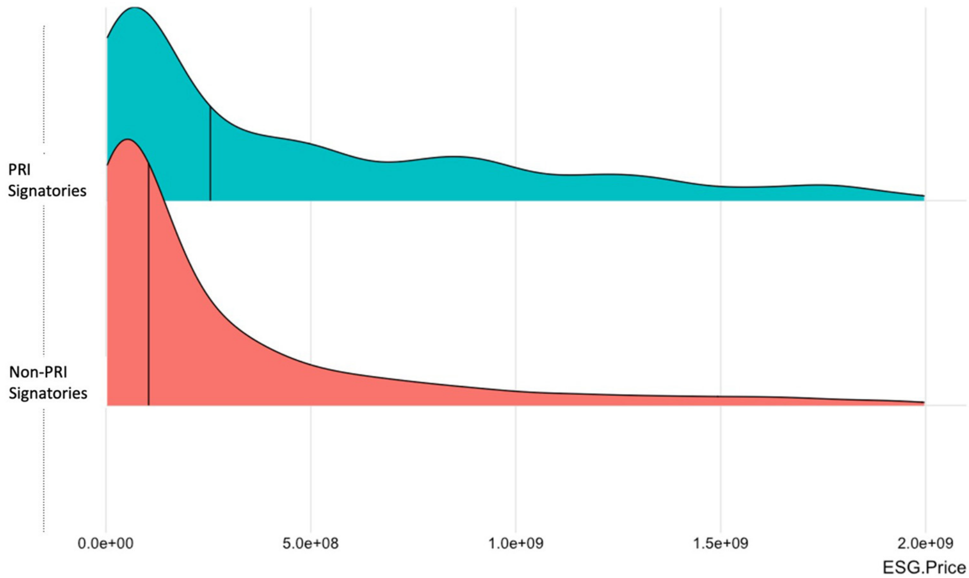


FIGURE 13 | Density ridgeline plots of ESG ownership score by the PRI signatory status.

FIGURE 14 | Average ESG tendency by the PRI signatory status.

the internal ESG investment volume in most North American countries, whereas the internal ESG investment volume has the largest share in Asia). Through the analysis, we find that the top 10% investors own approximately 98% of the total ultimate ESG ownership score and the key players differ among the leading countries. In the US, financial firms are the major players, whereas non-financial firms are the major players in Japan and China. Furthermore, the results of the study confirm the important role of corporations in ESG investment, which owns the most global shares (85.77%) measured by the ESG ownership score, considering the direct and indirect ownership chains. Even minority shareholders may be directly linked to adverse environmental and social impacts directly or indirectly caused by investee companies in their portfolios (PRI, 2017). Therefore, it is important for investors to undertake ESG risk-based due diligence and consider ESG risks in their investment processes. Moreover, investors can manage and influence the responsible business conduct of the investee companies through direct and indirect ownership (OECD, 2017). The results of this study highlight the need of wider implementation of investments considering ESG risks.

Using the results of the ultimate ESG ownership analysis, this study also investigated the relation between the calculated ESG ownership score and Investor's ESG commitment and ESG performance (i.e., the PRI signatory status, investors' ESG ratings). The study by Brandon et al. (2021) showed that the PRI signatory institutional investors have better ESG ownership score (calculated based on the direct ownership data) in general, barring the US. Our analysis includes institutional investors and other financial and non-financial corporations, and the results demonstrating that the PRI signatory investors have better ESG ownership score than non-signatory investors are consistent with Brandon et al. (2021). The analysis of the relation between the calculated ESG ownership scores and investors' ESG performances (ESG ratings provided by Refinitiv) reveals that investors with high ESG ratings do not necessarily have higher incorporation of the ESG investment strategies. This result supports the findings of previous studies examining fund flow reactions to ESG ratings (Hartzmark and Sussman, 2019) by providing the evidence that it is difficult to target the actual flow and ownership structure of ESG investment through ESG ratings of the investors themselves.

Although evaluating the ESG ownership score by considering the direct and indirect ESG ownership structure and covering both institutional investors and other corporations expand the literature on ESG investment and ownership structure in various ways, our study only evaluates the score of the latest year owing to difficulties in data collection and the time required to calculate the scores for multiple years. Future studies can address this issue by constructing the ESG ownership score as time-series data, which would help in investigating the causal relationship between the ESG ownership score and investors' ESG performances, such as the PRI signatory status and investors' ESG ratings, more accurately.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AK and CL carried out the experiment. ST conducted additional analyses using the calculated data. AK, CL, ST, and TG wrote the manuscript with support from SM. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to express our gratitude toward Mr. Okita and Mr. Ike from Vector Group for providing their insights and expertise on the ESG investment market during our discussions. This paper is based on results obtained from a project, JPNP14026, commissioned by the New Energy and Industrial Technology Development.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsus.2022.909239/full#supplementary-material

REFERENCES

Bauer, R., Derwall, J., and Otten, R. (2007). The ethical mutual fund performance debate: new evidence from Canada. *J. Bus. Ethics*. 70, 111–124. doi: 10.1007/s10551-006-9099-0

Brandon, R. G., Glossner, S., Krueger, P., Matos, P., and Steffen, T. (2021). "Do responsible investors invest responsibly?," in *ECGI Working Paper Series in Finance*, (Brussels: European Corporate Governance Institute).

Bryan, A., Choy, C. F. A. J., Lamont, K., and Sanzgiri, Z. (2020). *Passive Sustainable Funds: The Global Landscape 2020 - Choice Expands as Products and Assets Double in Three Years*. Chicago, IL: Morningstar.

Claessens, S., Djankov, S., and Lang, L. H. P. (2000). The separation of ownership and control in East Asian corporations. *J. Financ. Econ*. 58, 81–112. doi: 10.1016/S0304-405X(00)00067-2

Cortez, M. C., Silva, F., and Areal, N. (2009). The performance of European socially responsible funds. *J. Bus. Ethics*. 87, 573–588. doi: 10.1007/s10551-008-9959-x

Faccio, M., and Lang, L. H. (2002). The ultimate ownership of Western European corporations. *J. Financ. Econ*. 65, 365–395. doi: 10.1016/S0304-405X(02)00146-0

Funk, C. M., and Hento, E. (2022). *ESG Implications of the Russia-Ukraine War [Commercial Sector]*. State Street Global Advisors–Insights. Available online at: https://www.ssga.com/se/en_gb/institutional/ic/insights/esg-implications-of-the-russia-ukraine-war (accessed March 13, 2020).

Global Sustainable Investment Alliance (2021). *Global Sustainable Investment Review 2020*. Available online at: http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf

Hartzmark, S., and Sussman, A. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *J. Fin*.74, 2789–2837. doi: 10.1111/jofi.12841

Kenton, W., James, M., and Kvilhaug, S. (2021). *What Is Intercorporate Investment? Investopedia*. Available online at: https://www.investopedia.com/terms/i/intercorporate-investment.asp (accessed March 12, 2021).

Khan, S. A. R., Ponce, P., Tanveer, M., Aguirre-Padilla, N., Mahmood, H., and Shah, S. A. A. (2021a). Technological innovation and circular economy practices: business strategies to mitigate the effects of COVID-19. *Sustainability* 13:8479. doi: 10.3390/su13158479

Khan, S. A. R., Yu, Z., Umar, M., and Tanveer, M. (2021b). Green capabilities and green purchasing practices: a strategy striving towards sustainable operations. *Bus. Strategy Environ*. 31, 1719–1729. doi: 10.1002/bse.2979

La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (1999). Corporate ownership around the world. *J. Fin*. 54, 471–518. doi: 10.1111/0022-1082.00115

MorningStar (2022). *Country Risk*. Available online at: https://connect.sustainalytics.com/hubfs/INV%20-%20Reports%20and%20Brochure/Product%20Brochures/Country%20Risk%20Rating.pdf

OECD (2017). *Responsible Business Conduct for Institutional Investors: Key Considerations for Due Diligence Under the OECD Guidelines for Multinational Enterprises*. Paris: OECD.

OECD (2021). *ESG Investing and Climate Transition: Market Practices, Issues and Policy Considerations*. Paris: OECD.

PRI (2017). *Managing ESG Risk in the Supply Chains of Private Companies and Assets*. London: PRI. https://www.unpri.org/download?ac=1894

PRI (2022). *Signatory Directory*. Available online at: https://www.unpri.org/signatories/signatory-resources/signatory-directory (accessed March 2022).

Silver, C., Chen, J., and Kagan, J. (2021). *Accounting for Intercorporate Investments: What You Need to Know*. Investopedia. Available online at: https://www.investopedia.com/articles/fundamental-analysis/11/accounting-intercorporateinvestment.asp# (accessed March 13, 2021).

Widyawati, L. (2020). A systematic literature review of socially responsible investment and environmental social governance metrics. *Bus. Strategy Environ*. 29, 619–637. doi: 10.1002/bse.2393

Yoo, S., Keeley, A. R., and Managi, S. (2021). Does sustainability activities performance matter during financial crises? Investigating the case of COVID-19. *Energy Policy* 155:112330. doi: 10.1016/j.enpol.2021.112330

Author Disclaimer: Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

Conflict of Interest: TG was employed by Industrial Ecology Consultants.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Sustainability

TYPE Systematic Review
PUBLISHED 28 July 2022
DOI 10.3389/frsus.2022.920876

OPEN ACCESS

EDITED BY
Roberta Salomone,
University of Messina, Italy

REVIEWED BY
Teresa Maria Gulotta,
University of Messina, Italy
Christina Wulf,
Helmholtz Association of German Research Centres (HZ), Germany

*CORRESPONDENCE
Dyah Ika Rinawati
dyah.rinawati.272@s.kyushu-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 15 April 2022
ACCEPTED 30 June 2022
PUBLISHED 28 July 2022

CITATION
Rinawati DI, Keeley AR, Takeda S and Managi S (2022) Life-cycle assessment of hydrogen utilization in power generation: A systematic review of technological and methodological choices. *Front. Sustain.* 3:920876.
doi: 10.3389/frsus.2022.920876

Life-cycle assessment of hydrogen utilization in power generation: A systematic review of technological and methodological choices

Dyah Ika Rinawati[1,2]*, Alexander Ryota Keeley[1,3], Shutaro Takeda[1] and Shunsuke Managi[1,3]

[1]Department of Civil Engineering, Kyushu University, Fukuoka, Japan, [2]Department of Industrial Engineering, Diponegoro University, Semarang, Indonesia, [3]Urban Institute, Kyushu University, Fukuoka, Japan

Interest in reducing the greenhouse gas emissions from conventional power generation has increased the focus on the potential use of hydrogen to produce electricity. Numerous life-cycle assessment (LCA) studies of hydrogen-based power generation have been published. This study reviews the technological and methodological choices made in hydrogen-based power generation LCAs. A systematic review was chosen as the research method to achieve a comprehensive and minimally biased overview of hydrogen-based power generation LCAs. Relevant articles published between 2004 and 2021 were identified by searching the Scopus and Web of Science databases. Electrolysis from renewable energy resources was the most widely considered type of hydrogen production in the LCAs analyzed. Fuel cell technology was the most common conversion equipment used in hydrogen-based electricity LCAs. A significant number of scenarios examine the use of hydrogen for energy storage and co-generation purposes. Based on qualitative analysis, the methodological choices of LCAs vary between studies in terms of the functional units, allocations, system boundaries, and life-cycle impact assessment methods chosen. These discrepancies were likely to influence the value of the environmental impact results. The findings of the reviewed LCAs could provide an environmental profile of hydrogen-based electricity systems, identify hotspots, drive future research, define performance goals, and establish a baseline for their large-scale deployment.

KEYWORDS

life-cycle assessment, hydrogen, power generation, environmental impact, electricity

Introduction

Carbon dioxide (CO_2) emissions from fossil fuels have been increasing since the beginning of the Industrial Era. They became the main contributor of anthropogenic emissions to the air from around 1950 and their relative share has kept rising until the present (Friedlingstein et al., 2022). Global CO_2 emission reached a historical record of

33.1 gigatons (Gt) in 2018 (IEA, 2021a), at which time, coal-fired power generation alone emitted 10 Gt of CO_2 (IEA, 2019a). Renewable energy adoption has been accelerating and is expected to contribute to approximately 30% of all CO_2 emissions reduction that occurs between 2017 and 2060 (IEA, 2017). Renewable energy technologies are employed most intensively in the power sector, where it is intended to accelerate the transition to low carbon with applications in the transportation, building, and industry sectors (IEA, 2017). However, power generation with renewable energy sources like solar photovoltaic (PV) and wind power depends on the weather conditions and the time of day (i.e., day or night) and thus the same amount of electricity cannot be distributed constantly. As a consequence, flexible and reliable power generation relying on various energy sources is required to provide electricity when using renewable sources directly is not feasible (Peppas et al., 2021). Since hydrogen is a chemical energy carrier, it may be used as a storage option to balance seasonal fluctuations in providing electrical energy (IEA, 2019b; Peppas et al., 2021).

Numerous strategies have been prepared worldwide to encourage the deployment of hydrogen technologies, involving all stakeholders—government, business developers, investors, and citizens. As of March 2022, 21 governments have released hydrogen strategies, 27 countries have national hydrogen strategies in preparation, and 34 discuss their initial policies and pilot projects (Work Energy Council, 2022). In the IEA's Net Zero by 2050: A Roadmap for the Global Energy Sector, sectoral and technology milestones to guide the global journey to net-zero have been provided, where power is the primarily targeted sector, and hydrogen is one of the critical pillars of decarbonizing the global energy system (IEA, 2021b,c). Recently, Clean Hydrogen Joint Undertaking (JU) – a program intends to encourage research and innovation (R&I) initiatives in the European Union in clean hydrogen solutions and technologies, presents a set of prioritized actions divided into three pillars. The Clean Hydrogen JU Pillar 3, 'Hydrogen End Uses: Clean heat and Power', aims to support European supply chain actors in developing a portfolio of clean, renewable and flexible heat and power generation solutions for all end user's needs (Clean Hydrogen JU, 2022). This program will aid in developing several hydrogen technologies, which are currently undergone on R&I phase but are projected to contribute to making climate neutrality achievable by 2050.

The use of life cycle assessment (LCA) in R&I stages has gained attention in recent years (Cucurachi et al., 2018) due to its capability to assess environmental impact throughout a product's life cycle. The advantage of LCA is that it evaluates an entire product comprehensively, which prevents any suboptimization that may arise from focusing on only a few processes. LCA enables the comparison of potential environmental impacts from various alternatives (Varun et al., 2009). The European Union even requires LCA as an essential part of the R&I projects for funding proposals (Clean Hydrogen JU, 2022). Regarding emerging hydrogen technology, the FC-HyGuide document, a guideline for conducting LCA of hydrogen technologies, has been provided and recommended by European Union (Lozanovski et al., 2011; Masoni and Zamagni, 2011). Other than FC-HyGuide guidelines, there is a new ongoing EU project developing SH2E LCSA (Life Cycle Sustainability Assessment) guidelines for fuel cell and hydrogen (FCH) systems, including guidelines for LCA (SH2E, 2022).

With the growing number of the hydrogen LCA studies, numerous review paper of hydrogen LCAs have been published. Bhandari et al. (2014) reviewed 21 studies that addressed LCA hydrogen production technologies in which an aggregate comparison from an ecological perspective was discussed. Valente et al. (2016) reviewed more LCA studies on hydrogen energy systems. Koj et al. (2019) performed a review of 32 LCA studies on Power-to-X, where fuels and final use for transport applications were discussed most frequently instead of hydrogen for power generation purposes. Rinawati et al. (2022) placed a stronger emphasis on hydrogen for mobility use LCAs. These studies succeeded in identifying the relevant methodological trends. However, there is a lack of detailed overview about technical aspects on hydrogen for power generation application.

The objective of this review is to synthesize hydrogen for electricity generation LCAs. Our primary goal is to provide an overview of the technological aspects of addressed hydrogen uses for power generation LCAs. Our secondary goal is to analyze the methodological choices in preparing qualitative analyses of hydrogen-based power generation LCAs. Our third goal is to present a quantitative analysis of the environmental impacts of hydrogen-based power generation LCAs.

Materials and methods

This systematic review follows general systematic review principles (Tranfield et al., 2003) and "the STARR-LCA" methodology, which is a standardized technique for assessing and reporting LCA studies (Zumsteg et al., 2012). This methodology is discussed in this section.

Some delimitation criteria were applied in the selection of potential articles. The first delimitation is the origin of publications; only articles published in peer-reviewed journals were considered. The second criterion is year of publication; only articles published in 2000 or later were considered for the review since the first set of International Organization for Standardization (ISO) standards on LCA was completed in 2000. The third criterion is the language of publication; only articles in the English language were considered in this systematic review.

The procedure for selecting potential articles started with a search of the Scopus and Web of Science (WOS) databases using specific keywords and Boolean operators. The combination of keywords and operators was "Life-Cycle Assessment OR

Life-Cycle Analysis AND hydrogen AND Power Generation OR electricity Generation." The initial search generated 400 articles, which included 179 articles from Scopus and 221 publications from the WOS. The selection of articles comprising the systematic review followed several stages, which were conducted according to the method used by Rinawati et al. (2022), as shown in Figure 1. Firstly, the raw data were filtered, considering only papers published in journals (four publications of non-peer-reviewed journal articles were excluded). Then, a series of duplication checks were performed using Excel, which were manually rechecked; 101 duplications were removed using Excel, and seven repeated articles were recognized and eliminated using manual inspection.

Following Becheikh et al. (2006), a two-step practical screening was carried out with a set of inclusion and exclusion criteria adopted from Valente et al. (2016). All papers that reported the environmental implications of hydrogen for power generation applications based on the LCA method were included in this systematic review. We included only process-based LCA studies since they provide a precise and complete basis for analysis. Any variation in the technical features (e.g., feedstock, primary energy source, hydrogen production method, hydrogen storage, and hydrogen-based electricity conversion technology) or methodological choices (e.g., functional units, system boundaries, geographical scope, and impact assessment method) identified a unique case in this systematic review. The exclusion criteria of the articles included articles that (i) did not cover the hydrogen production phase, (ii) had a large number of case studies (more than 30), (iii) performed an environmental evaluation that did not employ the LCA method, (iv) involved hydrogen as a by-product from a background process, and (v) in which the conversion system configuration was combined with another energy storage system like batteries. The inclusion and exclusion criteria were applied in the title and abstract analysis for the first practical screening as well as during the analysis of the full text for the second screening. Based on this process, 228 articles were excluded from the first screening. Subsequently, four relevant articles were found in the literature during the second screening, so a total of 68 potential articles were identified. After the second screening, 47 articles were excluded. This left a total of 21 studies that met all the criteria for inclusion.

The most important information about the methodological choices and technical aspects of the selected hydrogen-based power generation LCA studies was extracted using a modified coding scheme (Muench and Guenther, 2013; Rinawati et al., 2022), as shown in the Supplementary material. We identified a case study or scenario for each variation in technological aspects such as hydrogen production method, feedstock, primary energy source, hydrogen storage, power conversion technology, and geographical scope. The predictive scenarios for sensitivity analysis were not recognized as unique case studies. Based on the data extracted from the 21 LCA studies that met the inclusion criteria, 71 case studies were identified (see Table 1). Furthermore, qualitative analysis of technological aspects in hydrogen-based power generation LCAs were performed based on the number of case studies described in Sections Hydrogen production methods, Hydrogen feedstock and energy sources, Hydrogen storage, Hydrogen-based power generation conversion technologies and applications, and Geographical context. Qualitative analysis of methodological choices were referred to the number of articles due to their homogeneity in a single article. These results provided in Sections System boundary, Functional unit, Allocation, Life-cycle inventory, and Life-cycle impact assessment method.

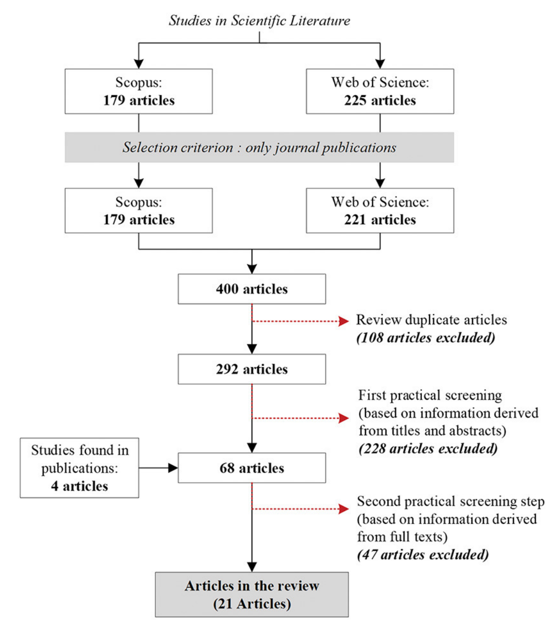


FIGURE 1
Illustration of the procedure for selecting articles for a systematic review.

Results

Scope of the reviewed studies

This section explains the scope of the reviewed hydrogen-based power generation LCAs, and provides an overview of the hydrogen production methods examined, the hydrogen feedstocks and energy sources considered, the hydrogen-based electricity conversion technologies addressed, and the geographical contexts analyzed.

Hydrogen production methods

Hydrogen can be produced by numerous techniques such as thermochemical, electrochemical, photochemical,

TABLE 1 The list of hydrogen-based electricity LCAs under reviewed (2004–2021).

Study	Year	Number of cases	Geographical scope	Hydrogen pathway	Hydrogen storage	Application	Power conversion technology
Lunghi et al. (2004)	2004	2	Italy	SR (landfill gas and natural gas)	Not specified	Power grid	MCFC
Khan et al. (2005)	2005	1	Canada	Electrolysis (wind power)	Not specified	Power grid	PEMFC
Spitzley et al. (2007)	2007	2	USA	Electrolysis (hydropower); SR (natural gas)	Liquid	Microgrid	PEMFC
Melamu and von Blottnitz (2009)	2009	2	South Africa	APR (maize) with and without heat integration	Compressed gas	CHP	Gas turbine
Strazza et al. (2010)	2010	3	Italy	Cracking; electrolysis; reforming	Not specified	Auxiliary power systems on-board ships	SOFC
Sevencan and Çiftcioglu (2013)	2013	6	Turkey	Electrolysis (solar PV, wind power, solar PV and wind power)	Compressed gas; metal hydride	Backup system in mobile home	Fuel cell
Mori et al. (2014)	2014	1	Slovenia	Electrolysis (renewable electrolysis)	Compressed gas	Backup power (UPS)	PEMFC
Oliveira et al. (2015)	2015	4	Belgium	Electrolysis (electricity mix UCTE 2004, Belgium 2011, PV mix, wind power)	Not specified	Power grid	PEMFC
Valente et al. (2015)	2015	1	Italy	Electrolysis (hydropower)	Metal hydride	Power grid	PEMFC
Walker et al. (2017)	2017	4	Canada	Electrolysis (electricity mix)	Not specified	Power grid	Gas turbine; combined cycle
Di Marcoberardino et al. (2018)	2018	12	Italia; Germany	SR and ATR (natural gas)	Compressed gas	CHP	PEMFC
Stropnik et al. (2018)	2018	2	Norway; Morocco	Electricity (electricity mix)	Compressed gas	Backup power (UPS)	PEMFC
Suwanmanee et al. (2018)	2018	4	Thailand	Gasification (biomass)	Compressed gas	Decentralized power generation	PEMFC
Ozawa et al. (2019)	2019	8	Japan	Electrolysis (renewable electricity)	Liquid; methylcyclo-hexane	Power grid	Combined cycle
Bicer and Khalid (2020)	2020	2	Europe	Electrolysis (wind power); SR (natural gas)	Not included	CHP	SOFC
Mori et al. (2021)	2020	1	Spain	Electrolysis (Hydropower)	Compressed gas	Microgrid	PEMFC
(Rossi et al., 2020)	2020	4	Italy	PEM electrolysis (solar PV)	Compressed gas	Nano-grid	PEMFC
Shimizu et al. (2020)	2020	8	Japan	Electrolysis (solar PV); SR (Fuel gas)	Compressed gas	Co-generation for household use	Fuel cell
Peppas et al. (2021)	2021	1	Greece	Electrolysis (combination of solar, wind power, and electricity mix)	Compressed gas	Micro CHP (tri-generation)	Hydrogen burner and fuel cell
Di Florio et al. (2021)	2021	1	Italia	Electrolysis (solar PV combined with national grid)	Compressed gas	Single-family house nano-grid	reverse SOFC
Zhang et al. (2021)	2021	2	China	SR (natural gas)	Compressed gas	Power grid	Combined cycle; PAFC
Total		71					

biochemical, photocatalytic, electrical-thermochemical, photonic-biochemical or photo-electrochemical processes (Balat, 2008; Dincer and Acar, 2014). The hydrogen production methods investigated in the selected hydrogen-based power generation LCAs comprised electrochemical and thermochemical technologies (Figure 2). Other possible hydrogen production technologies—such as photochemical, biochemical, electrical-thermochemical, photonic-biochemical, and electrical-photonic conversion—were not represented among the studies. Forty-one of 71 hydrogen-based power generation LCA case studies evaluated electrochemical conversion, all of which employed the electrolysis technique. In contrast, thermochemical technology was found in the remaining cases (30 of 71 case studies). Steam reforming (SR) was the most common technique among the thermochemical category (17 of 30). In addition to SR, there are other thermochemical technologies such as autothermal reforming (ATR) (6 of 30), gasification (4 of 30), aqueous phase reforming (APR) (2 of 30), and cracking (1 of 30). Electrochemical LCA studies place a major emphasis on energy sources, whereas thermochemical LCA studies focus on feedstocks.

Hydrogen feedstock and energy sources

Among the electrochemical LCA studies, 73% of them used renewable energy sources (30 of 41 cases), including wind power (5/41), solar photovoltaic (11/41), hydropower (3/41), a combination of solar PV and wind power (2/41), and an unspecified renewable energy source (9/41), as shown in Figure 3. The grid mix, which was generated from non-renewable sources, was considered the energy source for hydrogen extraction in eight cases. In addition, the combination of renewable electricity with the grid mix was investigated in two cases. There was only one case that did not clearly explain its energy source. Among the thermochemical LCA studies, natural gas was the most common hydrogen feedstock (17/30) used in SR (11/17) and ATR (6/17). In addition, there are four cases in which LPG is used as the feedstock for the SR production process. Biomass was used for hydrogen production *via* gasification (4/30) and aqueous phase reforming (2/30). Only one case considered landfill gas as a hydrogen feedstock in the SR pathway, and two cases did not specify the feedstock that was used in SR and cracking processes.

Hydrogen storage

Hydrogen contains 143 megajoules (MJ) of energy per kilogram, up to three times greater than liquid hydrocarbon-based fuels (Mazloomi and Gomes, 2012). However, storing the same amount of hydrogen requires a greater volume due to its low volumetric energy density. Conventional techniques for storing hydrogen include compressing it as gas in tanks, as cryogenic liquid, and storing it underground. Recently, material-based or solid-state hydrogen storage—which includes metal hydrides, complex hydrides, chemical hydrides, and adsorbents—has been rapidly developing (Yue et al., 2021). Hydrogen can be stored in gaseous form at higher pressure levels than 700 bar to increase the volumetric energy density, for example at hydrogen refueling stations the hydrogen is compressed to 900 bar to enable fast refueling (Reuß et al., 2017). Liquid hydrogen, which is achieved by cooling its temperature to −253°C, is another option for storing hydrogen. Despite significant improvements to volumetric density, the liquefaction process requires at least 35% of the fuel's energy content (Durbin and Malardier-Jugroot, 2013). Aside from compressing hydrogen in gas and liquid, hydrogen underground storage alternatives such as aquifers, depleted natural gas and oil reserves, and salt caverns are the main options for large-scale medium and long-term hydrogen storage (Yue et al., 2021). Storing hydrogen in a solid state is achieved by combining hydrogen with materials through absorption and adsorption. In adsorption, hydrogen attaches to the surface of material either as hydrogen atoms or hydrogen molecules. In absorption, hydrogen is split into H-atoms, and the H-atoms are then incorporated into chemical compounds (Durbin and Malardier-Jugroot, 2013; Tashie-Lewis and Nnabuife, 2021; Yue et al., 2021).

Our review identified 64 of 71 case studies that considered the storage stage (see Figure 2). The most common form of hydrogen storage was compressed gas (51/64). Liquid hydrogen was considered in six of 64 cases; apart from the two storage alternatives, numerous studies considered material-based hydrogen storage, including methylcyclohexane (MCH) (Ozawa et al., 2019) and metal hydrides for storing hydrogen (Hwang and Chang, 2010).

Hydrogen-based power generation conversion technologies and applications

The selection of hydrogen-based power generation conversion technology is mainly determined by the application for which the electricity is required. Hydrogen can directly react with oxygen in a fuel cell to supply electrical energy to a system or it can be burned in a combustion engine, such as a piston engine or a gas turbine (Tashie-Lewis and Nnabuife, 2021).

A fuel cell is an electrochemical cell that converts the stored chemical energy of hydrogen and oxygen directly into electricity. A fuel cell has four major components: this includes the anode, the cathode, the electrolyte, and the external circuit. At the anode, hydrogen is oxidized into protons and electrons, while at the cathode, oxygen is reduced to oxide ions, which subsequently react to produce water. Depending on the electrolyte, either protons or oxide ions are transmitted through an ion-conducting but electronically insulating electrolyte, while electrons flow around an external circuit delivering electrical energy. Based on the characteristics of the electrolyte, fuel

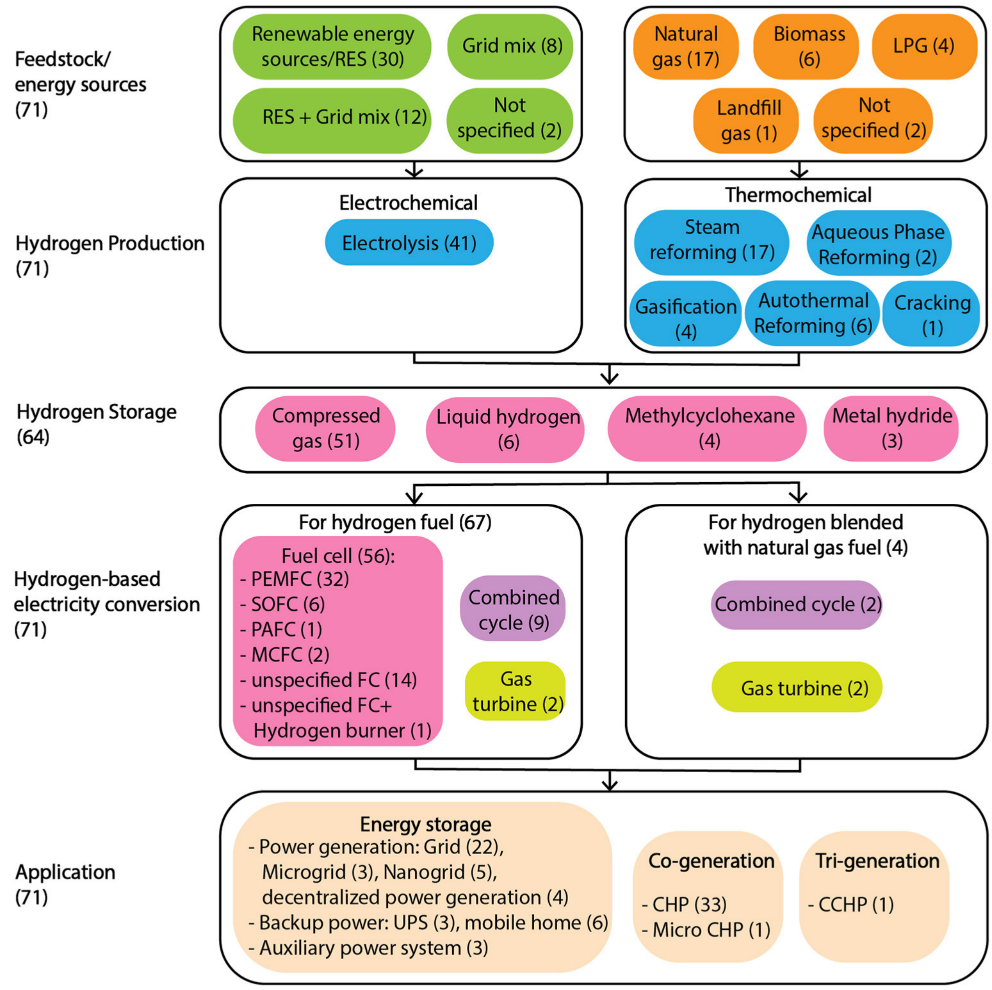


FIGURE 2
Overview of hydrogen-based power generation LCAs. The number in brackets indicates the number of cases in each particular area.

cells can be classified into several types. However, they all run according to the same basic principles (Ormerod, 2003; Mekhilef et al., 2012). The five types of fuel cells considered for application in hydrogen-based power generation systems are alkaline fuel cells (AFC), phosphoric acid fuel cells (PAFC), solid oxide fuel cells (SOFC), molten carbonate fuel cells (MCFC), and proton exchange membrane fuel cells (PEMFC). AFC uses an aqueous solution of either sodium or potassium hydroxide as the electrolyte, and the electrodes are made from carbon with a platinum electrocatalyst. AFC has an operating temperature of around 70°C. PAFC use carbon paper electrodes and phosphoric acid electrolyte and has an operating temperature of up to 200°C. SOFC uses a solid ceramic inorganic as the electrolyte, which operates at high temperatures, typically between 750°C and 1,000°C. MCFC uses a molten potassium lithium carbonate electrolyte and has an operating temperature of around 650°C. PEMFC uses a proton-conducting polymer electrolyte and operates at low temperatures between 60°C and 100°C. Fuel cells generate a range of power from 1 to 10 MW, making them suitable for practically any application that requires electricity (Ormerod, 2003; Mekhilef et al., 2012). Fuel cell can be used in household appliances, transportation, portable power and stationary power generation (such as combined heat and power (CHP), auxiliary power, and backup power) (Cottrell et al., 2011).

The technologies for hydrogen gas turbine power generation are designed for large-scale power generation. These technologies are classified into natural gas–hydrogen co-combustion and hydrogen-fired power generation. A gas turbine generates rotary motion by harnessing the energy contained in a gas—either the kinetic energy of the movement of a flowing gas stream or the potential energy of a gas under pressure. A modern gas turbine has three principal components: this includes a compressor, a combustion chamber, and a

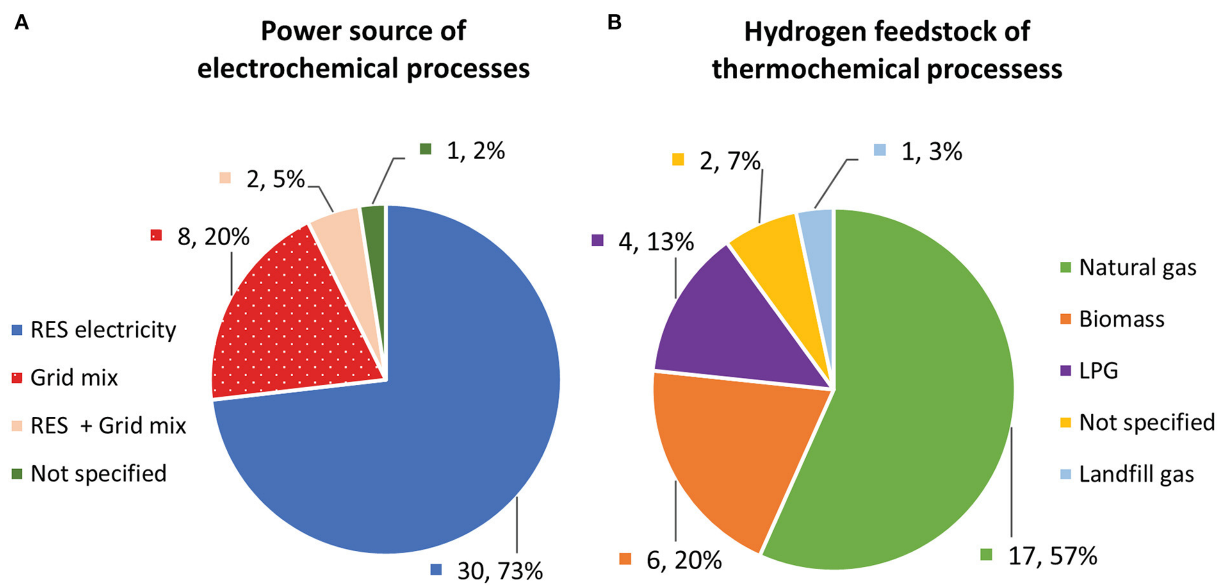


FIGURE 3
(A) Power sources of electrochemical cases and **(B)** hydrogen feedstock of thermochemical cases.

turbine stage. The turbine stage is the main energy-producing component, as it drives the compressor while also providing the energy to power the generator and produce electricity. The most important adaptation of the gas turbine cycle, however, is the combined cycle power plant. The efficiency of a gas turbine for power generation is always restricted by the fact that exhaust gases exit the turbine at a high temperature and therefore still contain a significant amount of energy that is not recovered. In a combined cycle power plant, the exhaust from the gas turbine is fed into a heat recovery steam generator, which converts the hot air into steam. Then, the steam is used to power a steam turbine generator, which generates an additional amount of electricity (Breeze, 2019).

The conversion technologies investigated in the hydrogen-based power generation LCAs examined in this study include gas turbines, gas and steam turbines (combined cycle power), and fuel cells, as shown in Figure 2. The fuel cell, which was assessed in 56 case studies, was the most common conversion technology. Among the application of fuel cells to converting hydrogen into electricity, thirty-two case studies assessed the application of PEMFC (Khan et al., 2005; Spitzley et al., 2007; Mori et al., 2014; Oliveira et al., 2015; Valente et al., 2015; Di Marcoberardino et al., 2018; Stropnik et al., 2018; Suwanmanee et al., 2018; Rossi et al., 2020). Of these, Six case studies assessed SOFC (Strazza et al., 2010; Bicer and Khalid, 2020; Di Florio et al., 2021), and two cases assessed MCFC (Lunghi et al., 2004). Only one case investigated PAFC (Zhang et al., 2021). Still, sixteen case studies do not specify the type of fuel cells in their evaluation (Sevencan and Çiftcioglu, 2013; Shimizu et al., 2020; Peppas et al., 2021). This review found fewer cases of hydrogen mono-combustion and co-combustion LCAs addressing either combined cycle or gas turbine conversion. Within the hydrogen mono-combustion category, combined cycle turbines were investigated in nine case studies, and only two case studies discussed gas turbine conversion technologies. Four cases focused on hydrogen-enriched natural gas co-combustion conversion (Walker et al., 2017). Two cases considered gas turbines, and two cases examined gas combined cycle turbines.

As hydrogen plays an important role in a variety of applications for storing and transferring energy, in this review three typical applications of using hydrogen in power generation systems are addressed with the LCA approach—energy storage, co-generation, and tri-generation (Figure 2). Hydrogen-based energy storage has recently attracted increased attention as it can satisfy a wide range of energy storage needs, from controlling short-term system frequencies to balancing the medium and long-term (seasonal) energy supply and demand (Parra et al., 2019). Compared to existing energy storage alternatives such as pumped hydro energy storage and batteries, hydrogen has the advantages of providing a high-capacity means of storing energy, the ability to store energy for a long time, as well as general flexibility (Bocklisch, 2016). To improve efficiency and lower costs, fuel cells and gas turbines can be used as the prime movers for CHP systems known as "co-generation systems," or it can be used for combined cold heat and power systems, known as "tri-generation systems." In a co-generation system, the prime mover produces both electricity and heat, with the energy being used for electrical needs and the released heat being used for heating.

Tri-generation is an application of co-generation that combines a primary mover with thermally driven equipment to generate cooling (Yue et al., 2021). In this review, some LCAs considered hydrogen as energy storage, including backup power (Sevencan and Çiftcioglu, 2013; Mori et al., 2014; Stropnik et al., 2018), power generation for the electrical grid (Lunghi et al., 2004; Khan et al., 2005; Oliveira et al., 2015; Valente et al., 2015; Walker et al., 2017; Ozawa et al., 2019; Zhang et al., 2021), decentralized power generation (Suwanmanee et al., 2018), power generation for a microgrid (Spitzley et al., 2007; Mori et al., 2021), power generation for nano-grid (Rossi et al., 2020; Di Florio et al., 2021), and auxiliary power (Strazza et al., 2010). Four studies provided LCAs of hydrogen for co-generation (Melamu and von Blottnitz, 2009; Di Marcoberardino et al., 2018; Bicer and Khalid, 2020; Shimizu et al., 2020), but only one study addressed the application of tri-generation systems (Peppas et al., 2021).

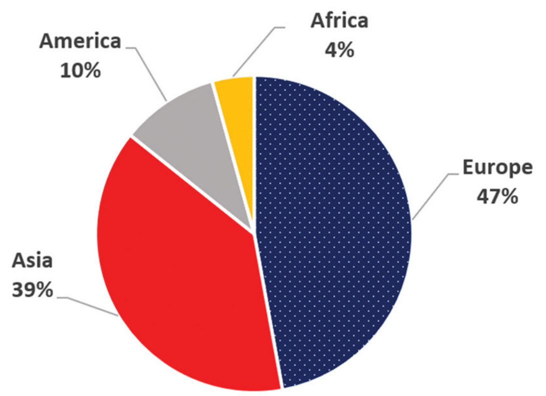


FIGURE 4
Geographical context of under-reviewed LCAs.

Geographical context

The geographical distribution of the LCAs is illustrated in Figure 4, along with their scope. The distribution is calculated based on the number of case studies rather than articles because one study considered cases from various regions (Stropnik et al., 2018) and provided LCAs in both the Norwegian and Moroccan contexts. More than half of the hydrogen-based power generation LCA case studies were examined in a European context (33 of 71). In this category, the countries that are most frequently represented include Italy (17 of 33), Germany (6 of 33), Belgium (4 of 33), Greece (1 of 33), Norway (1 of 33), Spain (1 of 33), and Slovenia (1 of 33). Yet, two case studies in the European context did not specify the country. Thirty-two of the 33 LCA cases on the European context focused on fuel cell applications, and only one case examined a gas turbine in mono-combustion conversion. Twenty-eight of the LCA case studies undertaken in the Asian context, including Japan (16/28), Turkey (6/28), Thailand (4/28), and China (2/28), considered fuel cell utilization and combined cycle turbines for mono-combustion conversion. By contrast, a limited number of case studies can be observed in the North American (7/71) and African (3/71) contexts.

Methodological choices

This qualitative analysis investigated the influencing methodological choices and was based on 21 hydrogen-based power generation LCAs.

System boundary

The system boundary specifies the specific system processes that are evaluated in a study. A one-dimensional definition of the system boundary is as either a "cradle-to-gate" or "cradle-to-grave" study; this classification is therefore insufficient for hydrogen-based power generation LCAs. Furthermore, the definition of the system boundary in hydrogen-based power systems is characterized by two paths. The first path is the hydrogen life cycle, which includes hydrogen production, storage, distribution, and use. In this case, a study is classified as cradle-to-grave if the scope includes hydrogen production and its use for power generation. The second path comprises the conversion system's manufacturing, operation, maintenance, and disposal. A study is labeled as a cradle-to-grave study if its scope covers all phases from manufacturing to disposal; otherwise, it is classified as a cradle-to-gate study.

Regarding the hydrogen life cycle of reviewed studies, hydrogen production and utilization are included in all LCAs (Figure 5). Hydrogen storage is generally included in LCAs, whereas only two of the studies excluded storage. Transportation was only included in eight studies. In terms of the conversion life-cycle equipment, all LCAs incorporated equipment operation. Therefore, equipment manufacturing is generally included in the scope of an LCA, whereas only three studies excluded this phase. Equipment maintenance and the end-of-life phase were only included in six and seven studies, respectively. Due to a lack of data on conversion system disposal and the fact that the studies are comparative LCAs, which would have a similar "end-of-life" phase for the various systems, excluding the end-of-life phase may be deemed an acceptable approximation of the goal of the studies (Di Florio et al., 2021).

Functional unit

The functional unit is an essential element of LCAs. It is a referencing unit for all environmentally relevant flows and impact assessment results. This review found that the functional units defined in LCAs of hydrogen for power generation

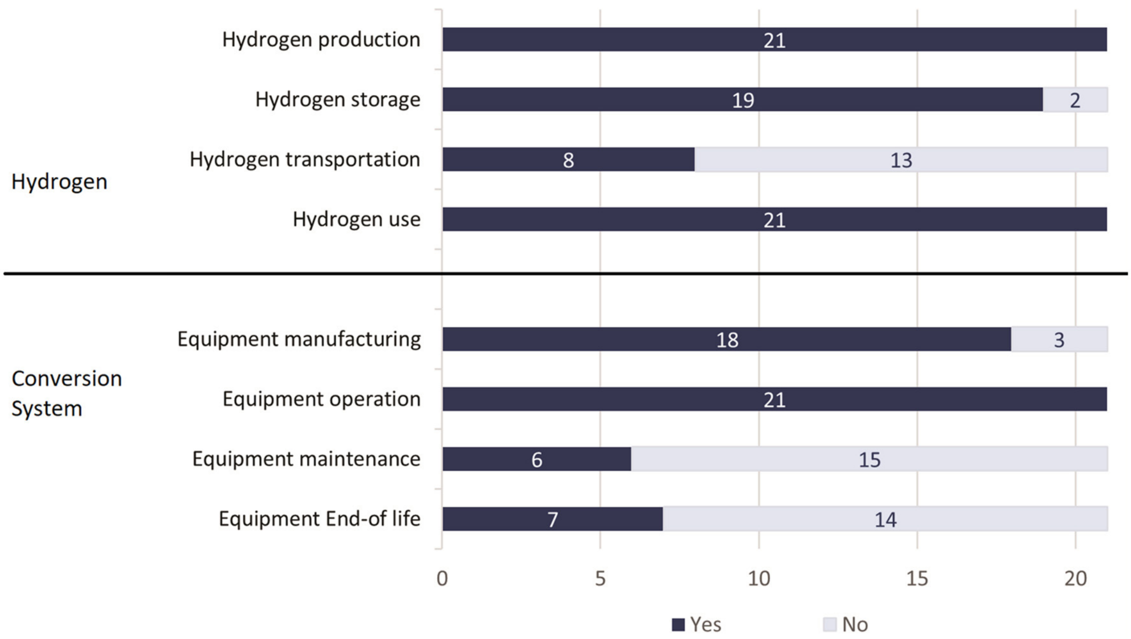


FIGURE 5
System boundaries of the reviewed studies.

are not uniform (Table 2). Figure 6 shows the number of times a functional unit was chosen in the reviewed studies. The amount of produced electricity in kilowatt hours (kWh) or megawatt hours (MWh) is the most extensively selected functional unit. Certain LCAs used case-specific of energy output-related functional units, such as the annual energy provided for one, two, or 10 dwellings (Di Marcoberardino et al., 2018), the use of hydrogen technologies in a specific region or area for 1 year (Shimizu et al., 2020), total power generation for 1 year (Sevencan and Çiftcioglu, 2013), and full coverage of energy demands for 20 years (Peppas et al., 2021). Only one study used exergy (in J) as a functional unit. Lastly, only one study used energy (in MJ) as a functional unit. Even though all functional units are energy output-related, a high discrepancy of functional units is seen. The purpose of hydrogen-based electricity technology application is responsible for the differences of functional units selected by reviewed studies. The main product of hydrogen-based power is electricity or electricity and useful heat, depending on their purposes. FC-HyGuide recommends using exergy as functional unit when both electricity and useful heat are generated and utilized (Lozanovski et al., 2011; Masoni and Zamagni, 2011). The exergy is defined as the sum of electricity (in MJ) plus the useful thermal energy (in MJ) times a Carnot coefficient. Adopting exergy could help avoid high discrepancies in selected functional units among the reviewed studies.

Allocation

Allocation rules were considered in LCA studies with multifunctional systems in which multiple outputs were produced. Despite the significance of allocation, many hydrogen-based power LCAs did not provide transparency regarding the allocation method applied. The information provided by the reviewed studies is scarce, since only two studies explicitly mentioned their allocation method. Energy efficiency allocation was selected to calculate the specific energy output in a combined heat and power plant (Bicer and Khalid, 2020). Physical allocation was used to examine the environmental impact of each reference flow of the system, which considers the hydrogen produced at the end of the equipment's life cycle as a by-product (Rossi et al., 2020). Additionally, an expansion system considered a multifunctional system was chosen in which conventional methods of producing electricity and transport fuel were applied, so that all scenarios were compared based on the same set of functions (Melamu and von Blottnitz, 2009).

Life-cycle inventory

The source of data for the life-cycle inventory step and their quality strongly affected the reliability of the assessment. Inputs and outputs for both foreground and background elements were used as representatives for the entire system. A definition of the foreground and background systems must be performed

TABLE 2 Methodological choices of the reviewed LCAs.

Study	Life-cycle boundary		LCIA method	Functional unit	Impact category	
	Hydrogen	Conversion equipment			Midpoint	Endpoint
Lunghi et al. (2004)	Cradle to grave	Cradle to grave	Eco-indicator 99	1 kWh of electricity	GWP, AP	EQ, HH, R
Khan et al. (2005)	Cradle to grave	Cradle to grave	Theoretical calculation	1 kWh of electricity	GWP	-
Spitzley et al. (2007)	Cradle to grave	Cradle to gate	Theoretical calculation	1 kWh of electricity	GWP	-
Melamu and von Blottnitz (2009)	Cradle to grave	Cradle to gate	Not specified	1 MJ of energy	GWP, AP, EP, HTP	
Strazza et al. (2010)	Cradle to grave	Cradle to gate	Not specified	1 kWh of electricity	GWP, AP, EP, POCP, ODP	-
Sevencan and Çiftcioglu (2013)	Cradle to grave	Cradle to grave	Eco-indicator 99	Total power generation for one year	-	EQ, HH, R
Mori et al. (2014)	Cradle to grave	Cradle to gate	CML 2001	1 kWh of electricity	GWP, AP, EP, ADP	
Oliveira et al. (2015)	Cradle to grave	Cradle to grave	ReCiPe 2018	1 kWh of electricity	GWP, HTP, PMF, FD	SS
Valente et al. (2015)	Cradle to grave	Cradle to gate	CML, CED (VDI), IPCC	1 MWh of electricity	GWP, AP, EP, ADP, POCP, ODP, HTP, CED, LU	-
Walker et al. (2017)	Cradle to grave	Cradle to gate	GREET	1 MJ of energy	GWP	-
Di Marcoberardino et al. (2018)	Cradle to grave	Cradle to grave	IMPACT 2002+ v2.2	Heat and electricity for 1, 2, and 10 dwellings over one year	GWP, WD	EQ, HH, R
Stropnik et al. (2018)	Cradle to grave	Cradle to grave	CML 2001	1 kWh of electricity	GWP, AP, EP, ADP, POCP, ODP, HTP, FETP, METP, TETP	-
Suwanmanee et al. (2018)	Cradle to grave	Cradle to gate	CML 2000; Eco-indicator 99	200 kWh of electricity	GWP	EQ, HH, R
Ozawa et al. (2019)	Cradle to grave	Cradle to gate	Theoretical calculation	1 kWh of electricity	GWP	-
Bicer and Khalid (2020)	Cradle to grave	Cradle to gate	ReCiPe 2018	1 kWh of electricity	GWP, POCP, HTP, PMF, WD, FD	-
Mori et al. (2021)	Cradle to grave	Cradle to gate	EF	Amount of energy provided in the form of heat and electricity during 1 year of operation	GWP, HTP, FETP, FEP, MEP, TEP, IR, LU, ODP, POCP, RUec, RUmm, RI, WS	
Rossi et al. (2020)	Cradle to grave	Cradle to grave	ReCiPe 1.11	1 MWh of electricity	-	SS
Shimizu et al. (2020)	Cradle to grave	Cradle to gate	LIME	One-year use of alternative technologies with10% replacement for the conventional technologies	GWP, AP, POCP, UAAP	HH, SA, PP, SS (LIME index)
Di Florio et al. (2021)	Cradle to grave	Cradle to gate	ReCiPe 2016; CED; IPCC	1 MJ of exergy	GWP, ODP, HTP, MA ETP, TETP, TEP, TAP, MEP, CED, PMF, LU, FR, MR, IR, WC	-
Peppas et al. (2021)	Cradle to grave	Cradle to gate	CML, ReCiPe	Full coverage of energy demands for 20 years	GWP, AP, POCP	
Zhang et al. (2021)	Cradle to grave	Cradle to gate	IPCC	1 MJ of energy	GWP	-

GWP, global warming potential; HTTP, human toxicity; POCP, photochemical ozone creation; AP, acidification potential; EP, eutrophication potential; PMF, particulate matter formation; ODP, ozone depletion potential; WD, water depletion; ADP, abiotic depletion potential; FD, fossil depletion; METP, marine ecotoxicity; FETP, freshwater ecotoxicity; TETP, terrestrial ecotoxicity; TAP, terrestrial acidification; FEP, freshwater eutrophication; MEP, marine eutrophication (MEP); IR, ionizing radiation; FR, fossil resources scarcity; MR, mineral resource scarcity; WS, water consumption; UAAP, urban area air pollution; LU, Land use; and EU, energy use. RUec, Resource us; energy carrier; RUmm, Resource use; mineral and metals; RI, Respiratory inorganics; WC, water consumption; WS, Water scarcity; HH, human health; EQ, Ecosystem quality; R, resource; SA, Social asset damage; PP, Primary production damage; SS, single score.

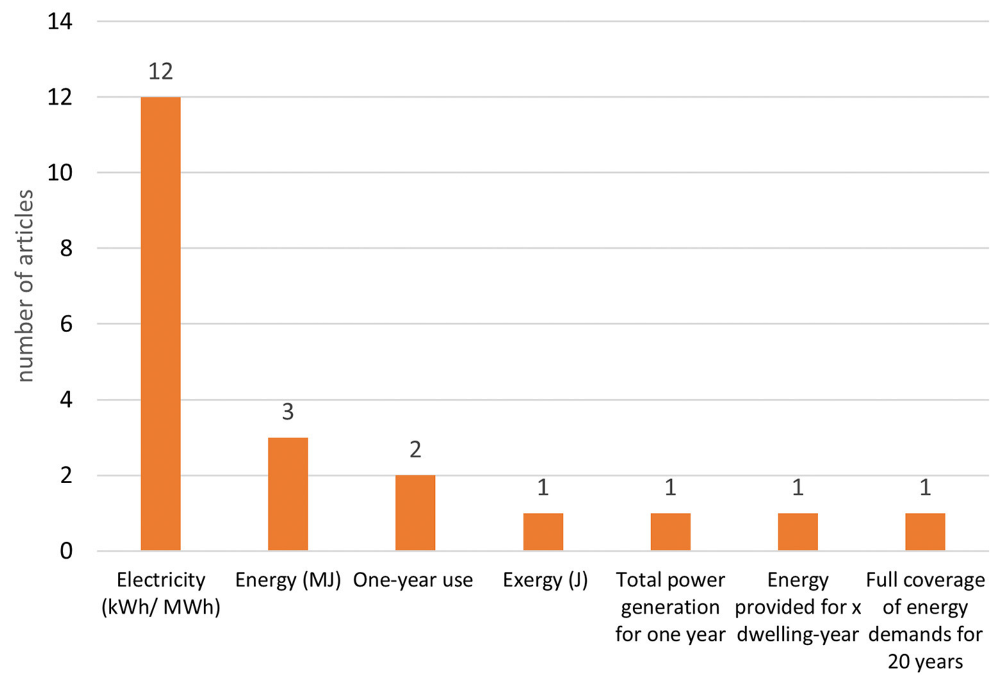


FIGURE 6
Prevalence of functional units in the reviewed studies.

prior to the selection of data sources in LCI stage. The foreground system, according to FC-HyGuide, consist of the main process phases and the related infrastructure processes such as manufacturing. The foreground system is supported by the background system which is made up of processes such as the infrastructure for the supply of energy including power plants and power lines.

In general, there are two types of data used in a LCA study: primary and secondary data. Primary data is recommended to be used for the main processes (foreground system). Primary data is provided from on-site measurement, project partners, manufacturer and/or operator of the system. However, the owner, project partner, manufacturer or operator of the system may not have all of the data required to perform LCA. In that case, secondary data is needed to fill data gaps. Secondary data is also used for the background system. Different data sources can be used for secondary data, such as LCA databases, scientific literature, non-scientific literature, simulations, calculations, assumptions, etc.

The data source distribution for both foreground and background processes is shown in Figure 7. Due to the homogeneity of the data source used in the study, this figure was set according to the source of a single article. “On-site measurement” refers to gathering data from actual measurements either in the field or lab scale. Data from the manufacturer was referred to as “manufacturer,” while “scientific literature” implies published articles. Data from the project partner was classified as “project documentation.” Expert estimation and assumptions were classified in the “other” category. Scientific literature and LCA databases, such as ecoinvent, GaBi, IDEA, BUWAL, and The Argonne’s The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) databases were widely used for the foreground process. For the background process, data were generally obtained from the LCA database. The database choice was based on various criteria, such as the geographical context and the LCA software.

Life-cycle impact assessment method

There are two distinct impact categories—the midpoint (problem-oriented) approach and the endpoint (damage-oriented) approach. The midpoint approach evaluates the environmental impact in the middle of the environmental cause-and-effect chain. In contrast, the endpoint approach concentrates on the damage that occurs at the end of the chain (i.e., human health, ecosystem quality, and resources) (Guinée, 2002). The midpoint approach provides a more detailed and scientific decision-making foundation, while the endpoint approach is easier to interpret and communicate to decision-makers (Bare et al., 2000; Dong and Ng, 2014). Several of the studies reviewed either applied a midpoint approach (14 of 21) or an endpoint approach (3 of 21), whereas four of the

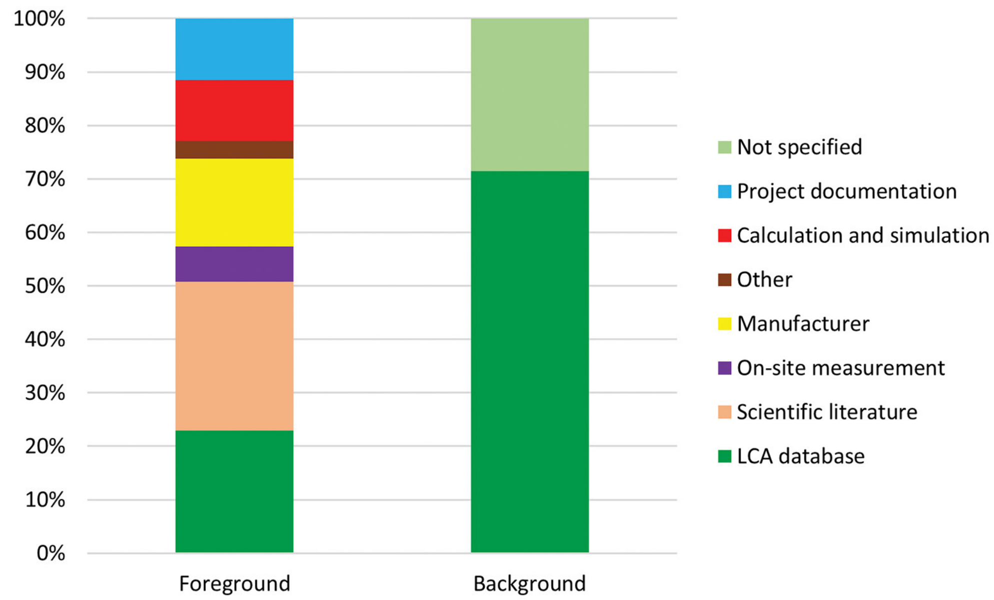


FIGURE 7
Data source distribution for foreground and background processes of the energy systems used in LCAs.

studies utilized both the midpoint and endpoint approaches (Table 2).

The commonly selected impact categories for midpoint and endpoint approaches and the life-cycle impact assessment (LCIA) methods used are illustrated in Figures 8, 9, respectively. These methodological families were recognized without regard for version discrepancies. When a study employed its formula or model in the quantification process, the LCIA method was classified as "other." Due to the homogeneity of the impact indicators and LCIA methods employed in a study, both figures were set according to the choices of a single article. The secondary axis of the figure indicates the total number of studies considering each impact category. Numerous LCIA methods were applied in the reviewed studies, including Centrum voor Milieukunde Leiden (CML), ReCiPe, IMPACT2002+, Environmental Footprint (EF), LIME2, Intergovernmental Panel on Climate Change (IPCC), GREET, cumulative energy demand (CED), and Eco-indicator 99.

Figure 8 presents commonly selected impact categories for the midpoint approach, including global warming potential (GWP), human toxicity potential (HTP), photochemical ozone creation potential (POCP), acidification potential (AP), eutrophication potential (EP), particulate matter formation (PMF), ozone depletion potential (ODP), water depletion (WD), abiotic depletion potential (ADP), and fossil depletion (FD). Based on the impact category choices in the LCIA phase, GWP was the most often examined impact indicator in hydrogen LCAs, followed by HTP and POCP. Regarding the LCIA methods for the midpoint approach, CML was the most widely used method for evaluating GHG emissions. IPCC was used to quantify GWP in three of the studies; CML also played a significant role in quantifying the other midpoint categories, including HTP, POCP, AP, EP, PM, ODP, WD, and ADP. ReCiPe was employed to characterize GWP, HTP, POCP, PM, ODP, WD, and FD. GREET was only used to quantify GHG emissions, whereas LIME2 was utilized to evaluate GWP, POCP, and AP. The EF method was applied to assess GWP, HTP, POCP, ODP, WD, and FD. Other than these impact categories, other midpoints were investigated in several of the reviewed studies, including marine ecotoxicity (METP), freshwater ecotoxicity (FETP), terrestrial ecotoxicity (TETP), terrestrial acidification (TAP), freshwater eutrophication (FEP), marine eutrophication (MEP) (Stropnik et al., 2018; Di Florio et al., 2021; Mori et al., 2021), ionizing radiation (Di Florio et al., 2021; Mori et al., 2021), fossil resources scarcity, mineral resource scarcity, water consumption (Di Florio et al., 2021, urban air pollution (Shimizu et al., 2020), land use (Valente et al., 2015; Di Florio et al., 2021), and energy use (Strazza et al., 2010; Valente et al., 2015; Di Florio et al., 2021). Two authors applied the CED method to evaluate energy use over the life cycle of the system under study (Valente et al., 2015; Di Florio et al., 2021). A few studies addressed endpoint categories, including human health, ecosystem quality, resource damage, and single scores (see Figure 9). Eco-indicator 99 was employed in all the indicators of the endpoint approach. Conversely, ReCiPe was only used to evaluate single-score indicators. IMPACT2002+ was applied to assess human health, ecosystem quality, and resources. In addition, LIME2 was utilized to evaluate human

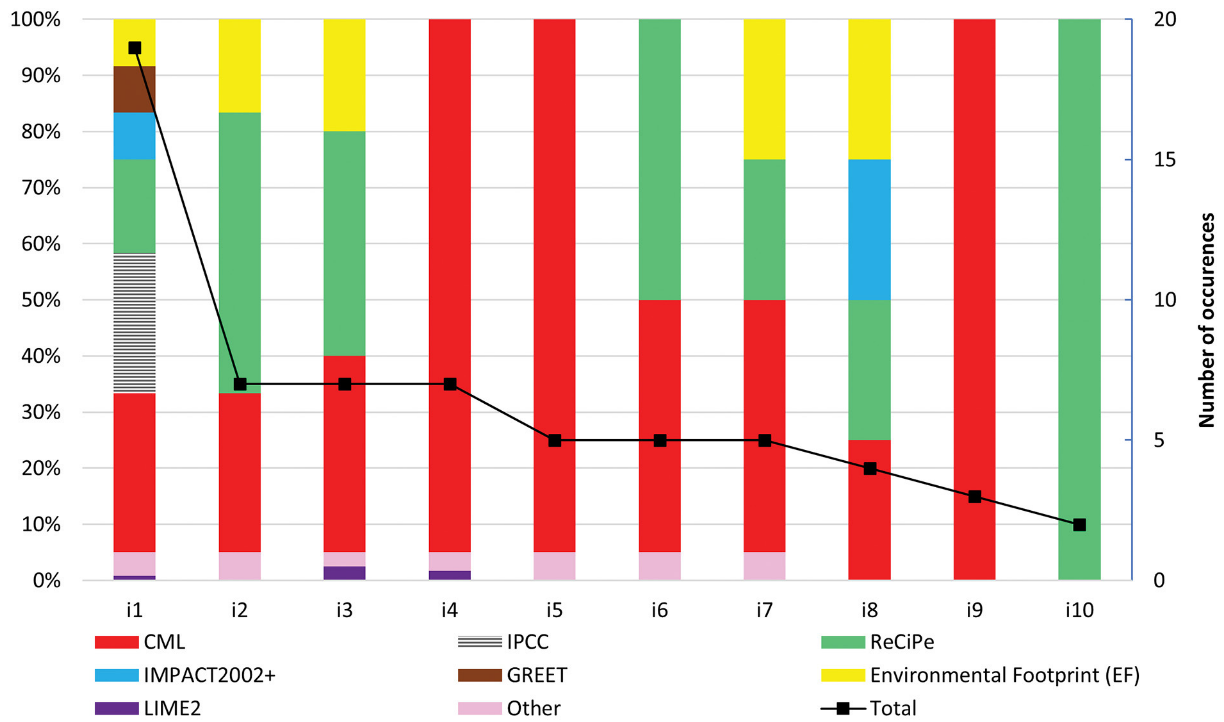


FIGURE 8
Choice of midpoint impact approach and their respective LCIA methods. i1: global warming potential (GWP), i2: human toxicity potential (HTP), i3: photochemical ozone creation potential (POCP), i4: acidification potential (AP), i5: eutrophication potential (EP), i6: particulate matter formation (PMF), i7: ozone depletion potential (ODP), i8: water depletion (WD), i9: abiotic depletion potential (ADP), and i10: fossil depletion (FD).

health and single-score indicators. Generally, the selection of the LCIA method for both midpoint and endpoint approaches in the reviewed studies depended on their environmental goals and ecosystem characteristics.

The distribution of the LCA software or tools used during the LCIA phase is shown in Figure 10. Because of homogeneity of the LCA software or tools employed in a study, a pie chart was created according to the choices of a single article. SimaPro and GaBi were widely used in the reviewed LCAs. Open LCA, GREET, and in-house software was utilized at similar frequencies. Many studies explicitly mentioned the software used.

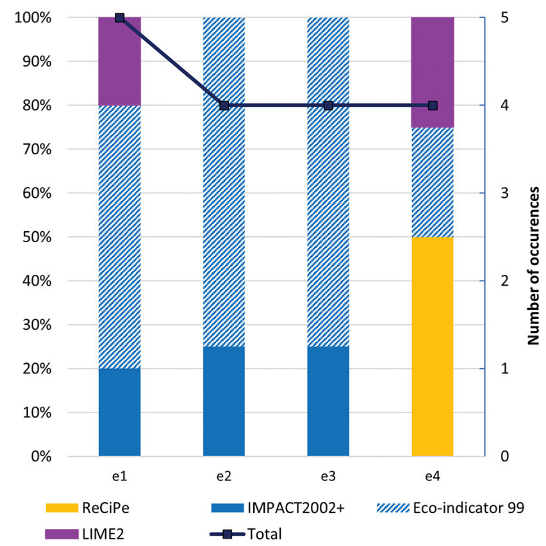


FIGURE 9
Choice of endpoint impact categories and their respective LCIA methods. e1: human health (HH), e2: ecosystem quality (EQ), e3: resources (R), e4: single score (SS).

Environmental impact of hydrogen-based power generation lcas

This section presents a quantitative assessment of the LCA results. Studies that did not provide comparable data were excluded from the quantitative analysis for the following reasons: (i) certain LCAs reported impact categories exclusively in graphical data, (ii) certain LCAs reported hydrogen-enriched natural gas power generation, and (iii) several LCAs used case-specific functional units, such as the annual demand for a specific system. The quantitative analysis concentrated on the commonly used midpoint impact categories of hydrogen-based power, including GWP, POCP, AP, and EP. The LCA results are

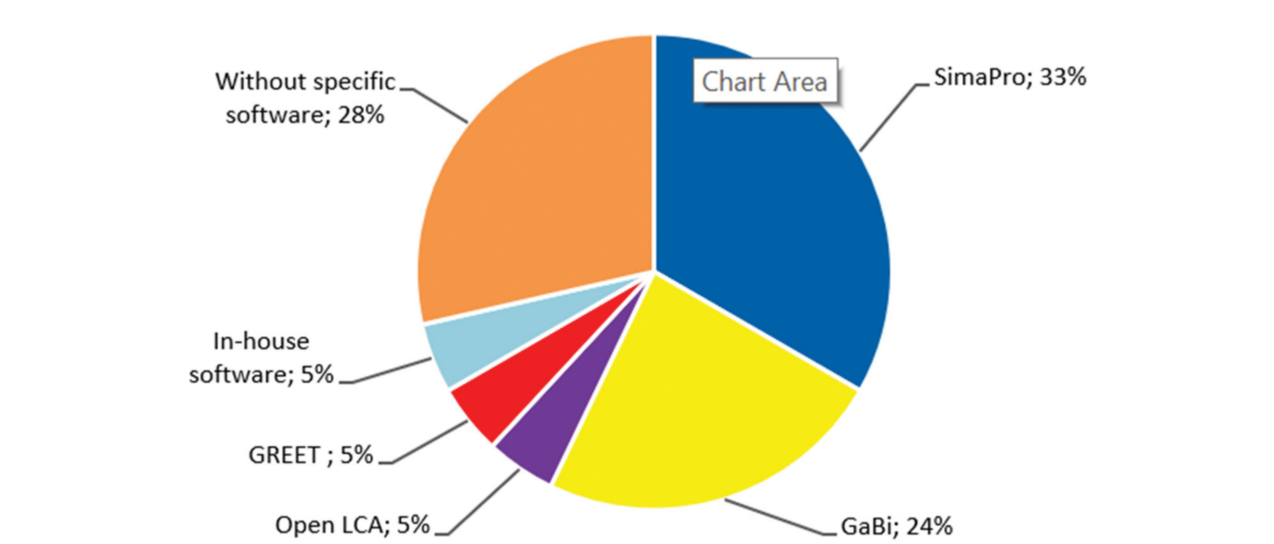


FIGURE 10
Types of LCA software.

presented according to homogeneous functional units and have been recalculated in the respective equivalency units.

Global warming potential

The analysis of GWP (Figure 11) was based on 19 case studies from eight studies. The GWP of a hydrogen-based power generation system has a median of 289 g/CO_2 eq. per kWh of produced electricity and ranges from 17.29 grams of carbon dioxide equivalent (g CO_2 eq.) per kWh to 4,040 g CO_2 eq. per kWh.

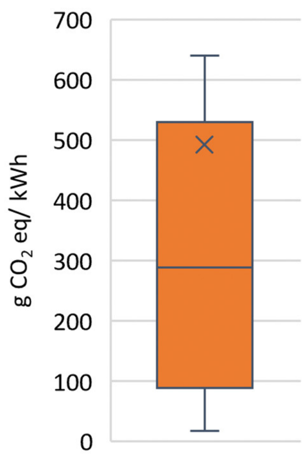


FIGURE 11
Global warming potential (in gCO_2 eq per kWh).

Other impact categories

The other impact results discussed in the LCAs include POCP, AP, EP, and ODP. However, the few case studies that did so prevented a comprehensive quantitative analysis. The results are presented in Table 3.

Discussion

Technical aspects of conversion technology

In this section, the description about the construction of fuel cells is discussed. We do not discuss the construction of gas turbine and combined cycle turbine due to absence of detailed description in the reviewed studies. The amount of electricity generated by a fuel cell is determined by various parameters, including the type of fuel cell, its size, operating temperature, and gas supply pressure. To boost the voltage, individual fuel cells are connected in series to form a stack. A fuel cell stack might comprise a few or hundreds of individual cells, depending on the purpose. The stack construction consists of electrodes, matrixes, and bipolar plates production and their relative assembly (Lunghi and Bove, 2003).

In the case of fuel cells stack/system, FC-HyGuide requests a brief description of the FC system or stack. Information about the major properties needs to be given by stating the FC standard being met, such as: IEC/TS 62282-1 and IEC 62282-2. If no standard applies, FC-HyGuide asks that the following properties be reported: trade name, type of electrolyte used, primary functions, electrical power (rated output), thermal power,

TABLE 3 The results of other impact categories.

Impact category	Functional Unit	Number of LCAs	Number of case studies	Max	Mean	Median	Min	Range
POCP (g C2H4 eq)	1 kWh electricity	5	9	4.60E-01	1.81E-01	1.03E-01	3.63E-03	4.56E-01
AP (g SO2 eq)	1 kWh electricity	4	8	1.51E+00	7.20E-01	4.05E-01	7.28E-02	1.44E+00
EP (g PO4 eq)	1 kWh electricity	4	7	1.30E+00	2.94E-01	1.43E-01	1.83E-02	1.28E+00
ODP (g CFC-11 eq)	1 kWh electricity	2	4	1.58E-06	1.30E-06	1.21E-06	1.21E-06	3.70E-07

efficiency, rated voltage, rated current, range of temperatures and operating temperature, weight, dimensions, duel used and its technical specifications, expected service life, and description of the intended use. Within the reviewed studies, highly detailed description about the construction and operation of selected fuel cells or gas turbine are scarce.

Within the reviewed publications, fuel cells can be used for a wide variety of application, PEMFC for power grid purposes (250 kW, 500 kW), SOFC for auxiliary power system (20 kW), PEMFC for backup power (3 kW), SOFC for CHP (250 kW) and PAFC for power grid (50 MW).

Level of accordance with fc-hyGuide

The result of the evaluation on methodological choices of reviewed studies regarding the level of agreement with a selected set of recommendations from the FC-HyGuide are summarized in Table 4. In this table, the level of accordance is classified as "very high" if more than 90% of studies followed the corresponding FC-HyGuide recommendation, "high" (60–90%), "intermediate" (40–60%), "low" (10–40%) and "very low" (<10%). According to the time of publication of the reviewed studies, the level of agreement is provided in the columns "before FC-HyGuide" and "after FC-HyGuide". Regarding product system information, it can be seen that the level agreement before and after the publish of the FC-HyGuide is quite similar. Furthermore, a highly satisfying level of agreement is generally obtained both before and after FC-HyGuide in terms of goal and scope definition, despite a contrary tendency in accordance with the use of an attributional modeling approach (from "high" to "intermediate"). On the other hand, dealing with using ISO hierarchy for solving multifunctional process, unsatisfactory level of agreement has found before the release of FC-HyGuide. In regard with LCI, the lack of information about data quality requirement and definition of foreground and background processes are found both before and after the release of FC-HyGuide. While the use of primary data was assessed as intermediate before and after FC-HyGuide. In contrast, a satisfactory level of agreement regarding filling data gaps with secondary data is found both before and after FC-HyGuide. Concerning LCIA, a downward trend is generally observed when comparing the studies before and after FC-HyGuide, despite an increase level of agreement for selecting the LCIA method and reporting non-normalized and non-weighted LCA results. In terms of impact categories, FC-HyGuide recommends the use of midpoint impact categories instead of endpoint categories, the use of GWP, AP, EP, ADP, primary energy demands (renewable and non-renewable), and other categories such as ODP, HTP, LU etc. (Lozanovski et al., 2011; Masoni and Zamagni, 2011). Regarding the LCIA method, FC-HyGuide suggests selecting recommended methods by JRC (European Commission - Joint Research Centre., 2011) or the midpoint CML method. JRC recommends LCIA methods for relevant environmental impact categories, for instance, the IPCC method is endorsed for evaluating GWP. Using the CML method or JRC recommendation reaches an intermediate level of agreement in the reviewed paper published after FC-HyGuide was announced. Despite the intermediate level of agreement in LCIA method with FC-HyGuide, the use of AP, EP and ADP impact categories are "low" since some studies only focused on GWP impact category.

LCA result

This section discusses the LCA results for the various studies included in this systematic review. We only included GWP to ensure that there were enough cases for an in-depth discussion. This measure does not imply that the other impact categories are less important. A wide range of reported results (Figure 11) can be attributed to an outlier above the upper fence. This outlier was derived from a case study reporting a system in which hydrogen was produced through electrolysis from a grid mix and converted into electricity in a PEMFC for UPS application (Stropnik et al., 2018). Moreover, the system boundary of the outlier case study was both cradle-to-grave for the hydrogen life cycle and conversion equipment life cycle, in contrast to other studies that conducted cradle-to-gate LCA analysis excluding the end-of-life phase. Furthermore, different LCA studies of similar hydrogen-based electricity pathways frequently provide divergent results, emphasizing the need to harmonize the LCA methodology in hydrogen-based electricity studies.

These discrepancies between the environmental impact results were likely caused by variations in technical and

TABLE 4 Accordance of observed trends with FC-HyGuide.

Topic	Corresponding recommendations from FC-HyGuide	Level of accordance[a]	
		Before FC-HyGuide[b]	After FC-HyGuide[c]
Product system information	State system configuration or description	Very high	Very high
	State the significant properties or technical characteristics of technology	High	High
Goal and scope definition	Clear define the goal of the study	High	High
	Show the chosen system boundary in a flow chart	High	High
	Use "the amount of energy or exergy defined" as the functional unit	Very high	Very high
	Use an attributional modeling approach in LCA studies	High	Intermediate
	Use the ISO hierarchy for solving multifunctional processes	Low	High
	The system boundary is consistent with the goal of the study	High	Very high
	In comparative studies, use the same rules for system boundary definition	Very high	Very high
	In comparative studies, methodological and data assumptions shall be analogous	Very high	Very high
	In comparative studies, harmonizing functional unit	Very high	Very high
	In comparative studies, harmonizing LCIA	Very high	Very high
Life cycle inventory analysis	Define the data quality requirement according to the goal and scope	Very low	Very low
	Define foreground and background processes taken into account	Low	Low
	Use primary data for the foreground system	Intermediate	Intermediate
	Fill data gaps with secondary data	Very high	Very high
Life cycle impact assessment	Use midpoint categories	Very high	High
	Show non-normalized and non-weighted results	Intermediate	High
	Use the Global Warming Potential impact category	Very high	High
	Use the Acidification Potential impact category	High	Low
	Use the Eutrophication Potential impact category	Intermediate	Low
	Use the Abiotic Depletion Potential impact category	Very low	Low
	Use renewable/non-renewable Primary Energy Demands categories	Very low	Low
	Use the midpoint CML methods or recommended methods by JRC	Low	Intermediate

[a] Level of agreement: very low, $\leq$ 10% of the case studies; low, 10–40 %; intermediate, 40–60 %; high, 60–90 %; very high, $\geq$ 90%.
[b] Based on five publications between 2004 and 2011.
[c] Based on sixteen publications between 2011 and 2021.

methodological choices in preparing the LCA such as functional units, allocation, system boundaries, and LCIA methods. These include options regarding hydrogen pathways and conversion equipment for technical decisions. Furthermore, a robust comparison of the environmental results from various studies with different methodologies cannot be conducted without a harmonization procedure.

Finally, it is essential to note that LCA merely provides an environmental profile of hydrogen-based power generation technologies, but other aspects are equally important, such as socioeconomic factors. Integrating LCA research with economic analysis, such as life cycle costing or techno-economic assessments, is highly recommended. Numerous reviewed studies have addressed economic analysis (Di Marcoberardino et al., 2018; Mori et al., 2021; Zhang et al., 2021). Integrating public acceptance and social effect assessments into the research and development process could be critical for the future deployment of hydrogen-based power generation. Several LCIA approaches for quantifying products' social and sociological impact have been developed in recent years. This LCIA method is known as the social LCA (S-LCA) method. Its typical impact categories represent the five main stakeholder groups of the product supply chain—workers, consumers, the local community, society, and value chain actors (United Nations Environment, 2020). However, no studies that addressed social LCA were encountered during this systematic review.

Conclusion and outlook

Hydrogen-based electricity generation LCAs were synthesized in this systematic review. Although the number of LCA studies has grown over the last 4 years, an inherent limitation of this systematic review is the small number of available studies and scenarios. The findings of the reviewed LCAs could help us evaluate the life-cycle environmental impact of hydrogen-based electricity systems, identify hotspots, direct future research, set performance goals, and provide a baseline for large-scale applications.

To address the first objective of this review, a qualitative analysis of the technological aspects of hydrogen-based electricity was performed. Electrolysis from renewable energy resources was the most widely considered hydrogen production method in the LCAs. In addition, storing hydrogen in compressed gaseous form was the most often used storage option. Fuel cell technology was the most common conversion equipment used in hydrogen-based electricity LCAs. Only a few studies have reported gas turbine and combined cycle with pure hydrogen fuel or hydrogen mixed with natural gas fuel. Many scenarios also focused on using hydrogen for energy storage and co-generation purposes. We identify some significant knowledge gaps and technology difficulties. Performing a LCA studies requires sufficient data and information about material and energy inputs and outputs, as well as the cause-effect relationships throughout the entire supply-chain of a technology which is generally obtainable in mature technology. The lack of data availability as well as uncertainty in the data and findings of LCA studies are common, especially for low Technology Readiness Level (TRL) technologies. Such is the case for many hydrogen-based power systems. On the other hand, early technology assessment offers a great chance to improve design and environmental profile. Hence, considering both level of technology maturity and the level of maturity of the market into which the technology will be implemented, are significant elements in emerging technology evaluations.

To address the secondary objective of this systematic review, a qualitative analysis of the methodological choices made in the hydrogen-based electricity LCAs was conducted. Based on the results, the methodological choices differed between studies, including functional units, allocation, system boundaries, and LCIA methods. Based on the observed methodological trends of reviewed studies and its accordance with FC-HyGuide, the quality of data and limitation of data availability should be reported. The selection of impact categories and impact assessment method should be done in accordance with FC-HyGuide. When conducting LCA with multi-functional concerns, ISO recommends using the multi-functionality hierarchy with system expansion or allocation (ISO, 2006). In the case of allocation of hydrogen-based power should be avoided, there is a possible way either by adopting exergy as functional unit or applying system expansion.

Due to uncertainty of endpoint indicators and single scores, FC-HyGuide suggest using midpoint indicators in the assessment. Therefore, the following ISO recommendations would be the appropriate approach: providing results with midpoint indicators (mandatory) and using endpoint indicators and particularly, single scores (optionally) only when comparing declaration are not to be disclosed to the public. LCA should not be restricted to climate change impacts only. Instead, a wide range of environmental impacts should be examined to prevent shifting the burden to other impact categories, such as reducing climate change impact yet raising human toxicity. These tradeoffs should be thoroughly examined.

To address the third aim, a quantitative analysis of environmental impact results was conducted. Based on GWP, the most extensively discussed impact indicator, hydrogen-based power generation systems, had a median of 289 g CO_2 eq. per kWh of electricity produced and ranged from 17.29 g CO_2 eq. per kWh to 4,040 g CO_2 eq. per kWh. These extreme values were likely caused by various technical and methodological choices in the preparation of the LCA. For this reason, a reliable comparison of the LCA results of multiple studies with varying methodological and technical options is not possible without first creating a harmonization procedure.

A lack of full traceability of results was noted from the reviewed studies. Only a few studies which identify the critical issues by quantifying which process/ flows are major contributors to the total impact. For future LCA studies, along with FC-HyGuide we recommend showing the contribution of significant processes to the total impact in stacked columns or pie-charts for identifying hotpots.

Lastly, it is essential to note that LCA only presents the environmental profile of a prospective product. However, other aspects, such as socioeconomic issues, are just as important. Economic considerations have been examined in a few of the reviewed studies. However, no studies that addressed social impact were encountered during this systematic review. Applying the social LCA (S-LCA) methodology to the research and development process may be beneficial in understanding the future impact of the hydrogen-based power generation more holistically.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

DIR carried out the review with support from AK and ST. DIR and AK wrote the manuscript. SM helped supervise the project. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the New Energy and Industrial Technology Development Organization under a Grant (number P14026).

Acknowledgments

DIR would like to thank Diponegoro University, The Ministry of Education and Culture of the Republic of Indonesia for the full support through the Doctoral Degree Scholarship Program (Grant No. 8205/UN7.P2/KS/2019).

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsus.2022.920876/full#supplementary-material

References

Balat, M. (2008). Potential importance of hydrogen as a future solution to environmental and transportation problems. *Int. J. Hydrog.* 33, 4013–4029. doi: 10.1016/j.ijhydene.2008.05.047

Bare, J. C., Hofstetter, P., Pennington, D. W., and Udo De Haes, H. A. (2000). State-of-the-art, LCIA midpoints versus endpoints, The sacrifices and benefits. *Int. J. Life Cycle Assess.* 5, 319–326. doi: 10.1007/BF02978665

Becheikh, N., Landry, R., and Amara, N. (2006). Lessons from innovation empirical studies in the manufacturing sector, a systematic review of the literature from 1993-2003. *Technovation.* 26, 644–664. doi: 10.1016/j.technovation.2005.06.016

Bhandari, R., Trudewind, C. A. and Zapp, P. (2014). Life cycle assessment of hydrogen production via electrolysis - a review. *J. Clean. Prod.* 85:151–163. doi: 10.1016/j.jclepro.2013.07.048

Bicer, Y., and Khalid, F. (2020). Life cycle environmental impact comparison of solid oxide fuel cells fueled by natural gas, hydrogen, ammonia and methanol for combined heat and power generation. *Int. J. Hydrog.* 45, 3670–3685. doi: 10.1016/j.ijhydene.2018.11.122

Bocklisch, T. (2016). Hybrid energy storage approach for renewable energy applications. *J. Energy Storage.* 8, 311–319. doi: 10.1016/j.est.2016.01.004

Breeze, P. (2019). *Power Generation Technologies (Third edition).* Oxford: Elsevier Ltd. doi: 10.1016/B978-0-08-102631-1.00014-6

Clean Hydrogen JU. (2022). *Clean Hydrogen Joint Undertaking (Clean Hydrogen JU) Work Programme 2022.* Available online at: https://www.clean-hydrogen.europa.eu/system/files/2022-02/Clean%20Hydrogen%20JU%20AWP%202022_0.pdf (accessed June 12, 2022)

Cottrell, C. A., Grasman, S. E., Thomas, M., Martin, K. B., and Sheffield, J. W. (2011). Strategies for stationary and portable fuel cell markets. *Int. J. Hydrog.* 36, 7969–7975. doi: 10.1016/j.ijhydene.2011.01.056

Cucurachi, S., Van Der Giesen, C., and Guinée, J. (2018). Ex-ante LCA of emerging technologies. *Procedia CIRP.* 69, 463–468. doi: 10.1016/j.procir.2017.11.005

Di Florio, G., Macchi, E. G., Mongibello, L., Baratto, M. C., Basosi, R., Busi, E., et al. (2021). Comparative life cycle assessment of two different SOFC-based cogeneration systems with thermal energy storage integrated into a single-family house nanogrid. *Appl. Energy.* 285, 116378. doi: 10.1016/j.apenergy.2020.116378

Di Marcoberardino, G., Manzolini, G., Guignard, C., and Magaud, V. (2018). Optimization of a micro-CHP system based on polymer electrolyte membrane fuel cell and membrane reactor from economic and life cycle assessment point of view. *Chem. Eng. Process.* 131, 70–83. doi: 10.1016/j.cep.2018.06.003

Dincer, I., and Acar, C. (2014). Review and evaluation of hydrogen production methods for better sustainability. *Int. J. Hydrog.* 40, 11094–11111. doi: 10.1016/j.ijhydene.2014.12.035

Dong, Y. H., and Ng, S. T. (2014). Comparing the midpoint and endpoint approaches based on ReCiPe - a study of commercial buildings in Hong Kong. *Int. J. Life Cycle Assess.* 19, 1409–1423. doi: 10.1007/s11367-014-0743-0

Durbin, D. J., and Malardier-Jugroot, C. (2013). Review of hydrogen storage techniques for on board vehicle applications. *Int. J. Hydrog.* 38, 14595–14617. doi: 10.1016/j.ijhydene.2013.07.058

European Commission - Joint Research Centre. (2011). *International Reference Life Cycle Data System (ILCD) Handbook- Recommendations for Life Cycle Impact Assessment in the European context. First edition.* Luxemburg: Publication Office of the European Union. Available online at: https://www.eplca.jrc.ec.europa.eu/ilcd.html

Friedlingstein, P., Jones, M. W., Sullivan, M. O., Andrew, R. M., Andrew, R. M., Bakker, D. C. E., et al. (2022). Global carbon budget 2021. *Earth System Science Data.* 14, 1917–2005, doi: 10.5194/essd-14-1917-2022

Guinée, J. B. (2002). *Handbook on Life Cycle Assessment*. Operational Guide to the ISO Standards. Kluwer Academic Publishers. doi: 10.1007/BF02978897

Hwang, J. J., and Chang, W. R. (2010). Life-cycle analysis of greenhouse gas emission and energy efficiency of hydrogen fuel cell scooters. *Int. J. Hydrog.* 35, 11947–11956. doi: 10.1016/j.ijhydene.2010.07.148

IEA. (2017). *Energy Technology Perspectives 2017*. Paris: Catalysing Energy Technology Transformations. Available online at: https://www.iea.org/reports/energy-technology-perspectives-2017 (accessed March 31, 2022)

IEA. (2019a). *Global Energy and CO2 Status Report 2019.* Paris. Available online at: https://www.iea.org/reports/global-energy-co2-status-report-2019 (accessed March 14, 2022)

IEA. (2019b). *The Future of Hydrogen*. Paris. Available online at: https://www.iea.org/reports/the-future-of-hydrogen (accessed March 10, 2022)

IEA. (2021a). *Global Energy Review 2021*. Paris. Available online at: https://www.iea.org/reports/global-energy-review-2021 (accessed June 9, 2022)

IEA. (2021b). *Global Hydrogen Review 2021.* Paris. Available online at: https://www.iea.org/reports/global-hydrogen-review-2021 (accessed June 9, 2022)

IEA. (2021c). *Net Zero by 2050, A Roadmap for the Global Energy Sector*. Paris. Available online at: https://www.iea.org/reports/net-zero-by-2050 (accessed June 9, 2022)

ISO. (2006). *ISO 14044 Environmental management - Life cycle assessment - Requirements and guidelines.* The International organization for standardization. (2006).

Khan, F. I., Hawboldt, K., and Iqbal, M. T. (2005). Life cycle analysis of wind-fuel cell integrated system. *Renew. Energy*. 30, 157–177. doi: 10.1016/j.renene.2004.05.009

Koj, J. C., Wulf, C., and Zapp, P. (2019). Environmental impacts of power-to-X systems - A review of technological and methodological choices in Life Cycle Assessments. *Renew. Sust. Energ. Rev.* 112, 865–879. doi: 10.1016/j.rser.2019.06.029

Lozanovski, A., Schuller, O., and Faltenbacher, M. (2011). *Guidance Document for Performing LCA on Hydrogen Production Systems. Guidance Document for Performing LCAs on Fuel Cells and H_2 Technologies. 139.* Available online at: https://www.hytechcycling.eu/wp-content/uploads/HY-Guidance-Document.pdf (accessed March, 26 2022)

Lunghi, P., and Bove, R. (2003). Life cycle assessment of a molten carbonate fuel cell stack. *Fuel Cells.* 3, 224–230. doi: 10.1002/fuce.200330124

Lunghi, P., Bove, R., and Desideri, U. (2004). Life-cycle-assessment of fuel-cells-based landfill-gas energy conversion technologies. *J. Power Sources.* 131, 120–126. doi: 10.1016/j.jpowsour.2004.01.006

Masoni, P., and Zamagni, A. (2011). *Guidance Document for performing LCAs on Fuel Cells and Hydrogen Technologies (Hyguide).* Available online at: https://www.hytechcycling.eu/wp-content/uploads/FC-Guidance-Document.pdf

Mazloomi, K., and Gomes, C. (2012). Hydrogen as an energy carrier, Prospects and challenges. *Renew. Sust. Energ. Rev.* 16, 3024–3033. doi: 10.1016/j.rser.2012.02.028

Mekhilef, S., Saidur, R., and Safari, A. (2012). Comparative study of different fuel cell technologies. *Renew. Sust. Energ. Rev.* 16, 981–989. doi: 10.1016/j.rser.2011.09.020

Melamu, R., and von Blottnitz, H. (2009). A comparison of environmental benefits of transport and electricity applications of carbohydrate derived ethanol and hydrogen. *Int. J. Hydrog.* 34, 1126–1134. doi: 10.1016/j.ijhydene.2008.11.020

Mori, M., Gutiérrez, M., and Casero, P. (2021). Micro-grid design and life-cycle assessment of a mountain hut's stand-alone energy system with hydrogen used for seasonal storage. *Int. J. Hydrog.* 46. doi: 10.1016/j.ijhydene.2020.11.155

Mori, M., Jensterle, M., MrŽljak, T., and Drobnič, B. (2014). Life-cycle assessment of a hydrogen-based uninterruptible power supply system using renewable energy. *Int. J. Life Cycle Assess.* 19, 1810–1822. doi: 10.1007/s11367-014-0790-6

Muench, S., and Guenther, E. (2013). A systematic review of bioenergy life cycle assessments. *Appl. Energy*. 112, 257–273. doi: 10.1016/j.apenergy.2013.06.001

Oliveira, L., Messagie, M., Mertens, J., Laget, H., Coosemans, T., and Van Mierlo, J. (2015). Environmental performance of electricity storage systems for grid applications, a life cycle approach. *Energy Convers. Manag.* 101, 326–335. doi: 10.1016/j.enconman.2015.05.063

Ormerod, R. M. (2003). Solid oxide fuel cells. *Chem. Society Rev*. 32, 17–28. doi: 10.1039/b105764m

Ozawa, A., Kudoh, Y., Kitagawa, N., and Muramatsu, R. (2019). Life cycle CO2 emissions from power generation using hydrogen energy carriers. *Int. J. Hydrog.* 44, 11219–11232. doi: 10.1016/j.ijhydene.2019.02.230

Parra, D., Valverde, L., Pino, F. J., and Patel, M. K. (2019). A review on the role, cost and value of hydrogen energy systems for deep decarbonisation. *Renew. Sust. Energ. Rev.* 101(October 2018), 279–294. doi: 10.1016/j.rser.2018.11.010

Peppas, A., Kollias, K., Politis, A., Karalis, L., Taxiarchou, M., and Paspaliaris, I. (2021). Performance evaluation and life cycle analysis of RES-hydrogen hybrid energy system for office building. *Int. J. Hydrog.* 46, 6286–6298. doi: 10.1016/j.ijhydene.2020.11.173

Reuß, M., Grube, T., Robinius, M., Preuster, P., Wasserscheid, P., and Stolten, D. (2017). Seasonal storage and alternative carriers, a flexible hydrogen supply chain model. *Appl. Energy.* 200, 290–302. doi: 10.1016/j.apenergy.2017.05.050

Rinawati, D. I., Keeley, A. R., Takeda, S., and Managi, S. (2022). A systematic review of life cycle assessment of hydrogen for road transport use. *Prog. Energy*. 4, 012001. doi: 10.1088/2516-1083/ac34e9

Rossi, F., Parisi, M. L., Maranghi, S., Basosi, R., and Sinicropi, A. (2020). Environmental analysis of a nano-grid, a life cycle assessment. *Sci. Total Environ.* 700, 134814. doi: 10.1016/j.scitotenv.2019.134814

Sevencan, S., and Çiftcioglu, G. A. (2013). Life cycle assessment of power generation alternatives for a stand-alone mobile house. *Int. J. Hydrog.* 38, 14369–14379. doi: 10.1016/j.ijhydene.2013.09.029

SH2E. (2022). *From FCH - LCA guidelines, through FCH-LCC and SLCA guidelines, to robust FCH-LCSA guidelines and tools.* Available online at: https://sh2e.eu/

Shimizu, T., Hasegawa, K., Ihara, M., and Kikuchi, Y. (2020). A region-specific environmental analysis of technology implementation of hydrogen energy in Japan based on life cycle assessment. *J. Industrial Ecol.* 24, 217–233. doi: 10.1111/jiec.12973

Spitzley, D. V, Keoleian, G. A. and Baron, S. G. (2007). Life cycle energy and environmental analysis of a microgrid power pavilion. *Int. J. Energy Res*. 31:1–13. doi: 10.1002/er.1223

Strazza, C., Del Borghi, A., Costamagna, P., Traverso, A., and Santin, M. (2010). Comparative LCA of methanol-fuelled SOFCs as auxiliary power systems on-board ships. *Appl. Energy*. 87, 1670–1678. doi: 10.1016/j.apenergy.2009.10.012

Stropnik, R., Sekavčnik, M., Ferriz, A. M., and Mori, M. (2018). Reducing environmental impacts of the UPS system based on PEM fuel cell with circular economy. *Energy*. 165, 824–835. doi: 10.1016/j.energy.2018.09.201

Suwanmanee, U., Saebea, D., Hacker, V., Assabumrungrat, S., Arpornwichanop, A., and Authayanun, S. (2018). Conceptual design and life cycle assessment of decentralized power generation by HT-PEMFC system with sorption enhanced water gas shift loop. *Energy Convers. Manag.* 171, 20–30. doi: 10.1016/j.enconman.2018.05.068

Tashie-Lewis, B. C., and Nnabuife, S. G. (2021). Hydrogen production, distribution, storage and power conversion in a hydrogen economy - a technology review. *Adv. Chem. Eng*. 8, 100172. doi: 10.1016/j.ceja.2021.100172

Tranfield, D., Denyer, D., and Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *Br. J. Management.* 14, 207–222. doi: 10.1111/1467-8551.00375

United Nations Environment, P.rogramme. (2020). *Guidelines for Social Life Cycle Assessment of Products 2020.* Available online at: https://www.unep.fr/shared/publications/pdf/DTIx1164xPA-guidelines_sLCA.pdf

Valente, A., Iribarren, D., and Dufour, J. (2016). Life cycle assessment of hydrogen energy systems, a review of methodological choices. *Int. J. Life Cycle Assess.* 22, 346–363. doi: 10.1007/s11367-016-1156-z

Valente, A., Iribarren, D., Dufour, J., and Spazzafumo, G. (2015). Life-cycle performance of hydrogen as an energy management solution in hydropower plants, a case study in Central Italy. *Int. J. Hydrog.* 40, 16660–16672. doi: 10.1016/j.ijhydene.2015.09.104

Varun, Bhat, I. K., and Prakash, R. (2009). LCA of renewable energy for electricity generation systems-a review. *Renew. Sust. Energ. Rev.* 13, 1067–1073. doi: 10.1016/j.rser.2008.08.004

Walker, S. B., Lanen, D., Van, M.,ukherjee, U., and Fowler, M. (2017). Greenhouse gas emissions reductions from applications of Power-to-Gas in power generation. *Sustain. Energy Technol. Assess.* 20, 25–32. doi: 10.1016/j.seta.2017.02.003

Work Energy Council. (2022). *World Energy Insights, Working Paper Regional Insights into Low Carbon Hydrogen. London.* Available online at: https://www.worldenergy.org/publications/entry/regional-insights-low-carbon-hydrogen-scale-up-world-energy-council

Yue, M., Lambert, H., Pahon, E., Roche, R., Jemei, S., and Hissel, D. (2021). Hydrogen energy systems, A critical review of technologies, applications, trends and challenges. *Renew. Sust. Energ. Rev.* 146, 111180. doi: 10.1016/j.rser.2021.111180

Zhang, J., Meerman, H., Benders, R., and Faaij, A. (2021). Techno-economic and life cycle greenhouse gas emissions assessment of liquefied natural gas supply chain in China. *Energy*. 224. doi: 10.1016/j.energy.2021.120049

Zumsteg, J. M., Cooper, J. S., and Noon, M. S. (2012). Systematic review checklist a standardized technique for assessing and reporting reviews of life cycle assessment data. *J. Industrial Ecol.* 16. doi: 10.1111/j.1530-9290.2012.00476.x

TYPE Original Research
PUBLISHED 12 August 2022
DOI 10.3389/frsus.2022.907821

OPEN ACCESS

EDITED BY
Sudhakar Rao,
Indian Institute of Science (IISc), India

REVIEWED BY
Conghui Meng,
Zhejiang University, China
Souknilanh Keola,
Japan External Trade Organization (JETRO), Japan

*CORRESPONDENCE
Akinori Kitsuki
kitsuki@artsci.kyushu-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 30 March 2022
ACCEPTED 14 July 2022
PUBLISHED 12 August 2022

CITATION
Kitsuki A and Managi S (2022) Toward a multidimensional indicator of resident-oriented sustainable development: The case of slum areas in Mumbai. *Front. Sustain.* 3:907821. doi: 10.3389/frsus.2022.907821

Toward a multidimensional indicator of resident-oriented sustainable development: The case of slum areas in Mumbai

Akinori Kitsuki[1]* and Shunsuke Managi[2]

[1]Faculty of Arts and Science, Kyushu University, Fukuoka, Japan, [2]Urban Institute & Department of Civil Engineering, Kyushu University, Fukuoka, Japan

This paper proposes a framework for weighting priority for the multidimensional domains of slum development from the viewpoint of residents. The weights estimated by our framework can be interpreted as marginal utility, and multiplying satisfaction scores in each domain of development by these weights yields residents' utility from slum development. The proposed approach is carried out by accessing residents' needs for slum development in Mumbai, India, where more than 5.2 million residents live in slum areas. Using the dataset obtained from a questionnaire survey that we conducted with slum residents in March 2019, we estimate marginal utilities for 23 domains of slum development. The results show that (1) slum residents most wish to improve development projects about human capital (health and education), electricity, and drinking water, even if they have already been highly satisfied, (2) they feel the least satisfaction with public toilets and place high priority on projects involving public toilets, (3) the sewage system has low priority with low satisfaction, but this priority increases as slum residents become better off, (4) the projects on social development and protection are highly satisfied and generally ranked low in terms of marginal utilities, and (5) air pollution and working conditions are also concerns of slum residents, especially as these residents become better off. It is expected that the framework used in this paper can be used to extract the problems of urban development and to track the progress of development plans from the viewpoint of residents.

KEYWORDS
multidimensional index of development, resident-oriented approach, subjective wellbeing, slum, India

Introduction

The improvement of slum conditions toward urban sustainable development is a global challenge. The number of slum dwellers is increasing in most developing countries, and in 2018, more than one billion people were living in slum areas (UN-Habitat, 2020), characterized by a lack of the minimum quality of life. Sustainable Development Goal (SDG) 11 seeks to ensure access for all to adequate, safe and affordable housing and basic services and to upgrade slum areas by 2030 (United Nations, 2015). To achieve this goal, appropriate indicators are required to trace the development progress. Objective measures, such as per capita net domestic product, literacy rate, and infant mortality rate, are commonly used as indicators. However, the importance of subjective evaluations that are directly associated with people's wellbeing has also received attention (e.g., Stiglitz et al., 2018). Incorporating subjective measures into the development indicator and reflecting these findings in the policy goals are thus urgently needed.

This paper aims to propose resident-oriented satisfaction measures of slum development. The important aspect of such an indicator is the multidimensionality of the domains of which it is comprised. For example, access to improved water sources, access to improved sanitation facilities, and security tenure all must be evaluated, and each evaluated indicator should be used to grasp the progress of each domain of slum development. However, to understand how much the improvement of one domain of slum development improves slum residents' overall satisfaction and to prioritize each domain, it is also necessary to understand the relative importance of each domain. We provide a theoretical framework through which to estimate weights to aggregate residents' satisfaction with the individual domain of slum development and apply it to the 2019 survey carried out with slum residents in Mumbai, India, where the total number of slum dwellers is 5,206,473, which is approximately 41.84% of the total population of Mumbai (Census, 2011).

One previous study on weights in multidimensional indices is that of Decancq and Lugo (2013), which used an extensive survey to broadly classify weighting approaches into three categories: data-based weighting, normative weighting, and a hybrid of the two. The first category is a data-based weighting approach such as principal component analysis (Noorbakhsh, 1998; Klasen, 2000), factor analysis and structural equation modeling (Krishnakumar and Nadar, 2008). The second category, normative weighting, is premised on a specific kind of value judgment and is based on expert opinions (Chowdhury and Squire, 2006; Mascherini and Hoskins, 2008) or on arbitrary external weight values (Diener and Suh, 1997; Lugo, 2007; Nilsson, 2010). This type of weighting also includes simple averages, such as the Human Development Index, which are most often used in practice. Data-based weighting has the problem of there being a possible mismatch between actual achievement and residents' value judgments, while normative weighting has the problem that it depends on who makes the value judgments. The third method presented to solve these problems is a hybrid of the two, which uses data including value judgment for weighting (hereafter referred to as "hybrid type"). This method can be broadly divided into two types. One is the hedonic type (Welsch, 2006; Schokkaert, 2007; Ferreira and Moro, 2010; Levinson, 2012), which regresses self-reported indicators of wellbeing such as happiness and life satisfaction on variables that capture its various domains and weights them through coefficients. The other is a stated preference method that directly asks residents about their priorities and importance for different domains (De Kruijk and Rutten, 2007; Bossert et al., 2013). We do not adopt the hedonic method because, as pointed out by Decancq and Lugo (2013), it would yield unstable weights due to the multicollinearity problem, as the number of domains that constitute wellbeing increases. In this study, we propose a hybrid-type weighting method with the stated preference method.

One problem faced by the previous research that has used the stated preference method (e.g., De Kruijk and Rutten, 2007; Bossert et al., 2013) is that the derivation of weights has been *ad hoc* and that the theoretical significance of the numbers themselves has not been established for a long time. Benjamin et al. (2014) conducted a pioneering study that provided a theoretical foundation for the weighting of these stated preferences. Based on the utility theory of microeconomics, the above authors regarded utility as a set of fundamental aspects of wellbeing and constructed a theoretical framework that connects the indicators whose value can be assessed through research with the welfare level. The present study is the first attempt at applying the theoretical framework presented by Benjamin et al. (2014) to weight the multidimensional domains of slum development.

The rest of the paper is organized as follows. Section Weighting approach for the multidimensional indicator discusses the weighting approach, Section Data collection describes the data, and Section Results provides the results. Finally, Section Conclusions concludes the paper.

Weighting approach for the multidimensional indicator

Theoretical background

This subsection provides the theoretical background concerning weighting on the multidimensional domains of slum development from the viewpoint of residents. Analogous to Benjamin et al. (2014), in which fundamental aspects of wellbeing were considered components of utility, this paper considers an agent's utility function obtained from multidimensional domains of slum development, $U(\mathbf{w})$, where $\mathbf{w} = (w_1, w_2, \cdots, w_J)$ represents the agent's satisfaction with J individual domains of slum development. Considering an arbitrary vector, $\bar{\mathbf{w}}$ (assumed to be satisfaction at the status quo), and the deviation $\Delta\mathbf{w} = (w_1 - \bar{w}_1, w_2 - \bar{w}_2, \cdots, w_J - \bar{w}_J)$ from it, a Taylor expansion around $\bar{\mathbf{w}}$ results in the following first-order approximation:

$$U(\mathbf{w}) = U(\bar{\mathbf{w}} + \Delta\mathbf{w}) \approx U(\bar{\mathbf{w}}) + \sum_{j=1}^{J} \frac{\partial U(\bar{\mathbf{w}})}{\partial w_j}(w_j - \bar{w}_j) \quad (1)$$

Note that since utility level $U(\mathbf{w})$ in Equation (1) is a first-order approximation of the Taylor expansion, the resulting approximation value becomes less accurate if evaluation vector $\mathbf{w}$ deviates largely from $\bar{\mathbf{w}}$. Since $U(\bar{\mathbf{w}})$ and $\sum_{j=1}^{J} \frac{\partial U(\bar{\mathbf{w}})}{\partial w_j} \bar{w}_j$ are constants, $\sum_{j=1}^{J} \frac{\partial U(\bar{\mathbf{w}})}{\partial w_j} w_j$, that is, the weighted average of the agent's satisfaction with J individual domains of slum development ($\mathbf{w}$) using weights as the marginal utility at the status quo, can be seen as the agent's index that tracks small changes in utility from slum development. Thus, once marginal utilities $\frac{\partial U(\bar{\mathbf{w}})}{\partial w_j}$ at a base period are derived from the survey, we can track residents' changes in utility from slum development if the survey that collects the data on individual satisfaction ($\mathbf{w}$) is regularly conducted. Thus, this paper focuses mainly on the derivation of marginal utilities. The following subsection discusses the procedures through which to estimate marginal utility from the survey.

Estimate of marginal utility

To estimate marginal utility, the following random utility is assumed for each domain of slum development j $(j = 1, \cdots, J)$:

$$U_{ij} = \beta_j x_j + v_i + \epsilon_{ij} \tag{2}$$

Here, subscript "i" represents an individual, subscript "j" represents a domain of slum development, and U_{ij} represents the utility that individual i obtains from domain j. x_j is a dummy variable that takes the value of 1 when the standards of slum development in domain j improve and 0 otherwise, v_i represents the utility from slum development in domain j at the status quo, and ϵ_{ij} represents the error terms. The coefficient β_j can be interpreted as marginal utility, which represents a change in utility when the level of utility obtained from slum development in domain j improves. To estimate the coefficient β_j, we take any two domains, s and t, out from the J domains of slum development and define U_i^* as

$$U_i^* \equiv \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_J x_J + (\epsilon_{is} - \epsilon_{it}) \tag{3}$$

where x_j $(j = 1, \cdots, J)$ takes the value of 1 if j is equal to s, -1 if j is equal to t, and 0 otherwise. Then, individual i prefers to improve domain s if $U_i^* = U_{is} - U_{it} > 0$ and prefers to improve domain t if $U_i^* = U_{is} - U_{it} < 0$. As explained in detail in the following section, in our questionnaire survey, we present two randomly chosen projects from the list of all domains and ask respondents to respond with which project they prefer in terms of improving their welfare from the status quo. The response choices are on a 4-point Likert-type scale: (1) definitely prefer project s, (2) slightly prefer project s, (3) slightly prefer project t, and (4) definitely prefer project t. Using Equation (4), this selection problem can be formulated as

$$U_i^* = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_J x_J + \epsilon_i \tag{4}$$

$$y_i = 4 \text{ } if \text{ } U_i^* > \alpha_1 \tag{5}$$

$$y_i = 3 \text{ } if \text{ } \alpha_1 \geq U_i^* \geq \alpha_2 \tag{6}$$

$$y_i = 2 \text{ } if \text{ } \alpha_2 \geq U_i^* \geq \alpha_3 \tag{7}$$

$$y_i = 1 \text{ } if \text{ } \alpha_3 > U_i^* \tag{8}$$

where U_i^* is the latent variable, y_i corresponds to the 4-point Likert-type response choice, and ϵ_i is $\epsilon_{is} - \epsilon_{it}$. Thus, the coefficient β_j can be estimated by ordered probit estimation with the common assumption that ϵ_{ij} follows the Gumbel distribution.

Data collection

This study relies on a face-to-face questionnaire survey conducted in Mumbai, India, in March 2019 through a third-party company (INTAGE INDIA). The stratified sampling method was adopted to collect the data. First, we obtained the list of all slums from the Slum Rehabilitation Authority, Mumbai, which included information on the location and size of each slum. Regarding the definition of a slum, the list follows that used in the Census of India (Census of India, 2001):

> *"For the purpose of Census of India (2001); the slum areas broadly constitute: (i) All specified areas in a town or city notified as 'slum' by State/Local Government and Union Territories (UT) Administration under any Act including a 'Slum Act'; (ii) All areas recognized as 'Slum' by State/Local Government and UT Administration, Housing and Slum Boards, which may have not been formally notified as slum under any act; and (iii) A compact area of at least 300 population or about 60e70 households of poorly built congested tenements, in unhygienic environment usually with inadequate infrastructure and lacking in proper sanitary and drinking water facilities."*

Slum size was classified into three categories: a slum was classified as large if its area was larger than 10,000 m^2, as medium-sized if its area was between 2,500 m^2 and 9,999 m^2, and as small if its area was smaller than 2500 m^2. Then, the slum was stratified by tehsil (a local unit of administrative division in Mumbai, which consists of 18 tehsils), slum size was proportional to the number of slums in the list, and a total of 102 slums were selected. From each slum, 10 samples for small slums, 30 samples for medium-sized slums, and 50 samples for large slums were randomly selected. To control quality of interviewers, brief session on survey procedures are hold before the survey, and qualification test interviews are conducted to check their understanding. A tablet-based digital

TABLE 1 List of development projects.

Project		Examples
A. Basic infrastructure		
1	Drinking water (piped water)	Expanding the coverage of a water work system (access to piped water), improving the quality of piped water
2	Sewage system (closed drainage)	Expanding the coverage of a sewage system (access to closed drainage), improving the quality (cleanliness) of closed drainage
3	Public toilets	Increasing the number of public toilets, improving the cleanliness of existing public toilets
4	Road improvement	Mitigating traffic jams on the main road, maintenance of narrow road width (e.g., nothing larger than a motorcycle can pass on the road outside)
5	Electricity	Expanding the coverage of the electricity power system, stabilizing the electricity power system (solving the problem of blackouts), lowering the prices of electric service, increasing the number of streetlights
6	Property rights for housing (possibility of your eviction)	Securing the right of residence (the government will not destroy your house)
B. Environmental management		
7	Garbage disposal	Increasing the number of garbage disposal spots, decreasing the amount of garbage on the street
8	Air pollution	Increasing air quality
9	Parks and greenery	Increasing the number of parks, increasing the number of roadside trees
C. Human capital		
10	Education (school)	Improving the quality of teaching, decreasing the costs for public school, improving school equipment
11	Health (hospital)	Improving the quality of doctors, improving the medical facilities in the clinic, decreasing the costs for the clinic, mitigating hospital congestion
D. Transfer for basic needs		
12	Financial support for elderly, widowed, and disabled people	Improving financial support policies for elderly, widowed, and disabled people
13	Support in emergencies (flood, fires, and so on)	Mitigating damage due to emergencies
14	Ration	Improving the ration system
E. Private sector management		
15	Job opportunities	Increasing the number of job opportunities
16	Working conditions (wages and working environment)	Improving physical working conditions (health and safety at work)
17	Access to microfinance	Being able to borrow money when you wish (with a moderate interest rate)
F. Law and order		
18	Safeness in the neighborhood	Improving neighborhood safety
19	Domestic violence/abuse	Improving policies to help the victims of domestic violence
20	Corruption, injustice, and abuse of power	Mitigating the corruption, injustice, and abuse of power of elected officials, Pradhan, policies, or the judicial system
G. Social development and protection		
21	Opportunities to participate in community meetings	Improving social connections
22	Freedom of religion and beliefs	Freedom of religion and beliefs
23	Freedom of speech	Freedom of speech

questionnaire is used in this study, and unrealistic answers and inconsistent responses are detected by the program and marked accordingly with a short alert message. The wording used in the questionnaire was carefully tested in the pre-survey so that all the slum residents could easily understand. A total of 3,111 respondents participated in this survey. The survey language was either Hindi, Marathi, or English. Among those 3,111 respondents, 2,323 respondents (74.67%) used Hindi, 786 (25.27%) used Marathi, and 2 (0.06%) used English. In addition to respondents' demographic information, the data about respondent preferences for development projects were collected as described below.

TABLE 2 Satisfaction with public projects.

		1. Completely dissatisfied	2. Somewhat dissatisfied	3. Neither satisfied nor dissatisfied	4. Somewhat satisfied	5. Completely satisfied	Average
1	Electricity	4.5%	5.4%	12.5%	33.2%	44.4%	4.08
2	Freedom of speech	4.1%	6.6%	14.9%	33.7%	40.7%	4.00
3	Education (school)	4.1%	6.8%	13.1%	38.0%	38.0%	3.99
4	Drinking water	5.9%	7.9%	11.6%	32.4%	42.2%	3.97
5	Health (hospital)	4.5%	7.2%	15.3%	37.8%	35.2%	3.92
6	Safeness in the neighborhood	5.0%	7.3%	14.2%	39.0%	34.6%	3.91
7	Freedom of religion and beliefs	5.4%	8.1%	15.5%	39.4%	31.7%	3.84
8	Opportunities to participate in community meetings	4.7%	8.6%	17.6%	37.5%	31.6%	3.83
9	Domestic violence/abuse	4.9%	9.7%	17.4%	36.2%	31.9%	3.80
10	Garbage disposal	8.2%	8.5%	16.8%	31.5%	35.1%	3.77
11	Job opportunities	7.9%	8.4%	19.4%	29.3%	34.9%	3.75
12	Support in emergencies (flood, fires, and so on)	6.0%	9.5%	18.5%	37.3%	28.6%	3.73
13	Air pollution	7.0%	8.9%	19.2%	35.9%	29.0%	3.71
14	Parks and greenery	8.3%	8.6%	17.6%	34.9%	30.6%	3.71
15	Access to microfinance	6.8%	9.5%	19.1%	35.1%	29.5%	3.71
16	Property rights for housing	7.0%	9.3%	18.0%	38.0%	27.8%	3.70
17	Working conditions (wages and working environment)	6.8%	10.2%	17.9%	37.7%	27.4%	3.69
18	Road improvement	7.5%	8.2%	18.1%	40.7%	25.5%	3.68
19	Ration	9.4%	10.9%	14.3%	36.4%	29.1%	3.65
20	Financial support for elderly, widowed, and disabled people	7.0%	11.1%	19.4%	36.4%	26.2%	3.64
21	Corruption, injustice, and abuse of power	7.4%	11.9%	17.6%	35.2%	27.9%	3.64
22	Sewage system	7.9%	10.3%	20.3%	35.9%	25.5%	3.61
23	Public toilets	11.8%	9.1%	17.7%	31.5%	29.9%	3.58

To estimate marginal utilities $\frac{\partial U(\bar{\mathbf{w}})}{\partial w_j}$ $(j = 1, \cdots, J)$, we first identified 23 development projects that are relevant to slum residents in Mumbai to compile the list of the domains of slum development. [1]One problem faced by the ordered probit estimation for Equations (4)–(8) was that only the differences from one base marginal utility could be identified due to the rank condition. To identify marginal utilities for all 23 domains of the slum project, a hypothetical project that plays no role was added to the list as a base domain (that is, $J = 24$), and its marginal utility was assumed to be zero for normalization. In the main survey, respondents were asked to compare two different development projects. First, two projects were randomly chosen from the list of 24 projects and presented to respondents with a brief example of the projects. Table 1 shows the list of development projects with brief examples described by the enumerators in the survey. Then, respondents were asked to respond as to which projects they preferred in terms of improving their welfare from the status quo. Response choices were on a 4-point Likert-type scale: (1) definitely prefer the first project, (2) slightly prefer the first project, (3) slightly prefer the second project, and (4) definitely prefer the second project. Questions were repeated 20 times for each respondent.

As a variable for individual satisfaction ($\mathbf{w}$), for all 23 development projects, respondents were asked to rate how satisfied they were with them. Response choices were on

1 To identify these 23 development projects, we first prepared the comprehensive list of 129 development projects based on the taxonomy that the World Bank has identified in July 2016 (https://projects.worldbank.org/en/projects-operations/project-theme?lang=en&page=). Among this list of the projects, we select development projects that are relevant to the slum dwellers in Mumbai through the discussion with several local field experts who were born in slums in Mumbai.

TABLE 3 Descriptive statistics (demographics).

	Full sample	By asset quantile					Regression analysis
		1	2	3	4	5	
Caste							
General	49.7%	35.0%	53.2%	50.6%	42.1%	69.8%	(Base)
Schedule caste (SC)	24.3%	35.1%	22.5%	23.2%	28.3%	10.9%	−0.331***
							(0.040)
Other backward caste (OBC)	20.7%	22.2%	19.8%	21.3%	24.8%	15.4%	−0.117***
							(0.041)
Schedule tribe (ST)	2.9%	5.5%	2.7%	1.8%	2.4%	1.9%	−0.270***
							(0.093)
Others	2.3%	2.3%	1.8%	3.2%	2.4%	1.9%	0.103
							(0.140)
Place of birth							
Same area living at present	32.9%	17.4%	27.8%	33.0%	28.3%	59.2%	(Base)
Other area within Mumbai	25.6%	9.5%	20.9%	27.1%	44.1%	28.0%	−0.120***
							(0.042)
Outside Mumbai but within Maharashtra	15.2%	13.5%	19.6%	15.5%	19.3%	8.8%	−0.355***
							(0.050)
Outside Maharashtra but within India	26.1%	59.1%	31.6%	24.5%	8.2%	3.9%	−0.860***
							(0.044)
Outside India	0.2%	0.4%	0.2%	0.0%	0.2%	0.2%	−0.608*
							(0.343)
Demographics of HH head							
Years in Mumbai	24.83	17.10	24.07	26.37	26.89	30.56	0.015***
							(0.001)
Years of education	10.40	8.70	10.05	10.48	10.44	12.55	0.062***
							(0.005)
Both can read and write	87.1%	78.2%	88.0%	87.2%	86.1%	97.1%	0.114***
							(0.031)
Constant							−0.733***
							(0.097)
Observations							2,894
R-squared							0.335

Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

a 5-point Likert-type scale: (1) completely dissatisfied, (2) somewhat dissatisfied, (3) neither satisfied nor dissatisfied, (4) somewhat satisfied, and (5) completely satisfied.

Results

Satisfaction with development projects

Table 2 shows the distribution of responses regarding satisfaction and its average score on each development. The projects were sorted by their average scores in descending order. The highest average score for satisfaction was for electricity (4.08), followed by freedom of speech (4.00) and education (3.99). Overall, the basic infrastructure of electricity and drinking water, infrastructure of human capital (school and hospital), and social development and protection (freedom of speech, freedom of religion and beliefs, and participation in community meetings) were rated as being highly satisfying, while the basic infrastructure of the sewage system and public toilets were the least satisfying for respondents.

TABLE 4 Descriptive statistics (housing).

	Full sample	By asset quantile				
		1	2	3	4	5
Housing structure						
Pucca house	45.5%	2.6%	8.3%	30.1%	89.1%	98.2%
Semipucca house	49.0%	85.2%	81.5%	65.2%	10.8%	1.8%
Kutcha house	5.5%	12.2%	10.2%	4.7%	0.2%	0.0%
Roofing material						
Tiles	19.6%	0.1%	0.4%	2.3%	23.0%	72.2%
Burnt brick, stone, slate, concrete	25.9%	2.4%	8.0%	27.9%	66.1%	26.0%
Grass, thatch, bamboo, wood, mud, plastic and polythene	5.9%	10.9%	11.8%	5.6%	0.8%	0.5%
Asbestos, tin sheets	48.5%	86.5%	79.9%	64.1%	10.1%	1.3%
Wall material						
Stone (packed with mortar)	15.0%	0.9%	2.9%	3.2%	8.2%	60.0%
Stone (not packed with mortar)	0.2%	0.3%	0.4%	0.3%	0.0%	0.0%
Burnt brick with cement plaster covered by tile, marble	10.6%	0.1%	1.1%	6.0%	22.8%	23.2%
Burnt brick with cement plaster	44.0%	28.5%	56.3%	68.9%	56.1%	13.5%
Burnt brick without cement plaster	24.9%	58.3%	29.6%	17.2%	12.7%	3.4%
Wood, asbestos sheets, tin sheets	2.6%	4.9%	5.1%	2.9%	0.0%	0.0%
Grass, thatch, bamboo, wood, mud, plastic and polythene	1.4%	3.7%	2.2%	0.6%	0.0%	0.0%
Mud unburnt brick	1.4%	3.3%	2.5%	0.8%	0.2%	0.0%
House ownership						
Owned with official documents	34.1%	6.0%	19.2%	34.0%	40.8%	71.9%

Asset ownership

To see the differences in the living conditions among slum residents, we measured asset index and created asset quantiles. An asset index was calculated using principal component analysis to assign indicator weights. For this analysis, household possessions [radio, black and white TV, color TV, refrigerator, washing machine, bicycle, motorcycle, car, landline phone, mobile phone, ceiling fan, liquefied petroleum gas (LPG) stove, laptop, air conditioner, mattress, pressure cooker, chair, cot or bed, sewing machine, or clock], house ownership, house structure, room material, wall material, access to electricity, access to drinking water, latrine facility, and drainage system were all considered. The average value of the constructed asset index was zero with a standard deviation of one. The poorest quantile of the asset index ranged from −2.428 to −0.819, the second- poorest quantile ranged from −0.818 to −0.368, the third-poorest quantile ranged from −0.368 to 0.216, the fourth-poorest quantile ranged from 0.218 to 0.848, and the fifth-poorest, that is, richest, quantile ranged from 0.851 to 2.450.

Table 3 represents the relation between slum residents' demographics and asset index. The first column reports the descriptive statistics of the full sample, and the second to sixth columns report those by asset quantile. While all the demographics are present throughout the within-slum asset distribution, those individuals whose history of residence in Mumbai was shorter, especially those from outside Mumbai, those whose education level was lower, and those who belonged to schedule caste (SC), other backward caste (OBC), or schedule tribe (ST) were disproportionately represented among the poorer asset quantiles. In particular, this tendency was strong for the poorest quantile of slum residents. To see how the asset index relates to these slum dwellers' demographics, we conducted ordinary least squares (OLS) regression analysis in which the dependent variable was the asset index, and the regressors were demographic variables. The results showed that all these variables significantly affect the asset index.

In India, there are three general housing classes: pucca houses, kutcha houses, and semipucca houses. A pucca house refers to a house that is designated to be solid and permanent, a kutcha house refers to a house that is designated to be temporary and less solid, and a semipucca house is a combination of the two. Usually, these classifications are based on house materials. Table 4 reports the information of the condition of respondents' houses. Most of the richest quantile of respondents lives in pucca houses with tile roofing and with walls made of stone backed with mortar or by burnt bricks with cement plaster covered by tile or marble. The second-richest quantile of respondents also lives in pucca houses, but with the roof made of burnt brick, stone, slate or concrete and with walls made of burnt

TABLE 5 Descriptive statistics (household asset possessions).

	Full sample	By asset quantile				
		1	2	3	4	5
Household electric appliances						
LPG stove	86.3%	69.8%	90.0%	89.7%	89.2%	95.2%
Color TV	79.3%	47.2%	83.1%	85.8%	89.9%	94.7%
Pressure cooker	76.1%	65.0%	87.8%	89.4%	90.7%	50.2%
Refrigerator	40.0%	2.3%	14.7%	42.7%	61.9%	79.9%
Sewing machine	15.8%	6.9%	10.9%	13.7%	13.0%	35.0%
Washing machine	15.0%	1.3%	4.7%	13.5%	13.3%	42.4%
Radio	5.7%	9.9%	6.9%	6.9%	2.9%	1.3%
Air conditioner	4.3%	1.6%	3.6%	3.9%	3.4%	9.3%
Black and white TV	2.4%	2.9%	4.4%	3.4%	1.4%	0.3%
Telecommunications device						
Mobile phone	93.3%	92.2%	92.9%	92.6%	93.4%	95.7%
Laptop	9.5%	2.2%	4.4%	6.1%	8.5%	26.4%
Landline phone	3.1%	3.5%	2.9%	3.1%	3.2%	2.7%
Vehicles						
Motorcycle	8.5%	1.3%	3.4%	7.2%	9.3%	21.4%
Bicycle	5.6%	2.9%	8.2%	6.8%	3.7%	6.9%
Car	1.3%	0.4%	1.8%	1.3%	1.6%	1.4%
House furnishings						
Ceiling fan	95.2%	87.2%	94.9%	96.5%	98.4%	99.8%
Watch or clock	69.8%	45.3%	75.9%	81.8%	76.8%	72.5%
Chair	55.1%	20.0%	51.4%	67.0%	66.7%	74.1%
Mattress	54.2%	38.4%	59.0%	68.8%	62.2%	45.2%
Cot or bed	48.0%	33.7%	35.9%	52.7%	51.6%	66.6%

brick. Moreover, most of the first- to third-poorest-quantile respondents live in houses with roofing covered by asbestos or tin sheets and with walls made of burnt brick without marble. The walls of the houses of more than half of the respondents in the poorest quantile are made of burnt brick that is not covered by cement plaster. Regarding housing ownership status, less than 10% of respondents in the poorest quantile own their own house, that is, have official documents verifying their ownership status, while more than 70% of respondents in the richest quantile own their own house.

Table 5 reports the percentage of respondents that possess each household asset. For household electric appliances, although, on average, more than 75% of all the respondents own LPG stoves, color TVs, and pressure cookers, less than 70% of respondents in the poorest quantile own them. Such disparities in asset possession rates are even worse for refrigerators, sewing machines, and washing machines. For all quantiles, only a few households possess radios and air conditioners, as well as black and white TVs. Regarding telecommunications devices, most respondents own mobile phones. Less than 10% of respondents in the first- to -fourth-poorest quantile own laptops or motorcycles, while 26.4 and 21.4% of those in the richest quantile possess them, respectively. Regarding house furnishings, approximately 50% of respondents on average do not own a mattress, cot or bed, which implies that they sleep directly on the floor. The situation is worse for the poorest quantile of respondents.

Table 6 reports slum residents' access to basic infrastructure. On average, more than 90% of respondents have access to electricity with a meter. However, if we look at the poorest quantile of respondents, only 65.5% have access to electricity with a meter. This unequal situation is also true for sources of drinking water. On average, 82.3% of respondents have access to tap water from treated sources, while only 49.6% of those in the poorest quantile have such access. In this poorest quantile, 42.6% of them drink tap water from untreated source, and 4.6% from hand pump and tube well. Looking at the location of drinking water, only 3.7% of respondents in this quantile have access to drinking water within their premise. In terms of access to a drainage system, we observe uneven access across asset quantiles. The percentage of those who have access to closed drainage, which consists of covered channels and prevents

TABLE 6 Descriptive statistics (access to public goods).

	Full sample	By asset quantile				
		1	2	3	4	5
Access to electricity						
Have electricity with a meter	90.1%	65.6%	92.4%	96.9%	99.2%	99.5%
Have electricity without a meter	9.5%	32.8%	7.4%	2.7%	0.8%	0.5%
No electricity	0.5%	1.6%	0.2%	0.3%	0.0%	0.0%
Source of drinking water						
Tap water from treated source/BMC pipeline	82.3%	49.6%	89.8%	94.5%	90.5%	91.8%
Tap water from untreated source	15.8%	42.6%	9.4%	5.2%	9.5%	8.2%
Hand pump and tube well/borehole	1.1%	4.6%	0.2%	0.0%	0.0%	0.0%
Water tanker	0.8%	3.2%	0.5%	0.0%	0.0%	0.0%
Covered/uncovered well	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%
Location of drinking water						
Within the premise	50.4%	3.7%	29.2%	56.8%	71.9%	93.2%
Near the premise (less than 100 m, approx. 1 minute by walk)	32.1%	65.2%	50.5%	21.4%	17.8%	3.9%
More than a 1-minute walk	17.5%	31.1%	20.3%	21.7%	10.3%	2.9%
Type of drainage system						
Closed drainage	60.8%	39.1%	45.4%	53.3%	72.0%	94.7%
Open drainage	37.3%	59.6%	51.5%	43.6%	25.9%	4.7%
No specific drainage facility available	2.0%	1.3%	3.1%	3.1%	2.1%	0.6%
Location of latrine/toilet						
Having a latrine/toilet facility within the premises	18.7%	0.4%	1.3%	6.1%	12.5%	73.5%
Use public latrine/toilet	77.9%	91.9%	93.8%	91.8%	86.7%	25.4%
Use public latrine/toilet during day time and open defecation at night	1.8%	3.6%	2.7%	1.3%	0.5%	0.8%
No latrine/toilet facility available, use open defecation	1.6%	4.0%	2.2%	0.8%	0.3%	0.3%

human exposure to harmful waste flows, is only 39.1% among those in the poorest quantile, but this rate increases, up to 94.7%, as the quantile of the asset index increases for those in the richest quantile. Regarding access to toilets, except for those in the richest quantile, most respondents rely on public toilets.

Estimates of marginal utility

Based on the current situation of slum residents discussed above, we address the needs for slum development by estimating marginal utility, the estimated results of which are reported in Table 7. The first column of Table 7 illustrates the marginal utilities of all the samples, and the projects in the table are ranked by these marginal utilities. The coefficients can be negative if respondents give a lukewarm answer, because they are the differences between marginal utility of each development project and that of hypothetical project that plays no role. However, the results show that all the coefficients are positive.

The estimated marginal utilities of each projects and its corresponding average satisfaction score reported in the last column of Table 2 are summarized in Figure 1. X-axis represents the average score of satisfaction, y-axis represents the marginal utilities, the horizontal dotted line represents the median value of marginal utilities, and the vertical dotted line represents the median value of average satisfaction score. First, the projects on drinking water as well as health and education are shown on the upper right part of Figure 1. Among the projects on basic infrastructure, clean drinking water is essential for people's life and health condition. These results imply that the projects that directly relate to human capital are most wished to be improved even if they have already been highly satisfied. Note that, according to Yesudhas (2021), the percentage of respondents of the survey of slum area in Mumbai that can use water less than 4 h per days reached up to 72 and 45% among those who use public taps and private taps, respectively. As a result, only 25% of their respondents have access to 20 or more liters of water per person per day, which is a minimum amount of requirement for water. Regarding the access to water, not only the expanding the coverage of a water work system, available time for water use should also be an important issue. Second, the projects on basic infrastructure are generally highly ranked in terms of marginal utilities. Among them, the project on public toilet is one of the projects that should be paid attention, because it

TABLE 7 Marginal utility estimates.

	Whole sample			Gender					
				Male			Female		
Variable	**Rank**	**Coef.**	**SE**	**Rank**	**Coef.**	**SE**	**Rank**	**Coef.**	**SE**
Health	1	0.064	0.004	1	0.065	0.006	3	0.064	0.005
Education	2	0.062	0.004	2	0.062	0.006	5	0.063	0.005
Drinking water	3	0.060	0.004	5	0.054	0.006	2	0.066	0.005
Public toilet	3	0.060	0.004	6	0.053	0.005	1	0.067	0.005
Ration	5	0.057	0.004	10	0.050	0.005	3	0.064	0.005
Garbage disposal	6	0.056	0.004	4	0.055	0.006	6	0.056	0.005
Job opportunities	7	0.055	0.004	3	0.057	0.005	7	0.053	0.005
Road improvement	8	0.049	0.004	7	0.052	0.005	11	0.046	0.005
Air pollution	8	0.049	0.004	9	0.051	0.006	9	0.048	0.005
Electricity	10	0.046	0.004	13	0.044	0.006	9	0.048	0.005
Working conditions	11	0.043	0.004	7	0.052	0.006	18	0.033	0.005
Safeness in the neighborhood	11	0.043	0.004	17	0.036	0.006	8	0.050	0.005
Parks and greenery	11	0.043	0.004	12	0.045	0.006	12	0.041	0.005
Financial support for elderly, widowed, and disabled people	14	0.042	0.004	14	0.043	0.006	12	0.041	0.005
Freedom of speech	14	0.042	0.004	10	0.050	0.006	17	0.034	0.005
Domestic violence/abuse	16	0.036	0.004	15	0.037	0.005	16	0.035	0.005
Property rights for housing	16	0.036	0.004	18	0.032	0.005	14	0.040	0.005
Corruption, injustice, and abuse of power	18	0.035	0.004	15	0.037	0.006	18	0.033	0.005
Sewage system	19	0.033	0.004	20	0.028	0.006	15	0.038	0.005
Support in emergencies (flood, fires, and so on)	20	0.028	0.004	19	0.030	0.005	20	0.025	0.005
Access to microfinance	21	0.024	0.004	21	0.025	0.005	21	0.023	0.005
Opportunities to participate in community meetings	22	0.022	0.004	22	0.024	0.005	22	0.020	0.005
Freedom of religion and beliefs	23	0.016	0.004	23	0.019	0.005	23	0.014	0.005
Observations		61,740			33,660			28,080	

All the coefficients are statistically significant at the 1% level.

is ranked as the third-highest marginal utilities, but with the worst satisfaction scores among all types of projects, shown in the upper left part of Figure 1. Overcrowded public toilets due to their limited number and the poor toilet conditions mainly caused by lack of public participation in their management are reported in the slum area of Mumbai (e.g., Biswas et al., 2020; Yesudhas, 2021). Since 77.9% of respondents rely on public toilets (Table 6), improving the public toilet should be an acute problem of slums in Mumbai. Another project that should be noted is that involving the sewage system. Among all the basic infrastructure projects, it ranks lowest in terms of marginal utilities. According to Subbaraman et al. (2013), although all bacterial contamination of drinking water occurred due to post-source contamination during storage in the household, point-of-source water contamination occurs in the monsoon season. In the slum area of Mumbai, only a few toilets are connected to a main sewer line and this situation could cause a contamination of groundwater (Biswas et al., 2020). Thus, the improvement of sanitation, as well as public education on household water treatment, would reduce water born disease caused by the contamination. In this sense, the importance of sewerage system might be under-evaluated by slum residents. Third, the projects on environmental management are also highly ranked in terms of marginal utilities with garbage disposal ranking 6th and air pollution ranking 8th, which are as high as the project on Job opportunities (7th rank) on road improvement (8th rank). Slum residents are found to care about environmental issues as greatly as some of the important projects which directly relate to their economic condition. Lastly, the projects on social development and protection are generally ranked low in terms of marginal utilities with opportunities to participate in community meetings ranking 22th, and freedom of religion and beliefs ranking the worst. These two projects are shown in the lower right part of Figure 1, indicating that slum residents are already satisfied with these projects.

To see whether these rankings are heterogeneous among respondents, we estimated marginal utilities separately for male respondents and for female respondents, with the second and

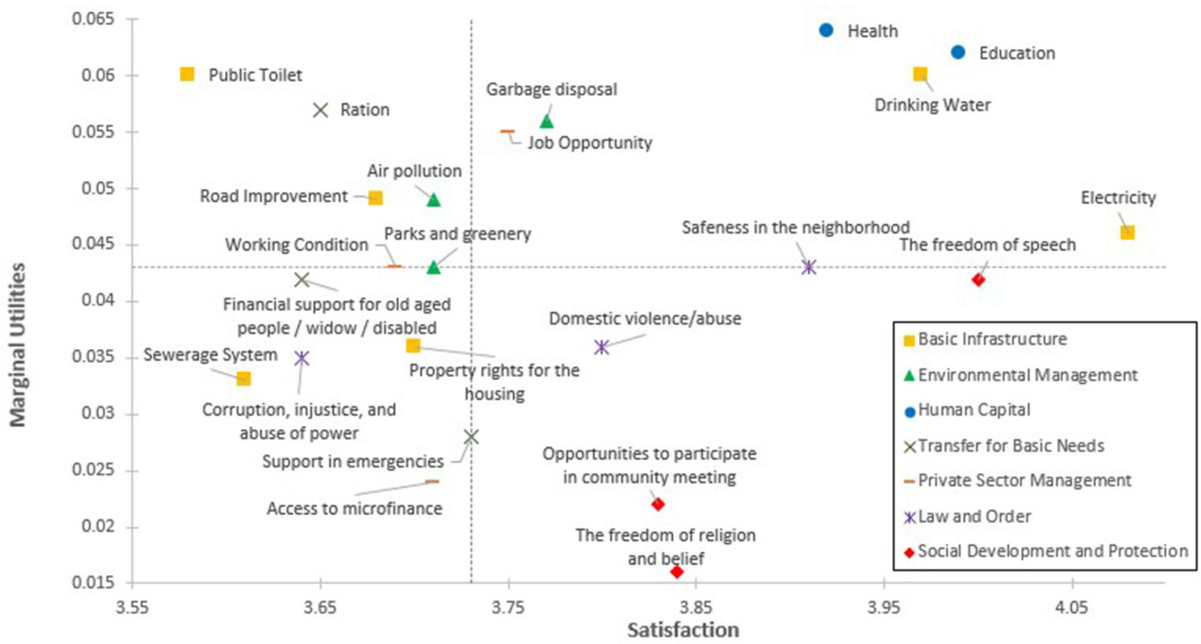


FIGURE 1
Marginal utilities and satisfaction with public projects.

third columns of Table 7 reporting these results, respectively. Although the rankings are generally similar for most domain, there are some variations among them. First, although the projects on health, education, drinking water, and public toilet, are highly ranked for both male and female respondents, female respondents put relatively higher weights on drinking water and public toilet. Moreover, a project on sewage system ranks higher among female respondents than male. These projects are closely related with a hygiene situation, implying that the projects about hygiene could be more appealing to female than male. Second, as for the project that directly influence their household budget, male respondents care more about job opportunities and working condition, and female respondents care more about ration, reflecting that men are the main wage earners. Lastly, other than these points, male respondents place more importance on freedom of speech and that female respondents place more importance on safeness in the neighborhood.

To see whether these rankings differ depending on slum residents' living conditions, we estimated marginal utilities by asset quantile, the results of which are reported in Table 8. Overall, the ranking results show a similar tendency across asset quantiles, but there are differences in some domains of development. For those projects on human capital, both health and education are the top-ranking problems for all asset quantiles, but the ranking of education is slightly lower among those in the poorest quantile (8^{th}) and in the richest quantile (10^{th}), despite the lowest average years of school and literacy rate of those in the poorest quantile and despite the highest average years of school and literacy rate of those in the richest quantile (Table 3). One possible explanation for this is that the importance of education may not be well recognized for the poorest residents and that the richest residents can send their children to private school instead of public school. Regarding basic infrastructure, as the asset quantile increases, the ranking of drinking water increases, while that of public toilets decreases. Another remarkable result is that although the ranking of the sewage system is low (19^{th}) for the whole sample, the ranking climbs up from 22^{nd} to 13^{th} as the asset quantile increases. For environmental management projects, garbage disposal ranks 1^{st} among those respondents in the poorest quantile. Conversely, the improvement of air pollution is ranked high among those respondents in the richer asset quantile. For private sector management, the improvement of job opportunities is anticipated to be ranked high among those in poorer quantiles, but the ranking is low (15th) among those in the richest quantile, as they wish to improve their working conditions (3rd). Overall, expectations for the role of quality of life increase as the asset quantile increases, although the needs for development projects are similar across asset quantile.

TABLE 8 Marginal utility estimates by asset quantile.

		By asset quantile									
		1		2		3		4		5	
	Variable	Rank	Coef.	Rank	Coef.	Rank	Coef.	Rank	Coef.	Rank	Coef.
1	Health	4	0.057	1	0.065	2	0.070	3	0.062	2	0.071
2	Education	8	0.051	5	0.051	1	0.076	1	0.069	10	0.054
3	Drinking water	7	0.052	9	0.047	3	0.066	4	0.061	1	0.079
4	Public toilet	2	0.060	2	0.061	5	0.061	5	0.059	8	0.056
5	Ration	9	0.050	7	0.050	8	0.051	2	0.068	5	0.063
6	Garbage disposal	1	0.064	3	0.052	6	0.054	8	0.053	6	0.060
7	Job opportunities	3	0.058	3	0.052	4	0.062	6	0.056	15	0.033
8	Road improvement	4	0.057	15	0.042	11	0.044	9	0.050	7	0.057
9	Air pollution	13	0.043	5	0.051	13	0.039	7	0.054	4	0.065
10	Electricity	6	0.055	16	0.038	10	0.045	12	0.042	9	0.055
11	Parks and greenery	13	0.043	12	0.043	6	0.054	13	0.040	18	(0.023)
12	Working conditions	12	0.046	8	0.048	15	0.037	18	0.034	3	0.066
13	Safeness in the neighborhood	15	0.039	12	0.043	12	0.043	10	0.046	12	0.041
14	Financial support for elderly, widowed, and disabled people	9	0.050	10	0.045	13	0.039	13	0.040	17	0.030
15	Freedom of speech	17	0.036	10	0.045	9	0.046	17	0.035	10	0.054
16	Property rights for housing	16	0.038	18	0.035	16	0.036	15	0.037	16	0.032
17	Domestic violence/abuse	9	0.050	17	0.037	17	0.034	20	0.029	14	0.034
18	Corruption, injustice, and abuse of power	17	0.036	12	0.043	18	0.033	15	0.037	19	(0.022)
19	Sewage system	22	0.020	21	0.028	19	0.031	11	0.044	13	0.040
20	Support in emergencies (flood, fires, and so on)	20	0.026	19	0.034	21	0.025	19	0.030	22	(0.019)
21	Access to microfinance	19	0.027	20	0.032	20	0.029	22	0.015	21	(0.021)
22	Opportunities to participate in community meetings	20	0.026	21	0.028	22	0.021	23	0.015	20	(0.022)
23	Freedom of religion and beliefs	23	0.017	23	0.027	23	(0.005)	21	0.024	23	(0.001)
	Observations	13,740		10,960		12,320		12,360		12,360	

The coefficients in parentheses are not significant at the 10% level.

Conclusions

This paper proposes a framework for weighting priority on multidimensional domains of slum development from the viewpoint of residents by extending the approach presented by Benjamin et al. (2014). We then demonstrate this approach by accessing residents' needs for slum development in Mumbai, India. We identify 23 domains of slum development that are relevant to the slum dwellers in Mumbai and estimate residents' marginal utilities obtained from the improvement of each domain of slum development. We show that (1) slum residents most wish to improve development projects about human capital (health and education), electricity, and drinking water, even if they have already been highly satisfied, (2) they feel the least satisfaction with public toilets and place high priority on projects involving public toilets, (3) the sewage system has low priority with low satisfaction, but this priority increases as slum residents become better off, (4) the projects on social development and protection are highly satisfied and generally ranked low in terms of marginal utilities, and (5) air pollution and working conditions are also concerns of slum residents, especially as these residents become better off. Based on these results on priorities across different area of development projects, administration and NGO's need to put forward individual projects in each area. Especially, projects to reduce the disease burden caused by poor sanitation and unsafe drinking water are found to be urgent issues. Longer and stable water supply, distribution of chlorine tablets to residents to disinfect drinking water, providing information on the importance of washing hands after using toilets, and people's participation in the management of public toilets, for example, are all to move into action immediately.

Using marginal utilities derived from this approach as weights, the weighted average of satisfaction scores in each domain of development can be interpreted as residents' utility from slum development. Once the marginal utilities are estimated, this composite index of weighted averages can be used to trace the progress of overall development plans.

Our results suggest two implications in the measurement of development progress. First, in many cases, simple averages are used as a composite index of development (e.g., Decancq and Lugo, 2013). However, we find some discrepancy of the evaluation between marginal utilities and satisfaction scores. This implies that only the satisfaction scores are insufficient for the understanding of overall subjective wellbeing of residents' welfare. The information on marginal utilities of each domain, as well as satisfaction scores, should be collected. Second, this paper suggests to conduct regular surveys that ask about satisfaction with each domain of slum developments in order to trace the progress of overall development plans. In addition to this regular survey, this paper recommends to conduct a periodic reassessment of marginal utilities, because our results suggest that, although the needs for development projects are similar across asset quantile to some extent, expectations for the role of quality of life increase as the asset quantile increases.

Of course, like other approaches, this approach has several weaknesses that should be noted. First, as discussed in Section Weighting approach for the multidimensional indicator, the composite index represents a first-order approximation of utility derived from the Taylor expansion and thus becomes less accurate if residents' perceived satisfaction from slum development deviates largely from the status quo. Therefore, the use of this index should be limited to when evaluating small changes in slum development. Second, although the proposed approach has a strong advantage in evaluating the relative importance across different domains of slum development, it does not provide any information on each individual domain. Thus, once the domain to be prioritized is identified, a deeper assessment of each individual domain should be conducted.

Nevertheless, we believe that these weaknesses are outweighed by the merit of the resident-oriented nature of this approach. As Naess (2001) argues, planning for a sustainable development should make use of both expert and layman knowledge. Experimental knowledge of local residents about local environmental quality could compensate for expert knowledge, and on the other hand, residents' value priorities on development projects might be required to change for sustainable development. However, in many cases, slum residents are not legally notified by the government, and even if they are notified, their needs hardly ever reach the government. Different from other data-driven approaches or from normative approaches that judge priorities by the analyst or other third parties, such as community leaders or politicians, this approach directly asks slum residents about their preferences in the domains of slum development. In this sense, this approach gives them a voice with which to disclose their actual needs for slum development, and is expected to promote dialogue between urban planners and slum residents.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

AK designed and executed the survey and wrote the manuscript. SM contributed to the concept and helped to write the manuscript. Both authors read and approved the final manuscript.

Funding

This research was funded by the Grant-in-Aid for Scientific Research 26000001, 19K15919, and 20H00648 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The results and conclusions of this article do not necessarily represent the views of the funding agencies.

References

Benjamin, D., Heffets, O., Kimball, M., and Szembrot, N. (2014). Beyond happiness and satisfaction: Toward well-being indices based on stated preference. *Am. Econ. Rev.* 104, 2698–2735. doi: 10.1257/aer.104.9.2698

Biswas, R., Arya, K., and Deshpande, S. (2020). More toilet infrastructures do not open defecation: a perspective from squatter settlements in megacity Mumba. *Appl. Water Sci.* 10, 96. doi: 10.1007/s13201-020-1169-4

Bossert, W., Chakravarty, S., and D' Ambrosio, C. (2013). Multidimensional poverty and material deprivation. *Rev. Income Wealth.* 59, 29–43. doi: 10.1111/j.1475-4991.2012.00519.x

Census of India. (2001). *Metadata and brief highlights on slum population*. Available online at: https://censusindia.gov.in/Data_Products/Data_Highlights/Data_Highlights_link/metadata_highlights.pdf (accessed 11 March, 2022)

Census. (2011). *Mumbai (Greater Mubai) City Population 2011–2022*. Available online at: https://www.census2011.co.in/census/city/365-mumbai.html#:~:text=Mumbai%20Slums%202011,total%20population%20of%20Mumbai%20city (accessed 12 March, 2022)

Chowdhury, S., and Squire, L. (2006). Setting weights for aggregate indices: an application to the commitment to development index and human development index. *J. Develop. Stud.* 42, 761–771. doi: 10.1080/00220380600741904

De Kruijk, H., and Rutten, M. (2007). "Weighting dimensions of poverty based on people's priorities: constructing a composite poverty index for the Maldives," in *Q-Squared Working Paper,* Vol. 35 (Centre for International Studies, University of Toronto).

Decancq, K., and Lugo, M. A. (2013). Weights in multidimensional indices of wellbeing: an overview. *Econ. Rev.* 32, 7–34. doi: 10.1080/07474938.2012.690641

Diener, E., and Suh, E. (1997). Measuring quality of life: economic, social and subjective indicators. *Social Indicat. Res.* 40, 189–216. doi: 10.1023/A:1006859511756

Ferreira, S., and Moro, M. (2010). On the use of subjective well-being data for environmental valuation. *Environ. Resource Econ.* 46, 249–273. doi: 10.1007/s10640-009-9339-8

Klasen, S. (2000). Masuring povety and deprivation in South Africa. *Rev. Income Wealth* 46, 33–58. doi: 10.1111/j.1475-4991.2000.tb00390.x

Krishnakumar, J., and Nadar, A. (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. *Social Indicators Res.* 86, 481–496. doi: 10.1007/s11205-007-9181-8

Levinson, A. (2012). Valuing public goods using happiness data: the case of air quality. *J. Public Econ.* 96, 869–880. doi: 10.1016/j.jpubeco.2012.06.007

Lugo, M. A. (2007). Comparing multidimensional indices of inequality: methods and application. *Res. Econ. Inequality* 14, 213–236. doi: 10.1016/S1049-2585(06)14010-7

Mascherini, M., and Hoskins, B. (2008). "*Retrieving Expert Opinion on Weights for the Active Citizenship Composite Indicator*," in, European Comission, Joint Research Center, Institute for the Protection and Security of the Citizen.

Naess, P. (2001). Urban planning and sustainable development. *Eur. Plan. Stud.* 9, 503–524. doi: 10.1080/09654310120049871

Nilsson, T. (2010). Health, wealth and wisdom: exploring multidimensional inequality in a developing country. *Social Indicators Res.* 95, 299–323. doi: 10.1007/s11205-009-9461-6

Noorbakhsh, F. (1998). The human development index: some technical issues and alternative indices. *J. Int. Develop.* 10, 589–605. doi: 10.1002/(SICI)1099-1328(199807/08)10:5<589::AID-JID484>3.0.CO;2-S

Schokkaert, E. (2007). Capabilities and satisfaction with life. *J. Hum. Develop. 8,* 415–430. doi: 10.1080/14649880701462239

Stiglitz, J. E., Fitoussi, J. P., and Durandm, M. (2018). *Beyond GDP Measuring What Counts for Economic and Social Performance.* Paris: OECD Publishing. doi: 10.1787/9789264307292-en

Subbaraman, R., Shitole, S., Shitole, T., Sawant, K., O'Brien, J., Bloom, D. E., et al. (2013). The social ecology of water in a Mumbai slum: failures in water quality, quantity, and reliability. *BMC Public Health* 13, 173. doi: 10.1186/1471-2458-13-173

UN-Habitat. (2020). *World Cities Report 2020*. The Value of Sustainable Urbanization. United Nations Human Settlement Programme (UN-Habitat), Nairobi.

United Nations. (2015). *Transforming Our World: the 2030 Agenda for Sustainable Development*. Available online at: https://sdgs.un.org/2030agenda (accessed 11 March, 2022)

Welsch, H. (2006). Environment and happiness: valuation of air pollution using life satisfaction data. *Ecol. Econ.* 58, 801–813. doi: 10.1016/j.ecolecon.2005.09.006

Yesudhas, R. (2021). *Access to Water, Sanitation and Hygiene in Slumes of Mumbai*. Available online at: https://www.researchgate.net/profile/Ronald-Yesudhas/publication/349493721_Access_to_Water_Sanitation_and_Hygiene_in_Slums_of_Mumbai/links/60336dc54585158939beb3ed/Access-to-Water-Sanitation-and-Hygiene-in-Slums-of-Mumbai.pdf (accessed 7 July, 2022).

frontiers | Frontiers in Sustainability

TYPE Original Research
PUBLISHED 31 August 2022
DOI 10.3389/frsus.2022.908822

Check for updates

OPEN ACCESS

EDITED BY
Alexander Ryota Keeley,
Kyushu University, Japan

REVIEWED BY
Changyong Cao,
National Oceanic and Atmospheric Administration (NOAA), United States
Meiqin Che,
Nantong University, China

*CORRESPONDENCE
Go Kazawa
kazawa.gou.62x@st.kyoto-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 31 March 2022
ACCEPTED 03 August 2022
PUBLISHED 31 August 2022

CITATION
Kazawa G, Seki D, Keola S, Iwasaki F and Yamashiki YA (2022) Possible correlation between nighttime lighting data and building height. *Front. Sustain.* 3:908822. doi: 10.3389/frsus.2022.908822

Possible correlation between nighttime lighting data and building height

Go Kazawa[1]*, Daikichi Seki[1], Souknilanh Keola[2], Fusanori Iwasaki[3] and Yosuke Alexandre Yamashiki[1]

[1]Graduate School of Advanced Integrated Studies in Human Survivability (Shishukan), Kyoto University, Kyoto, Japan, [2]Institute of Developing Economies, Japan External Trade Organization, Chiba, Japan, [3]Economic Research Institute for ASEAN and East Asia (ERIA), Jakarta, Indonesia

This study examined whether nighttime light can be a proxy for building height. It is important for estimating an urban environment and urban design and understanding industrial clusters. However, it is necessary to purchase height data for a wide range of buildings, and it is difficult to obtain data in a time series. In this study, we examined whether it is possible to estimate the height of buildings using nighttime light data. This research focuses on the fact that as the number of windows increases, the amount of light leaking from the windows during the night increases for the entire building. This increases the amount of light emitted by the building. We conducted the first Japanese correlation analysis using a three-dimensional dataset with a resolution of 1 meter that was released by Hyogo Prefecture on 10 January 2020. We also used nighttime light data from the Suomi NPP satellite, which has been in operation since 2011 and is jointly managed by National Oceanic and Atmospheric Administration and NASA. The percentage of land occupied by buildings within a 450-meter square of a nighttime light resolution was varied to obtain a higher correlation between buildings and nighttime light. The correlation between the average height of buildings and the luminous intensity of nighttime light within that pixel was calculated. The coefficient of determination was 0.425, which was the best value when the percentage of land occupied by buildings in a pixel was between 70 and 80%. This study found a high correlation between night light and building height. We believe that if a building has a certain size, the night light can be used as a proxy. The results of our research will contribute to understanding the changes and characteristics of urban development and population distribution as urbanization progresses in various regions. Rapid public transportation services, education, and environmental improvement will contribute to urban development and changes in population distribution, which will greatly relieve urban poverty and improve urban security and the environment.

KEYWORDS

night light data, building height, satellite image analyses, GIS, urban environment

Introduction

Research background

Estimating economic activities or Gross Domestic Products (GDPs) is one of the critical challenges to promote the research of economics and other fields of social sciences (James, 2020). However, we face various obstacles such as lack of data or data accessibilities and improving statistical methods (Henderson and Adam, 2012). In line with the research trends, the authors aimed to examine the relationship between the nighttime light and building height to understand the amounts of economic activities in urban areas by using the case of Japan.

In Japanese urban planning, floor area ratios are set for each building site in order to control industrial concentration and population density. The floor area ratio is the total floor area ratio to the site area. In Japan, floor area ratios are set by local governments according to the characteristics of each region. Local governments set floor area ratios according to the characteristics of each region. In central areas, where people and businesses tend to congregate, floor area ratio regulations are more relaxed, making it easier to construct relatively large buildings [see the website of the Ministry of Internal Affairs Communications (2021)]. In this way, the development of a city is controlled by its floor area ratio, which is closely related to its industrial concentration and population density.

When planning urban facilities, such as transportation, parks and green spaces, and educational and cultural facilities, it is important to understand the changes and characteristics of urban development and population distribution changes and characteristics. If public transportation services, education, and environmental development lag behind urban development and changes in population distribution, there may be serious problems in terms of poverty, security, and the environment. We believe it is important to understand the differences and characteristics of urban development and population distribution to reduce these risks.

The total floor area of buildings and the building height are critical factors in understanding the characteristics of a city. However, the data on gross floor area and building height needed to determine floor area ratios are not available every year. The only way to obtain accurate data over a wide area is to purchase it. The data collection method is another reason for the high cost of data. Building height data is measured by aerial laser surveying. The measurements were taken by airplane over a large area over about two years, so it took a lot of time and money to collect the data, which increased the cost of the data.

On the other hand, the amount of satellite data downloaded free of charge has increased significantly over the past decade. The accuracy of earth observation using satellites has been improving in recent years, and the scope of open data has been expanding. In addition, geographic information systems (GIS) is a technology that enables advanced analysis and quick decisions on information indicating the location of specific points or areas in space (positional information). Information on various events associated with the location-based on computer processing has been dramatically developed. The combination of these technologies has made it possible to understand and predict the spatial structure of various regions [see the website of the Geospatial Information Authority of Japan (2021)].

This study examined the relationship between nighttime light and building height. Suppose nighttime light data can be applied as a proxy variable for building height. Building height data—which is difficult to obtain in a time series—can be referred to free of charge. The characteristics of urban industrial clusters and population distribution can be grasped on a finer timescale, contributing to urban design and planning.

Review of previous studies

Henderson and Adam (2012) used the World Bank's global GDP growth data for 1992–1993 and 2004–2005 and found a correlation between changes in nighttime light intensity and GDP growth. The coefficient of determination is 0.288, which is a scatter plot of 170 countries worldwide with the difference in GDP without processing on the Y-axis and the logarithm of the difference in night light on the X-axis. This research approaches the three main problems that have been identified in developing countries: the existence of unofficial transactions that are not reflected in GDP, the lack of infrastructure for collecting statistics, and the existence of countries whose GDP is not published. Research is still being actively conducted. Since the publication of this paper as a working paper in 2009, the value of using nocturnal satellite imagery has been fully recognized by the economist community, and its use has expanded rapidly since then (Mellander et al., 2015).

On the other hand, it is difficult to estimate GDP in developing countries where the main industry is rural because nighttime light cannot be measured (Nordhaus and Chen, 2012). Previous studies have proven that nighttime light has a certain effect on improving economic production statistics at the national level in certain countries, but little research has been conducted at the local level within a country (Keola and Andersson, 2015). In particular, it has been challenging to estimate the economic activity of people living in or near the international poverty line ($1.90 per person per day). Since the level of luminous intensity in these areas was very low and varied very little, the night light was not effective in estimating the economic activity in these areas. At the time, a method of using cell phone data to infer poverty was used to solve this problem by Blumenstock et al. (2015).

Blumenstock et al. (2015) shows that they used an anonymous database containing records of billions of interactions on Rwanda's largest cell phone network to infer and map the socioeconomic status of individuals based on their past cell phone usage history. They also randomly selected 856 people from the database and conducted a telephone questionnaire about their assets and housing characteristics. Using cell phone metadata, they were able to reconstruct the distribution of wealth for the entire country accurately and infer the distribution of assets in micro-regions consisting of only a few households (Blumenstock et al., 2015). This approach was a very effective way to visualize the distribution of wealth in countries with limited information resources, where census and household surveys are rarely conducted.

On the other hand, extending this research beyond a country is very difficult because it relies on a unique data set, with data bias depending on the cell phone company that provides the data.

Therefore, recent studies have used deep learning techniques to infer economic activity in rural districts and other districts where it is challenging to extract production at night light. Jean et al. (2016) used survey data from five African countries (Nigeria, Tanzania, Uganda, Malawi, and Rwanda) and satellite data were used to train convolutional neural networks to explain up to 75% of the economic variation at the regional level. Jean et al. (2016) obtained characteristics of poor areas and made economic forecasts by overlaying daylight and nightlight satellite images of areas that appear almost black with no light when viewed in nightlight. We extracted objects such as houses and rivers from the daytime satellite images using deep learning. We found that these features, such as the material of the roof and the distance to the urban area, were proportional to the expenditure of the village and inferred economic activity from them.

In general, deep learning models such as convolutional neural networks can, in principle, be trained to estimate economic outcomes from satellite images directly. Still, when it comes to poor regions in Africa, the lack of these training data makes it difficult to apply this method (Christian, 2019). However, Jean et al. (2016) used a method called transfer learning, which is capable of inferring even on a small training data set, allowing for the extraction of village objects.

Thus, research has been conducted on the use of night light to infer GDP from night light (Chen and Nordhaus, 2019). However, previous studies have shown that it is difficult to estimate GDP from nightlight data alone since GDP is estimated based on primary statistics covering as many economic transactions as possible. Therefore, in developing countries with many data gaps in the primary statistics, GDP as a teacher data, has low credibility and needs to be substituted with other figures (Nordhaus and Chen, 2012). There are many regions where no data exists when it comes to GDP. Gaps in teacher data can be an obstacle to numerical estimation. On the other hand, for height data, the accuracy of the teacher data is high, and the sample size is large.

Barr et al. (2015) is the first paper to examine how skyscraper height and output co-move rigorously. Here he explains that building height can be a leading indicator of GDP, while at the same time explaining that height does not cause production (Barr et al., 2015). Therefore, instead of using nighttime lighting directly to estimate GDP, understanding industrial structure from building height could contribute to closing the GDP data gap. Therefore, this study proposes a proxy variable for building height as a new use value for nighttime lighting.

Research purpose

This study aims to conduct a correlation analysis with night light data using building height data based on prefecture-wide 3D data with 1-m mesh accuracy, which Hyogo Prefecture in Japan released for the first time in Japan on January 10, 2020. Hyogo Prefecture is located in the center of Japan. It has more than five million populations, including Kobe city (about 1.5 million people), one of the largest cities in Japan according to the statistics[1]. Therefore, the data of Hyogo prefecture is suitable for our research purpose to examine the urban area.

The hypothesis of this study is that night light can be used as a proxy for building height as a new use night light. This is inspired by the fact that the number of windows in a building increases in proportion to the height of the building. As the number of windows increases, the amount of light that leaks out of the windows at night increases for the establishment, and the amount of light emitted increases. This research is focused on this point. In this study, we will examine whether a positive correlation can be found between nighttime light and the height of a building.

Data

e-Stat (https://www.e-stat.go.jp/)

The General Contact Point for Government Statistics (e-Stat) is a portal site for government statistics in Japan. It has been in operation since 2008 and was created by the Liaison Conference of Chief Information Officers of Ministries and Agencies. It aims to provide a one-stop service for information on government statistics in Japan [see the websites of e-Stat (2021) "About this Site"].

This site consolidates statistics-related information that has been scattered across independent websites operated by ministries and agencies. The site provides various statistics registered by each ministry and agency, such as statistical table files, statistical data, publication schedules, new information, survey item information, and statistical classifications [see the websites of e-Stat (2021) "About this Site"]. Its purpose

1 See the website of Hyogo Prefecture, http://web.pref.hyogo.lg.jp/kk11/jinkou-tochitoukei/jinnkounouno.html (access 30 march 2022).

is to provide statistical results, which are the information infrastructure of society, in a form that is easily accessible to all.

In this study, we used the boundary data of Hyogo Prefecture from e-Stat. The boundary data is used to cut the satellite data and GIS data such as Night-Time Light Data and building information by coordinate mesh. The boundary data is the data for the whole Hyogo Prefecture in the 2014 national census and subregions (town, district, character, etc.). The data format is a shapefile in the world geodetic system rectangular plane coordinate system.

It is used in large scale maps, such as "1/2,500 National Land Map," "1/5,000 National Land Map," and "1/10,000 Topographic Map" published by the Geographical Survey Institute. The projection method is the Gauss-Kruger method (also called "horizontal Mercator projection"). In order to reduce the distortion caused by projecting an ellipsoid surface onto a plane, the whole country of Japan is divided into 19 regions, each of which has its coordinate origin.

In order to reduce the distortion caused by projecting an ellipsoid surface onto a plane, the whole country is divided into 19 regions, each of which has its own coordinate origin. One-hundred-thirty km east and west from the coordinate origin is the applicable range of each coordinate system [see the website of esri Japan., 2021, "Coordinate Systems Used in Japan"].

National land data information (https://nlftp.mlit.go.jp/ksj/)

The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan has developed numerical national land information to support the formulation and implementation of national land planning. The ministry provides a variety of information on national land, including administrative areas, railroads, roads, rivers, official land price announcements, land use mesh, and public facilities [see the website of MLIT Ministry of Land Infrastructure Transport Tourism (2021a), GIS Home Page].

This study used the 2014 Hyogo Prefecture data. To investigate the light intensity of building sites in Hyogo Prefecture, we overlaid the building site data on top of the nighttime lighting data.

Fundamental geospatial information (https://www.gsi.go.jp/kiban/)

Fundamental Geospatial Information has been developed by the Geospatial Information Authority of Japan since the enactment of the Basic Act on the Utilization of Geospatial Information in 2007[2]. At present, the initial development of base map information is almost complete, with about 100,000 square kilometers of urban planning areas at a scale level of 2,500 and other areas at a scale level of 25,000.

Basic items of fundamental geospatial data

In order to extract the area of buildings and location information by coordinates, all the data of Hyogo Prefecture was downloaded and used. The downloaded file format was JPGIS (GML) format.

The JPGIS format, officially called the "Japan Profile for Geographic Information Standards," is a set of data formats developed in accordance with international and national standards. The main purpose is to ensure the compatibility of the GIS spatial data for mutual use among different systems. It is a practical standard that specifies rules for data design, quality, description methods, and specification writing methods. These are international and national standards.

Therefore, by promoting the use of JPGIS, an environment that facilitates the mutual use of data can be established. The effects of sharing data maintained by different maintenance entities, reducing system dependency, and eliminating duplication of investment can be expected. This will enable cost reduction and business efficiency and will contribute to the formation of a more efficient information society and the realization of a safe and secure society (see the website of GSI 2 (2021), "What can be done with the practical use of geospatial information (JPGIS)").

The basic items of Fundamental Geospatial Information are 10 items taken from the 13 items of Fundamental Geospatial Information. They are survey reference points, coastlines, boundaries and representative points of administrative divisions, road edges, track centerlines, elevation points, water lines, perimeter lines of buildings, boundaries and representative points of towns and cities, and boundaries and representative points of city blocks (Fundamental Geospatial Data Fundamental Geospatial Data Basic Items., 2021).

Building data

The extent of the buildings was determined from the building perimeter data published by the Geospatial Information Authority of Japan. The data was downloaded for the entire area of Hyogo Prefecture and was used to identify the actual locations of buildings. The data items are defined by the acquisition

2 See website, https://www.gsi.go.jp/common/000002047.pdf (access 30 march 2022).

standard used in the large scale topographic map format, and the buildings are indicated by the following categories.

Futsu-Tatemono

Buildings with less than three floors and buildings with three or more floors constructed of wood or other materials.

Kenrou-Tatemono

Buildings constructed of reinforced concrete, etc., more than three stories above ground or with a height equivalent of more than three stories.

Futsu-Muhekisha

Buildings without side walls, greenhouses, and similar buildings in factories with less than three floors; for example, plastic greenhouses.

Kenrou-Muhekisha

Buildings without side walls constructed of reinforced concrete or similar materials and having a height of more than three stories above ground or more than the equivalent of three stories (for example, self-propelled multistorey parking lots, etc.) (Ministry of Land Infrastructure Transport Tourism, 2021b).

G-space information center (https://www.geospatial.jp/gp_front/)

The G-Spatial Information Center is a platform built by the National Institute of Information and Communications Technology. the University of Tokyo and Hitachi, Ltd. have built a platform to freely combine and obtain G-Spatial information held by the public and private sectors in a one-stop service. It was established based on the Basic Plan for the Promotion of Geospatial Information Utilization approved by the government in March 2012[3]. The Council for Promotion of Social Infrastructure Information Distribution has been conducted since 2016.

The merits of the G-Spatial Information Center are abilities to register and centrally search for G-Spatial information held by industries, governments, and academia; the ability to convert registered G-Spatial information into a form that is easy for users to use and obtain; and the ability to obtain G-Spatial information by both downloading and API. Making the registered G-Spatial information available *via* API, may mean that G-Spatial information will be utilized in a wide range of fields [see the website of Japan Council for Promotion of Information Sharing on Infrastructures Operation of the "G.-Geospatial Information Center. (2021)].

Hyogo prefecturewide digital surface model

On January 10, 2020, Hyogo Prefecture released the first prefecture-wide 3D dataset with 1-m mesh accuracy in Japan. The Digital Surface Model (DSM) is also known as a numerical surface model, which is the numerical surface layer of aerial laser survey results that include the ground surface, buildings, and vegetation. It is the original aerial survey data.

The data is in the form of a grid with 1-m intervals, and it is text data with XYZ coordinate values. It was created using aerial laser survey data of a 1-m mesh produced by Hyogo Prefecture and the Rokko Erosion Control Office, Kinki Regional Development Bureau, and the Ministry of Land, Infrastructure, and Transport of Japan. The coordinate system is the Japan Geodetic System 2000.

Hyogo-wide digital height model

This is the height data of the ground above the ground surface, which is created by the difference between the DSM and the Digital Elevation Model (DEM). This includes everything above the ground surface, including buildings and trees and other plants. The observation period was from 2012 to 2013.

National oceanic and atmospheric administration (NOAA)/NASA Suomi NPP satellite day/night band (https://eogdata.mines.edu/download_dnb_composites.html)

NASA and the NOAA launched the Suomi NPP Earth observation satellite in November 2011, equipped with the Visible and Infrared Imager and Radiometer (VIIRS). Most satellites have a daytime light source, and most satellites observe only during the daytime, but Suomi NPP also performs observations at night, and the data used in this study is nighttime data.

The monthly data has not been filtered to remove auroras, fires, boats, and other transient lights. Also, the resolution is a 15-arcsecond grid (about 450 m) ranging from −180 to 180 degrees, covering latitudes 65S to 75N, and is available in geotiff format [Version 1 VIIRS Day/Night Band Nighttime Lights].

This paper uses annual data for the year 2016. In this study, we used the data from the vcm-orm-ntl file, where the nighttime light outliers are removed and the non-light portion is set to

3 See website, https://www.gsi.go.jp/common/000195830.pdf (Last view on 30 March 2022).

zero. The values in this data are proportional to the actual luminous intensity.

Research methods

Creating a correlation between building height and nighttime light in hyogo prefecture

Using the VIIRS annual composite, DSM, and DEM data, we calculated the height (DSM–DEM) in each night light (NL) pixel. The coordinates are standardized to ESPG 2447 (JGD2000, Japan Plane Rectangular CS V). The mean, standard deviation, maximum, minimum, and number of samples for the height of each building per NL pixel were calculated using Fundamental Geospatial Data. Additionally, building heights were derived for *Futsu-Tatemono* and *Kenrou-Tatemono*.

Then, the NL pixels overlapped the land use meshes that occupied more than 50%, more than 60%, more than 70%, more than 80%, and more than 90% of the total building land area and were analyzed, respectively and compared with Google map to investigate the characteristics of the land. Then, we accepted the best-fit building land ratio as the result of this study.

Spatial merging and feature extraction using geopandas 0.8.0

GeoPandas is an open-source project that enables geospatial data manipulation in Python. GeoPandas extends the data types used to enable geospatial manipulation with Pandas. In this paper, we used the Digital Height Model for the entire Hyogo area (2010–2018), masked with the building perimeter frame obtained from the base map information, to extract the nighttime light allocated to it. The data from the SUOMI NPP was used to extract the NL allocated to each of the 23 wards of Tokyo and to extract how much nighttime light was allocated to each building in these 23 wards.

The data in the SUOMI NPP is "created in a 15-second geographic grid [Version 1 VIIRS Day/Night Band Nighttime Lights "README" file]." Note that 1 second angle = $1^{\circ}/3{,}600.0$. Therefore, the distance between the axes of the longitude and latitude lines is expressed as ddeg = 15.0/3,600.0. Using this, prepare longitude lines at intervals of 15.0/3,600.0 for 60° to 180°, which is the range of the NL data, and latitude lines at intervals of 15.0/3,600.0 for 0° to 75°. Cut the NL data at the pixel that you want to study.

Analysis results

The elevation data of the ground surface and the overlying land surface were obtained with a spatial resolution of 1 m by subtracting the DSM and DEM data published by Hyogo Prefecture during the observation period 2012–2013. Data with either DSM or DEM missing were not used. The data were spatially combined with the perimeter frame data of ordinary and solid buildings in Hyogo Prefecture in 2014. It was extracted from the base map information to obtain the data for buildings only.

Because the spatial resolution of the NL data is 460 × 460 m, we used this unit as 1 pixel. By spatially combining the building data, we derived the number of buildings contained in each pixel and extracted the amount of light from each building in each pixel. Subsequently, when we checked the data on the map, we found that forests and gardens were included in the building data.

Therefore, the building site information was extracted from the land use information, and by spatially combining this information, only the buildings were extracted from the part where the site was selected. After that, the average height of the buildings in the pixel was derived and became the Y-axis measurements. The annual composite value of the number of observations for each month weighted by the number of observations for each pixel in NL VIIRS 2013 was used as the X-axis.

In order to remove data where forests and fields occupied most of the pixel, we restricted the data by the percentage of land used for buildings in the pixel so that we could more accurately analyze the correlation between building height and NL. To find the best correlation between NL and building heights, the percentage of building sites in the pixel was changed to 50% or more, 60% or more, 70% or more, 80% or more, and 90% or more, respectively, and the correlation analysis was conducted.

The average height of the buildings in the pixel was derived and placed on the Y-axis, and the annual composite value was placed on the X-axis, weighted by the number of observations in each month for each pixel in NL VIIRS 2013. Figure 2 shows the plots of the pixels with more than 50% of building sites. The data have a linear distribution, but there are some outliers where the average height of the buildings is more than 25 m. It is also clear that the data is mainly concentrated in the area where the average height is $<$15 meters.

Figure 1 shows 1,769 observations; the slope is 0.126, the intercept is 5.3500, and R-squared is 0.335. The F-stat is sufficiently large, and the prob (F-stat) is close to 0, indicating that the coefficient of the explanatory variable is unlikely to be 0, and the regression equation is meaningful. Both t-statistics are more significant than 2, which means that both coefficients are significant.

Next, we looked at the plots with more than 60% of the building sites in the pixel (Figure 2). They were scattered

10.3389/frsus.2022.908822

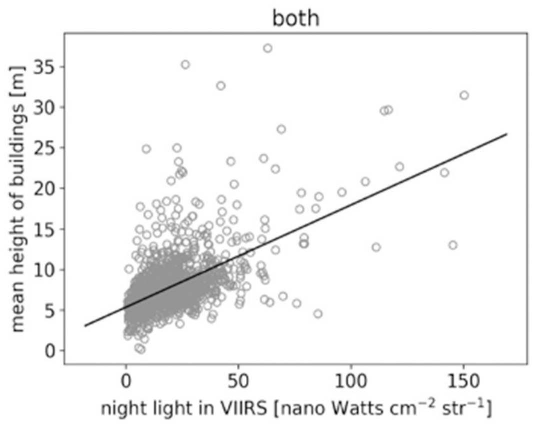


FIGURE 1
Scatter plot of the night light and building height for land use ratio of more than 50%.

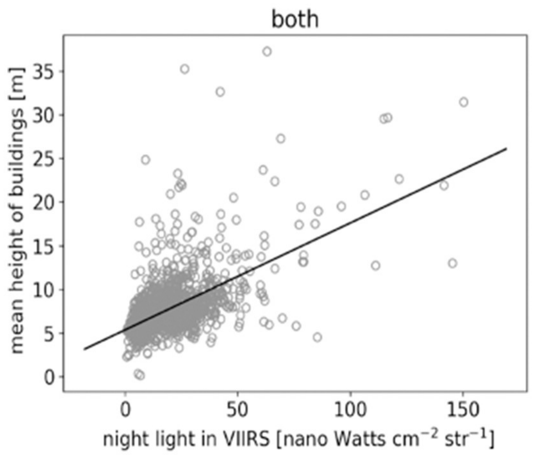


FIGURE 3
Scatter plot of the night light and building height for land use ratio of more than 70%.

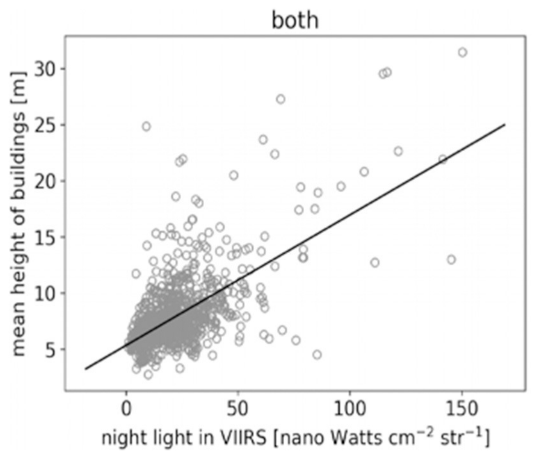


FIGURE 2
Scatter plot of the night light and building height for land use ratio of 60% and above.

linearly, but there were some outliers where the average height of the buildings was more than 25 meters. We also found that the data was mainly concentrated in the area where the average height was below 15 meters. There were 1,470 observations; the slope was 0.1225, the intercept was 5.3791, and R-squared was 0.331. The F-statistic was sufficiently large, and the prob (F-statistic) was close to 0, so it was very unlikely that the coefficient of the explanatory variable was 0, indicating that the regression equation was meaningful. Both t-statistics are >2, which means that both coefficients are significant.

The number of observations decreased, and R-squared also decreased slightly when compared with more than 50% of building sites in the pixel. There is not much difference in the data between the case where the land within the pixel is more than 50% and the case where the land within the pixel is more than 60%.

Here is a plot with more than 70% of the building sites in pixels (Figure 3). Similarly, the data is linearly distributed. We also found that the data is mainly concentrated in areas where the average height is <15 meters. The number of observations was 1,028, the slope was 0.1161, the intercept was 5.3668, and the R-square was 0.392.

The F statistic was sufficiently large, and the prob (F statistic) was close to zero, indicating that the coefficients of the explanatory variables were very unlikely to be zero and that the regression equation was meaningful. Both t-statistics were more outstanding than two, and the coefficients were significant. Buildings with >50 or 60% land use had the highest Pixel R-squared values, although the number of data was reduced.

We plot pixels with >80% of building sites (Figure 4). Overall, the distribution is linear. We also see that the data is mainly concentrated in areas with an average height of 15 m or less. The number of observations was 595, the slope was 0.0920, the intercept was 5.5951, and the R-square was 0.290. The F statistic is sufficiently large, and the prob (F statistic) is close to 0, indicating that the coefficient of the explanatory variable is very unlikely to be 0 and the regression equation is meaningful. The t statistics are all above two, and both coefficients are significant.

Comparing the number of building lots in pixels with more than 50%, 60%, and 70% building lots, the observed values decrease, and the R-square decreases significantly; when data with more than 80% of buildings in a pixel is superimposed on a Google map, the data for Kobe City Hall disappears due to the increase in the percentage of residential lots. Data with high building height is critical to see the relationship with NL,

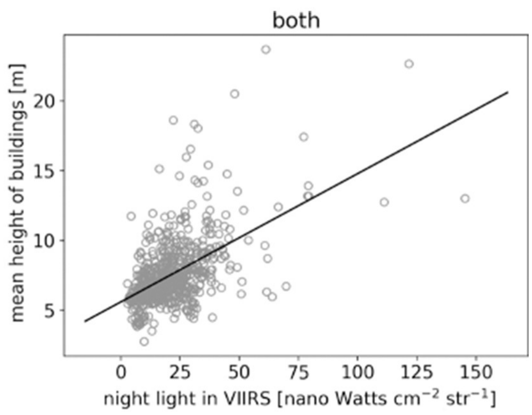


FIGURE 4
Scatter plot of the night light and building height for land use ratio of 80% and above.

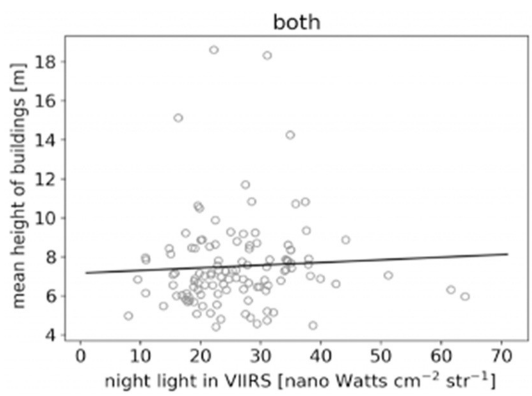


FIGURE 5
Scatter plot of the night light and building height for land use ratio of 90% and above.

and it is challenging to adapt the data with more than 80% of residential land.

We can see a plot with more than 90% of the buildings in pixels (Figure 5). Overall, the distribution is linear. We also found that the data is mainly concentrated in areas with an average height of 8 meters or less. The number of observations was 116, the slope was 0.0136, the intercept was 7.1641, and the R-square was 0.003.

The F-statistic is sufficiently large, and the prob (F-statistic) is close to 0, so it is very unlikely that the coefficient of the explanatory variable is 0, indicating that the regression equation is meaningful. Both t-statistics are larger than 2, and both coefficients are significant.

The number of observations and the R-squared decreased significantly when comparing the data of more than 50%, more than 60%, more than 70%, and more than 80% of building sites in a pixel. In the case of low-rise buildings such as those in

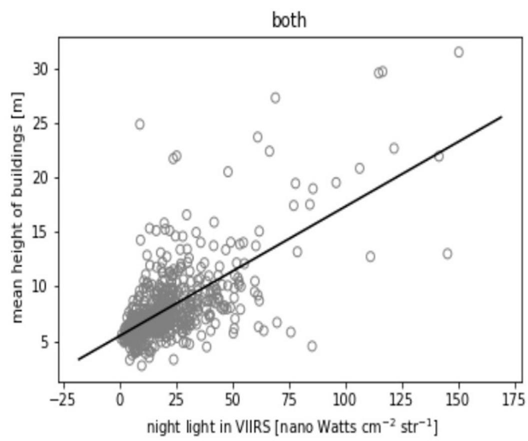


FIGURE 6
Scatter plot of the night light and building height for land use ratio between 70 and 80%.

residential areas, the correlation between building height and nighttime light is not significant.

Finally, we extracted and examined the percentage of building sites with a high relationship between building height and nighttime light, looking at the percentage of building sites in a pixel between 70 and 80% (Figure 6). The number of observations was 599, the slope was 0.1185, the intercept was 5.4652, and the R-squared was 0.425. The F-statistic was large enough, and the prob (F-statistic) was close to 0, so it was very unlikely that the coefficient of the explanatory variable was 0, indicating that the regression equation was meaningful. Both t-statistics were >2, which means that both coefficients are significant; R-squared was able to produce the highest results.

Discussion

The correlation between nighttime light and DCM data in Hyogo Prefecture was analyzed by dividing the percentage of land use within one nighttime light pixel into 50%, 60%, 70%, 80%, and 90%. The results show that the best correlation is obtained when the percentage of land use within one pixel is 70%.

As the land use ratio increased, the grid became more dominated by residential areas with many single-family homes. The best results were obtained for between 70 and 80%, except for building coverage ratios above 80%, where residential areas dominate and important landmarks are excluded.

This paper assumes that as building height increases, the number of windows increases, and the amount of light increases. Thus, nighttime light can be a proxy variable for building height. However, in low-rise residential areas, the number of windows is not proportional to height, and a good correlation cannot be found. Thus, light intensity is not proportional to height.

FIGURE 7
Figure inferring buildings from nighttime light data in Hyogo Prefecture.

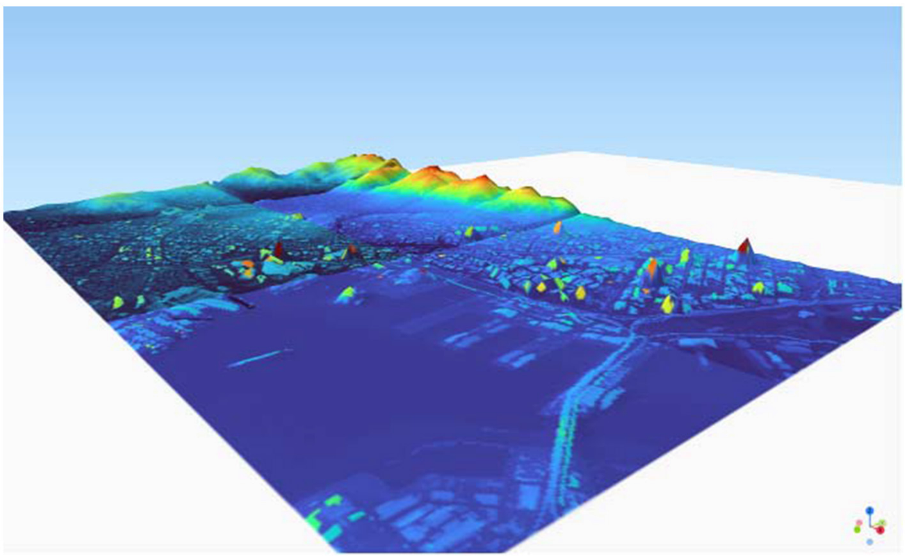


FIGURE 8
Reproduction of buildings in the vicinity of Sannomiya to Nagata in Hyogo Prefecture from DSM data.

Using the R-square of 0.288 of Henderson and Adam (2012) as a benchmark, this study obtained a good value of R-square of 0.425 for a building coverage of 70%-80%. However, this study was able to show that nighttime lighting correlates with buildings of a certain height but not with buildings with low-height building structures, such as those found in residential areas.

Using the present results, the height of buildings was estimated from the nighttime light in Hyogo Prefecture, and the buildings are depicted in Figure 7. For comparison, Figure 8 uses DSM data to represent actual building heights; in Figure 8, blue, yellow, and red represent taller buildings as they become more elevated. This view shows that the data allows predictions for buildings of a certain height. On the other hand, Figure 7.

The residential area to the north of Sannomiya, the center of Kobe city, is blank and not represented in Figure 4, indicating that a certain height is necessary to estimate the height of buildings by nighttime light.

The spatial resolution of NL needs to be increased for more accurate estimation. 460 × 460m resolution is still too coarse, so the data currently needs to be corrected by hand, for example, by visually checking buildings on site. It can be expected.

Another possible reason for the low accuracy of the data is the time it took for the satellite to capture the NL. In Hyogo Prefecture, nighttime light was captured for about 2 h a day, which is not enough to capture the lights of the buildings accurately. In addition to using satellite imagery, it is also important to go to the site and understand the differences between the NL data and the satellite imagery to improve accuracy.

Conclusion and future works

This study examines whether nighttime light can be a proxy for building height. Building height is a very important measurement for estimating urban environment and design and for understanding industrial clusters. However, only tall buildings such as landmarks can be obtained from OpenStreetMap. To obtain height data and include low-rise buildings across the country, data would need to be purchased. Another major disadvantage is the difficulty in acquiring time-series data. Time-series data on building heights can be used to visually check changes in municipalities over time and can be applied to a variety of studies.

Therefore, we examined the possibility of using "nighttime lighting data," which has recently become available free of charge on a detailed time scale, for estimating the height of buildings, which is basically provided for a fee.

Due to the structure of buildings, the number of windows increases in proportion to the size of the building. As the number of windows increases, the amount of light leaking from the building windows at night increases and the amount of light emitted by the building increases. This study focused on this correlation.

If we can correlate NL with building height, we can replace NL with the height of all buildings in the world.

This study obtained a good value of R-square of 0.425 for a building coverage of 70–80%. Calculations were performed for buildings with building percentages of 50, 60, 80, and 90% or more within the mesh, but the most accurate results were obtained for buildings with building percentages of 70% or more within the mesh. When the percentage of buildings in the mesh was 80% or higher, the mesh was more residential and no correlation was observed. This indicates that further work is needed to find a correlation for buildings with a height of about 10–12 m, which is the height of an average Japanese house. Thus, NL was found to have a certain correlation with building height.

The challenge for this study is to test the results in other regions and countries in order to verify the generality of the results of this study. However, Hyogo Prefecture is one of the most suitable regions for this study; we plan to investigate whether there are regional differences in the correlation between NL and building height.

Another major issue to be investigated is the source of nighttime light. Nighttime light is a mixture of various light sources. For example, the light from streetlights reflects off the ground, is dispersed by aerosols, and is very strongly detected in satellite imagery. We consider it a very important task to analyze the spectra received by satellite sensors in order to investigate more precise correlations between buildings and nighttime light. We believe that with further analysis of this spectrum, it will be possible to determine not only the height of buildings from nighttime light alone, but also the outer frames of various objects, such as traffic signals and roads.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

GK: concept creation, methodology, drafting, and diagramming. DS: data analysis, peer review, and editing. SK: concept making, peer review, and figure preparation. FI and YY: concept making, peer review, and editing. All authors contributed to the paper and approved the submitted version.

Funding

This research was supported by (Research and Development Project on the Development of Earth Observation System Using Space Technology) of Mori Memorial Foundation.

Acknowledgments

We sincerely thank the members of the GSAIS Empirical Research Group for their various comments and suggestions and for their significant contributions to the completion of the paper.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsus.2022.908822/full#supplementary-material

References

Barr, J., Mizrach, B., and Mundra, K. (2015). Skyscraper height and the business cycle: separating myth from reality. *Appl. Econ.* 47–2 doi: 10.1080/00036846.2014.967380

Blumenstock, J., Cadamuro, G., and On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science.* 350. 1073–1076. doi: 10.1126/science.aac4420

Chen, X., and Nordhaus, D. W. (2019). VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP. *Remote Sens.* 11, 1057doi: 10.3390/rs11091057

Christian, O. (2019). Industrial growth in Sub-Saharan Africa: evidence from machine learning with insights from nightlight satellite images. *J. Econ. Stud.* 48–8. doi: 10.1108/JES-05-2020-0201

esri Japan. (2021). *Coordinate Systems Used in Japan.* Available online at: https://www.esrij.com/gis-guide/coordinate-and-spatial/coordinate-system-japan/ (accessed June 22, 2021).

e-Stat. (2021). *About this Site.* Available online at: https://www.e-stat.go.jp/about (accessed June 22, 2021).

Fundamental Geospatial Data Fundamental Geospatial Data Basic Items. (2021). Available online at: https://fgd.gsi.go.jp/download/ref_kihon.html (accessed June 22, 2021).

Geospatial Information Authority of Japan. (2021). *What is GIS*... Available online at: https://www.gsi.go.jp/GIS/whatisgis.html (accessed June 22, 2021).

GSI 2. (2021). *What can be done with the practical use of geospatial information (JPGIS).* Available online at: https://www.gsi.go.jp/GIS/jpgis-wj_about.html (accessed June 22, 2021).

Henderson, J. V., and Adam, D. N. (2012). Measuring economic growth from outer space. *Am. Econ. Rev.* 102, 994–1028 doi: 10.1257/aer.102.2.994

James, H. (2020). *Alternative Economic Indicater*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. doi: 10.17848/9780880996778

Japan Council for Promotion of Information Sharing on Infrastructures Operation of the G.-Geospatial Information Center. (2021). *A Key Infrastructure for the 21st Century - Toward the Distribution and Business Creation of Various G-Geospatial Information from Industry, Government and Academia.* Available online at: https://aigid.jp/web/wp-content/uploads/2016/11/geospatial_pressrelease.pdf (accessed June 22, 2021).

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science.* 353, 790–4. doi: 10.1126/science.aaf7894

Keola, S., and Andersson, M. (2015). Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. *World Dev.* 60, 322–334. doi: 10.1016/j.worlddev.2014.08.017

Mellander, C., Lobo, J., Stolarick, K., and Matheson, Z. (2015). Night-time light data: a good proxy measure for economic activity?. *PLoS ONE.* 10:e0139779. doi: 10.1371/journal.pone.0139779

Ministry of Internal Affairs and Communications. (2021). *Budget of National Statistics (Outline of Expenditure Budget for Statistical Services)*. Available online at: https://www.soumu.go.jp/toukei_toukatsu/index/seido/3.html (accessed June 27, 2021).

Ministry of Land Infrastructure Transport and Tourism. (2021a). *Appended Table 1::List of geographical objects [Appended Tables 1-2, 1-3]* Available online at: https://www.mlit.go.jp/crd/tosiko/GISguidance/pdf/11.pdf (accessed June 22, 2021).

Ministry of Land Infrastructure Transport and Tourism. (2021b). *GIS Home Page, National Land Data Download.* Available online at: https://nlftp.mlit.go.jp/index.html (accessed June 22, 2021).

Nordhaus, W. D., and Chen, X. (2012). *Improved Estimates of Using Luminosity as a Proxy for Economic Statistics: New Results and Estimates of Precision.* Cowles Foundation Discussion Paper No. 1857. Available online at: https://ssrn.com/abstract=2030960 or http://dx.doi.org/10.2139/ssrn.2030960 doi: 10.2139/ssrn.2030960 (accessed March 29, 2012).

frontiers | Frontiers in Sustainability

TYPE Systematic Review
PUBLISHED 21 September 2022
DOI 10.3389/frsus.2022.920955

Check for updates

OPEN ACCESS

EDITED BY
Dingsheng Li,
University of Nevada, Reno,
United States

REVIEWED BY
Elena Escrig Olmedo,
University of Jaume I, Spain
Mark Anthony Camilleri,
University of Malta, Malta

*CORRESPONDENCE
Alexander R. Keeley
keeley.ryota.alexander.416@m.kyushu-u.ac.jp

SPECIALTY SECTION
This article was submitted to
Quantitative Sustainability Assessment,
a section of the journal
Frontiers in Sustainability

RECEIVED 15 April 2022
ACCEPTED 30 August 2022
PUBLISHED 21 September 2022

CITATION
Keeley AR, Chapman AJ, Yoshida K,
Xie J, Imbulana J, Takeda S and
Managi S (2022) ESG metrics and
social equity: Investigating
commensurability.
Front. Sustain. 3:920955.
doi: 10.3389/frsus.2022.920955

ESG metrics and social equity: Investigating commensurability

Alexander R. Keeley[1,2]*, Andrew J. Chapman[3], Kenichi Yoshida[1], Jun Xie[1], Janaki Imbulana[1], Shutaro Takeda[2] and Shunsuke Managi[1,2]

[1]Department of Civil Engineering, Kyushu University, Fukuoka, Japan, [2]Urban Institute, Kyushu University, Fukuoka, Japan, [3]International Institute for Carbon-Neutral Energy Research, Kyushu University, Fukuoka, Japan

During the past two decades, the world has seen exponential growth in the number of companies reporting environmental, social, and governance (ESG) data, and various ESG metrics have been proposed and are now in use. ESG metrics play a crucial role as an enabler of investment strategies that consider ESG factors, which are often referred to as "ESG investments". The ESG metrics and investment market are evolving rapidly, as investors, corporations, and the public are giving more priority to the "S" in ESG, including social equity issues, such as diversity, income inequality, worker safety, systemic racism, and companies' broader role in society. In this critical, systematic review, utilizing in-depth assessments, we investigate and compare the approaches employed in major ESG metrics and studies, then, we shed light on the "S" aspect by reviewing existing approaches used to assess social equity to clarify commensurability with ESG. Through the systematic review, this paper confirms that ESG investments can be expected to provide stable and high returns especially over the long term. This paper also clarifies how elements considered in social equity studies are largely reflected in major ESG metrics.

KEYWORDS

ESG metrics, social equity, systematic review, sustainability index, sustainable investment

Introduction

Environmental, social, and governance (ESG) investment has become an opportunity for businesses to tap into the growing social demand for lasting change and the emerging ESG market. According to a report by the Global Sustainable Investment Alliance (Global Sustainable Investment Alliance, 2021), total global ESG investments in 2020 reached $35.3 trillion, which is an increase of 15% from 2018 and 55% from 2016. The $35.3 trillion figure represents 35.9% of the $98.4 trillion in assets managed by all of the institutional investors surveyed[1]. A comprehensive literature review by Camilleri (2020) confirms that the providers of financial capital are increasingly allocating funds toward positive impact and sustainable investments. Because of the growth in environmental and ethical consciousness, both consumers and investors want companies to consider these values. And the growth in such demand

1 GSIA has surveyed institutional investors in the five regions: Europe; the US; Canada; Australia; NZ; and Japan.

increases the importance of developing sound ESG metrics to evaluate ESG activities. As shown by the recent adoption of a proposal for a Directive on corporate sustainability due diligence by the European Commission on February 23, 2022 that aims to foster not just environmentally but also socially responsible corporate behavior throughout global value chains, investors, corporations, and the public are giving more priority to the “S” in ESG, including social equity issues, such as diversity, income inequality, worker safety, systemic racism, and companies’ broader role in society. Despite the growing importance of ESG metrics and social aspects of ESG, there is a lack of academic scholarship investigating the commensurability of these metrics, and especially how important social elements are reflected in these metrics.

In recent years, there has been a growing awareness of the problem of the disarray of standards for disclosing ESG information, which is important in assessing the ESG initiatives of companies. The main disclosure standards for ESG information vary, depending on the purpose of the disclosure, such as the areas to be disclosed, whether it is principle or detailed based, the assumed stakeholders, disclosure channels, principles to be followed, and disclosure items. There has been a move toward the unification of standards, including a joint statement by standard-setting bodies and a proposal by the International Financial Reporting Standards Foundation to set sustainability reporting standards. Many studies have critiqued the lack of common theorization and commensurability among the ESG metrics mainly used in the market (Chatterji et al., 2016). Some studies have pointed out that the divergence of ESG ratings is mostly due to the differences in the scope, measurement, and weights of the metrics (Berg et al., 2022), but a critical analysis of ESG metrics and studies that have employed the metrics is required to solve the lack of common theorization and commensurability. We conduct this critical analysis in this study through a systematic review and a detailed examination of ESG metrics, and we investigate and compare the assessment approaches employed in the major ESG metrics and studies. Through the systematic review, we also examine the impact of ESG performance on financial performance and how the results differ among studies using different ESG metrics for the analyses. Additionally, to further examine how the major ESG metrics incorporate important social elements, we shed light on the “S” aspect by reviewing existing approaches used to assess social equity and examine how the elements considered in existing approaches are reflected.

Through the systematic review, this paper confirms that ESG investments can be expected to provide stable and high returns especially over the long term. Regarding the commensurability of the metrics, based on accessible methodology descriptions for four leading ESG metrics widely used in academic research, and business, this paper finds that the elements assessed have a significant divergence across the metrics: only four elements are common among all four ESG metrics, with the ratios of exclusive elements being 37.3, 38.1, 4.4, and 7.1% for the four metrics. This paper also clarifies how the elements considered in the social equity studies are reflected in the major ESG metrics. Some of the common factors that we find in the studies that evaluated social equity quantitatively are the concept of employment, such as relations, unemployment ratios and age groups, as well as income and education, which are also important elements in ESG metrics (e.g., gender balance, salary, and training). This paper also clarifies that access is a factor that can be quantified and used frequently in social equity studies, including access to energy, transportation, and essential facilities, whereas quantifying access is hardly observed in the major ESG metrics. The results of this paper contribute to advancing the research community’s and practitioners’ knowledge by providing a detailed examination of commensurability of the major ESG metrics, and how the ESG metrics capture important social elements.

The remainder of this paper is structured as follows. Section Systematic review of the literature that has employed ESG metrics provides the results of the systematic review of the ESG and social equity literature. Section ESG metrics and social equity: A closer look at the methodology and elements assessed provides a closer look at the major ESG metrics and critical elements of social equity studies. The discussion and conclusions, which are based on the systematic review and a detailed examination of the elements assessed, are presented in Section Discussion and conclusion.

Systematic review of the literature that has employed ESG metrics

This section provides the result of the systematic review of the ESG and social equity literature to capture the trends in the literature, such as investigated issues, geographical region, industries, and research fields.

Data

The ESG articles reviewed in this study are collected from Scopus, Thomson Reuters’ Web of Science, and the top-ranked journals in finance. Considering the fact that ESG studies have been increasingly undertaken in the past decades, we set the search period from January 1, 2000 to December 31, 2021. Following the keywords in previous systematic reviews, the keywords of the ESG topic include “CSR,” “corporate social responsibility,” “ESG,” and “environmental social* governance” (Kong et al., 2020; Widyawati, 2020), where the * stands for any other patterns of the word. The keywords of the ESG database include “MSCI,” “KLD,” “Kinder Lydenberg Domini,” “Refinitiv,” “Thomson Reuters Asset4,” “Bloomberg,” “FTSE Russell,” and “Arabesque S-Ray,” which are the major ESG data providers

10.3389/frsus.2022.920955

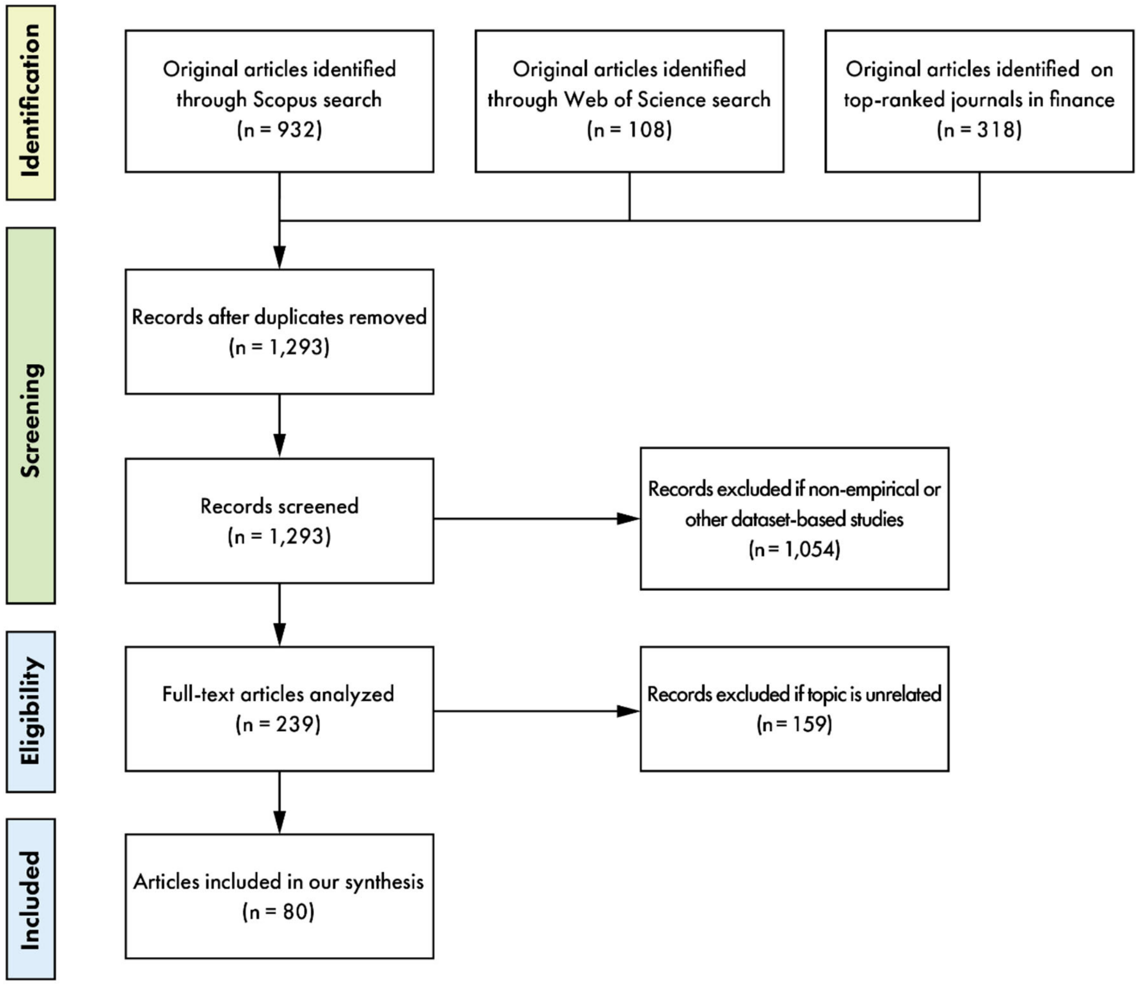


FIGURE 1
Screening procedure of ESG literature reviewed.

in the global market (Escrig-Olmedo et al., 2019). Keywords of financial performance that usually appear in the literature include "CFP," "financial performance," "stock return," "ROA," "ROE," and "Tobin's Q." We perform two search strategies in both Scopus and Web of Science[2]. Strategy 1 is keywords of ESG topic and financial performance, and Strategy 2 is keywords of ESG topic and ESG database. After filtering research articles in English and highly cited or hot papers on the Web of Science, Strategy 1 found 90 results, and Strategy 2 found 18 results. Moreover, after combining Strategies 1 and 2, 932 results were found on Scopus. We then searched for papers in the top-ranked journals in finance (Journal of Banking and Finance, Journal of Corporate Finance, Journal of Finance, Journal of Financial Economics, and Review of Financial Studies) to supplement the results from Scopus and Web of Science. This screening process was conducted by one author and verified by another author.

Figure 1 presents the screening procedure, starting from original articles in Scopus, Web of Science, and top-ranked journals in finance. After removing duplicated articles, we identified 1,293 articles. We manually checked all articles to filter out empirical studies that used the ESG database we are focusing on, and we had 239 articles. Figure 2 presents the number of publications in the selected empirical studies. Most of the studies that used ESG metrics were in the field of corporate finance, followed by specialized CSR journals. We then listed studies that discussed the impact of ESG activities on corporate financial performance. ESG activities are proxied by ESG scores or any specific ESG indicators in the ESG databases. Corporate financial performance is proxied by ROA, ROE, Tobin's Q, and other indicators of market return. After excluding studies that are not our focus, the final sample is 80 articles.

2 The search date is January 6th 2022.

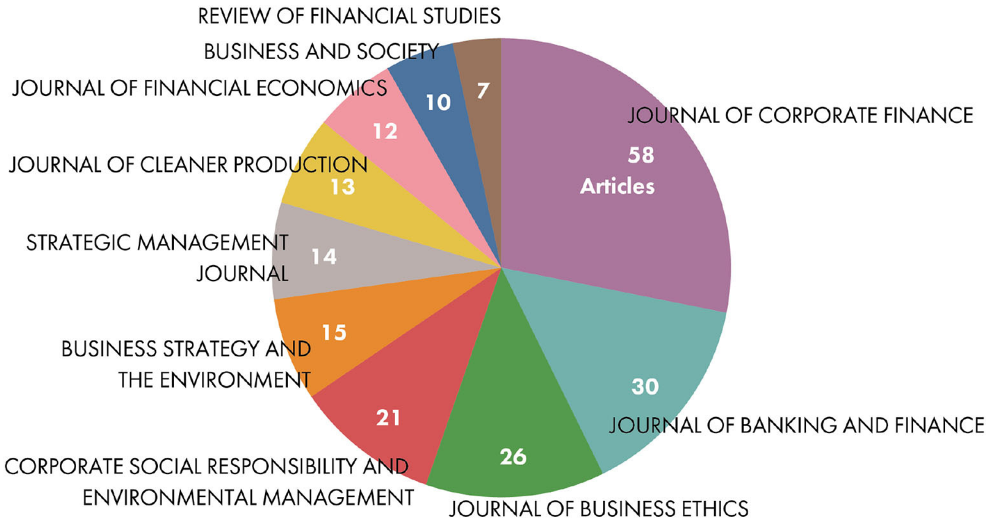


FIGURE 2
The number of publications among the selected empirical studies.

Regarding the social equity literature, following the review process for the ESG literature, the articles reviewed are from Scopus and Thomson Reuters' Web of Science, with the search period set from January 1, 2000 to December 31, 2021. As one of the main objectives of this study is to investigate how the elements considered in social equity evaluation studies are reflected in major ESG metrics, the keywords for this systematic review are "social equity," "assessment/evaluation," and "quantitative." The screening procedure is presented in Figure 3. Initially, there were 29 papers from the Web of Science and 170 papers from Scopus. After dropping duplicates, we obtained 172 articles. To identify papers with high impact, we employed a selection strategy where the databases were grouped by publication year—Group 1 ends in 2015; Group 2 is from 2016 to 2020, and Group 3 is 2021. In Group 1, papers whose citation count is less than the 50% average level are excluded. In Group 2, papers whose citation count is less than the 25% level are excluded. Regarding papers in 2021, all articles reflected the time-function nature of citations. After filtering out using our citation count quota approach, 128 articles were left. Finally, after excluding studies that did not focus on social equity evaluation, we had 24 articles and 26 case studies for the review.

Analysis and discussion

Based on the final articles selected from 2014 to 2021, we review how the conclusions and implications change across different topics and databases. Figure 1 depicts how the frequency of using ESG metrics increased in the reviewed period. Three databases that are mostly used in the literature are MSCI, Bloomberg, and Refinitiv. The number of publications increased in 2014 after MSCI's ESG database became available and kept growing in subsequent years. Initially, MSCI's ESG database was the most used. However, the number of studies that use Refinitiv's ESG database surged in 2021, becoming comparable to that of MSCI's ESG database. The use of Bloomberg's ESG database had a steady growth in the past 5 years (Figure 4).

Table 1 presents four panels that focus on different topics about the effect of ESG factors on corporate performance. In all panels, ESG factors (overall or each factor) are used as independent variables. Table 1A summarizes studies that used accounting measures, such as ROA, ROE, and EBITDA, as dependent variables; Table 1B summarizes studies that used market evaluation Tobin's Q, that is, firm value as dependent variables; Table 1C summarizes studies that used stock return as dependent variables; Table 1D summarizes studies that used the cost of capital and risk indicators as dependent variables.

We now discuss the systematic review results. Table 1 uses the notation "positive (negative)," "mostly positive (negative)," "partially positive (negative)," "mixed," or "not significant" for the conclusion of each study. Most of the studies considered in this systematic review estimated the relationship between dependent and independent variables multiple times under various models, with minor changes, to test for robustness. In Table 2, positive (negative) means that a "positive (negative)" coefficient value is observed in all the estimation models in each of the papers. In addition, "mostly" indicates a case in which most of the estimation models are positive (negative), whereas "partially" indicates a case in which positive (negative) results are reported in a few of the estimation models. However, "mixed" refers to cases where the study had different trends

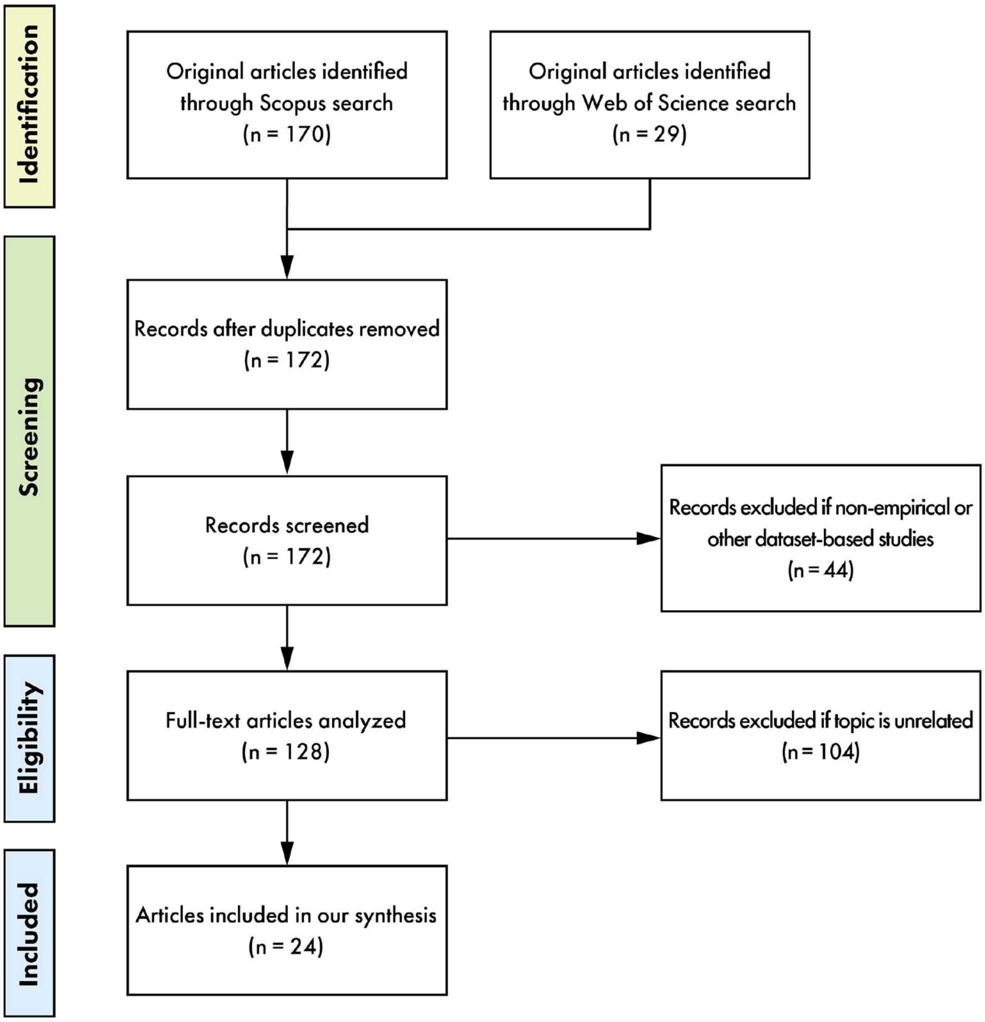


FIGURE 3
Screening procedure of social equity literature reviewed.

(positive and negative) depending on the estimation model. The square frames in Table 2 mean that the enclosed variables are estimated using the same formulas, and the variables enclosed in this square frame contain interaction terms. In this study, to simplify the discussion, the results are considered "mixed" even when the variables in the framework lack consistent trends (positive and negative) due to the influence of specific elements. In addition, for independent variables, most of the studies employed variables in which the greater the value, the higher the degree of ESG management. In contrast, some studies used non-ESG management variables (e.g., CSR concern, negative CSR, toxic firm dummy, and SIN stock), where the greater the value, the lower the degree of ESG management. Finally, the following discussion captures the whole trend of individual papers. If the same study reports both positive and negative trends, we count it as mixed. For studies that report both positive (negative) and non-significant trends, we count them as positive (negative). However, studies that employ more than two ESG metrics are excluded from the count. We also present the results of the reviewed studies on selected dependent variables in the form of heatmaps in Figure 5 (ROA, ROE, and EBITDA), Figure 6 (Stock Return), and Figure 7 (Tobin's Q). To show the trend in more simple way, "mostly positive" and "partially positive" results are presented as positive in the heatmaps and "mostly negative" and "partially negative" are presented as negative, while in the case of "mixed", one count is added to both positive and negative. The heatmaps presents the breakdown of the results by showing the count for ESG, E, S, G (and the combinations) as explanatory variables.

Regarding the relationship between ESG and profitability (see Table 1A; Figure 5), the results are mixed. We find that seven studies used Bloomberg as ESG metrics, among which

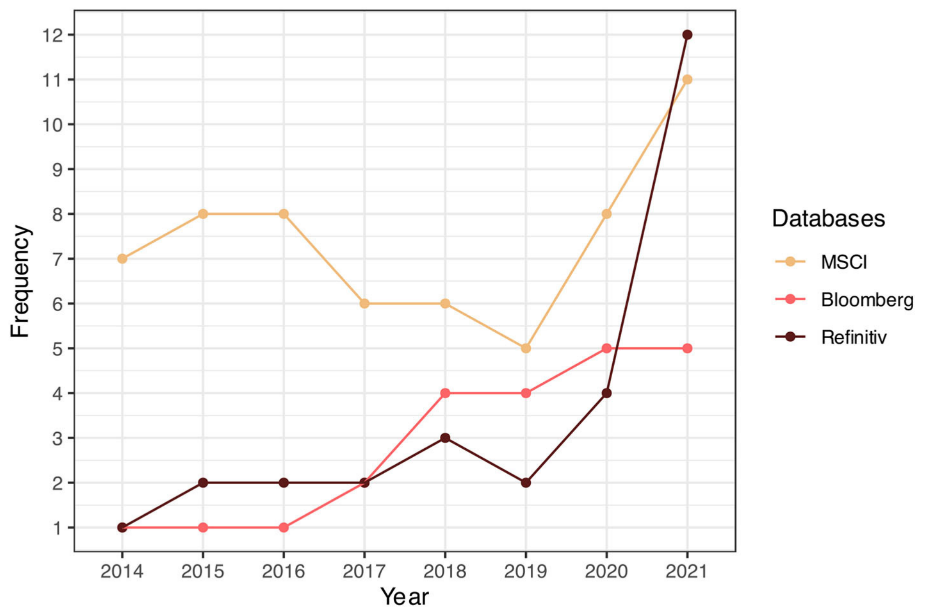


FIGURE 4
The frequency of using ESG metrics in reviewed papers. This figure shows the frequency of using ESG databases from 2014 to 2021. There are several studies using multiple databases.

three studies reported a positive relationship; one study found a non-significant relationship; two studies reported a mixed relationship, and one study reported a negative relationship. Similarly, seven studies used Refinitiv or Asset4 as ESG metrics, among which four studies reported a positive and significant trend, two mixed, and one negative. However, 17 studies used MSCI or KLD as ESG metrics, among which four studies suggested a negative relationship; nine studies suggested a positive relationship, and the remaining four studies found a mixed relationship. From the heatmap, we can observe that for ROA, ROE, and EBITDA "Positive" ≥ "Nagative" holds in all of the cases except for EBITDA with overall ESG score as explanatory variable. However, there are still statistically non-significant results and negative results that cannot be neglected. Thus, in the short term, it is hard to prove the positive effects of ESG factors on profitability.

Regarding Tobin's Q (see Table 1B; Figure 6), the studies revealed that ESG factors have a positive effect on firm value. Among the 20 studies that used MSCI and KLD as ESG metrics, all reported positive trends, except for six studies that reported a mixed result. This trend is also similar for studies that used Bloomberg, Refinitiv, and Thomson Reuters' Asset4 as ESG metrics, which generally confirms a positive direction, except for two studies that did not find statistical significance and five studies that reported a mixed trend. Figure 6 shows that "Positive" > "Nagative" holds in all of the cases. Therefore, regardless of the type of ESG metrics for Tobin's Q, it is confirmed that the most recent studies report a positive direction. In summary, ESG factors are found to have robust and positive effects on corporate performance in the long term.

Table 1C summarizes the effect of ESG factors on stock return. Refinitiv's Thomson Reuters Asset4 was used in five studies, among which three reported a positive direction; one suggested a mixed trend, and one did not find any significant result. However, 12 studies used MSCI and KLD as ESG metrics, among which seven reported positive effects; two reported a negative result; two suggested a mixed trend, and the remaining one reported a non-significant relationship. Based on these results, studies that used Refinitiv's Thomson Reuters Asset4 tend to have positive results, whereas those that used MSCI and KLD found more complicated results. From the heatmap, as shown in Figure 7, "Positive" ≥ "Negative" holds in most of the cases. However, the impact of ESG factors on stock returns naturally depends on the research period, samples, and other environmental factors. In Table 1C, studies that used interaction models mostly had complicated results, indicating the external contingency of ESG factors. Some studies focused on the time trend of negative shocks to the stock prices of many firms, but it is not necessary to compare this in an analysis that focuses on normal stock returns. We recognize these limitations, but we do not generalize and make comparisons for discussion.

Regarding the cost of equity and other risks (see Table 1D), the studies revealed a negative trend, which means that ESG factors are effective in reducing financial risk and cost. In the 12 studies that used MSCI and KLD as ESG metrics,

TABLE 1 List of studies using ESG metrics reviewed.

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
(A) Dependent variable: *ROA, ROE, EBITDA*					
Duque-Grisales and Aguilera-Caracuel (2021)	ROA	Thomson Reuters	Overall	ESG score	(Mostly) Negative
[-4mm]		Asset4	Overall	ESG score × slack	(Mostly) Positive
Griffin et al. (2021)	ROA	Thomson Reuters Asset4	E, S	E/S	(Mostly) Positive
Bátae et al. (2021)	ROA	Refinitiv	E	Env_RU (*resource use efficiency*)	Not significant
			E	Env_EM (*emission and waste reductions*)	Positive
			E	Env_IN (*environmental innovation*)	Not significant
			S	Soc_WF (*workforce*)	Not significant
			S	Soc_HRights (*human rights*)	Not significant
			S	Soc_COM (*community*)	Not significant
			S	Soc_PRD (*product responsibility*)	Not significant
			G	Gov_MN (*management and oversight*)	Negative
			G	Gov_SH (*shareholder rights*)	Not significant
			G	Gov_CSR (*CSR strategy*)	Not significant
	ROE	Refinitiv	E	Env_RU	Not significant
			E	Env_EM	Positive
			E	Env_IN	Not significant
			S	Soc_WF	Not significant
			S	Soc_HRights	Not significant
			S	Soc_COM	Not significant
			S	Soc_PRD	Not significant
			G	Gov_MN	Not significant
			G	Gov_SH	Not significant
			G	Gov_CSR	Not significant
Kuzey et al. (2021)	ROA	Refinitiv	Overall	ESGs	Not significant
			Overall	ΔESGs	Not significant
			Overall	CSRcom (*CSR committee*)	(Partially) Negative
			Overall	ESGs × CSRcom	Not significant
			Overall	ΔESGs × CSRcom	Not significant
	ROE	Refinitiv	Overall	ESGs	Not significant
			Overall	ΔESGs	Not significant
			Overall	CSRcom	Not significant
			Overall	ESGs × CSRcom	Not significant
			Overall	ΔESGs × CSRcom	Not significant
Atif et al. (2021)	ROA	Bloomberg	G	WOBP (*% of women on the board*)	Positive
			G	WOBP × REN/TC (*Total renewable energy consumption as a percentage of total energy use*)	Positive
	ROE	Bloomberg	G	WOBP × REN/TC	Positive
Naseem et al. (2020)	ROA	Thomson Reuters	Overall	CSR	Positive
		Asset4	Overall	PCSRhat (*predicted value of CSR*)	Positive
	ROE	Thomson Reuters	Overall	CSR	Positive
		Asset4	Overall	PCSRhat (*predicted value of CSR*)	(Partially) Positive
Cai et al. (2020)	ROA Low = 1, ROA High = 2 (*multinomial probit regressions*)	MSCI, KLD	Overall	Net adjusted CSR score	Positive

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
Liu et al. (2020)	ROA	MSCI, KLD	Overall	CSR	Positive
	ROE	MSCI, KLD	Overall	CSR	Not significant
Nguyen et al. (2020)	ln (1 + profitability)	MSCI, KLD	Overall	CSR proxy	Positive
			Overall	Long-term investor ownership × CSR proxy	Not significant
Devie et al. (2020)	CFP (*corporate financial performance*)	Bloomberg, other reliable sources	Overall	CSR	(Mostly) Positive
Alareeni and Hamdan (2020)	ROA	Bloomberg	Overall	ESG index	Positive
			E	EVN index	Negative
			Overall	CSR index	Negative
			G	CG index	Positive
	ROE	Bloomberg	Overall	ESG index	Positive
			E	EVN index	Negative
			Overall	CSR index	Negative
			G	CG index	Negative
Hoang et al. (2020)	ROA	Bloomberg	E	EDS (*environmental disclosure score*)	(Mostly) Positive
			E	GHG (*greenhouse gas emissions per unit of revenue*)	(Mostly) Positive
			E	WATER (*total water uses per unit of revenue*)	Not significant
			E	WASTE (*total waste per unit of revenue*)	Not significant
Saleem et al. (2021)	ROA	Bloomberg	G	GDev-index-index (*the governance deviance index*)	Negative
Albuquerque et al. (2019)	Change in ROA	MSCI, KLD	Overall	CSR1 variable	Not significant
			Overall	CSR1 × GDP growth	Negative
			Overall	CSR2 variable	Not significant
			Overall	CSR2 × GDP growth	Negative
Luffarelli et al. (2019)	EBITDA	MSCI, KLD	Overall	CSP	Not significant
			Overall	CSP × PMP (*product-market profile*)	Negative
Xie et al. (2019)	ROA	Bloomberg	E	Verification type	Not significant
			E	Green building policy	Positive
			E	Sustainable packaging	Positive
			E	Environmental quality management policy	Not significant
			E	Environmental supply chain management	Not significant
			E	Climate change policy	Not significant
			E	Climate change opportunities discussed	Not significant
			E	Risks of climate change discussed	Not significant
			E	Emissions reduction initiatives	Not significant
			E	New products climate change	Not significant
			E	Energy efficiency policy	Not significant
			S	Equal opportunity policy	Not significant
			S	Human rights policy	Not significant
			S	Training policy	Not significant
			S	Employee CSR training	Negative

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
			S	Health and safety policy	Not significant
			S	Fair remuneration policy	Not significant
Fauver et al. (2018)	ROA	Thomson Reuters Asset4	S	EF Index	Positive
Brogi and Lagasio (2019)	ROA	MSCI, KLD	Overall	ESGSCORE	Not significant
			E	ESCORE	Not significant
			S	SSCORE	Not significant
			G	GSCORE	(Partially) Positive
Byun and Oh (2018)	ΔROA	MSCI, KLD	S	Net CSR coverage (*positive CSR-related news articles minus negative CSR-related news articles. its articles covering topics in community, diversity, and employee relations*)	Positive
			Overall	KLD index	Not significant
Hoi et al. (2018)	ROA	MSCI, KLD	Overall	CSP	(Partially) Positive
			Overall	CSP × high social capital	Positive
			Overall	Positive CSR	Positive
			Overall	Positive CSR × high social capital	Positive
			Overall	Negative CSR	Not significant
			Overall	Negative CSR × high social capital	Negative
Bhandari and Javakhadze (2017)	ROA	MSCI, KLD	Overall	KLD	Negative
Wang and Sarkis (2017)	ROA	Bloomberg	G	CSRGOV	Not significant
			E	CSRENV	Not significant
Cornett et al. (2016)	ROA	MSCI, KLD	Overall	ESG index	Positive
			Overall	ESG index × small	(Partially) Positive
	ROE	MSCI, KLD	Overall	ESG index	Positive
			Overall	ESG index × small	(Partially) Positive
Harrison and Berman (2016)	ROA	MSCI, KLD	Overall	CSP (*total strengths*)	Negative
			Overall	CSP (*total concerns*)	(Partially) Negative
Tebini et al. (2016)	ROA	MSCI, KLD	E	Envt	Not significant
			E	Envt (-1)	Positive
			E	Envt (-2)	Positive
			E	Envt (-3)	Positive
			E	Envt × size	Positive
			E	Envt × invest	Negative
			E	Envt × beta (*systematic risk*)	Negative
Lys et al. (2015)	ΔROA	Thomson Reuters Asset4	Overall	CSR	Positive
			E	ENV_COMP	Not significant
			S	SOC_COMP	Positive
			G	CORPGOV	Not significant
Nguyen and Nguyen (2015)	ROA	MSCI, KLD	Overall	Aggregate strengths	Positive
			Overall	Aggregate concerns	Positive
Boesso et al. (2015)	EBITDA	MSCI, KLD	S	Community	(Mostly) Positive
			G	Governance	(Mostly) Positive

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
			S	Diversity	(Mostly) Positive
			S	Employee	(Mostly) Positive
			E	Environment	(Mostly) Negative
			S	Human rights	Negative
			S	Product	Negative
Di Giuli and Kostovetsky (2014)	ROA Δ next 3 years	MSCI, KLD	Overall	KLD strengths	Negative
			Overall	KLD concerns	Not significant
Kumar et al. (2016)	EBITDA	MSCI, KLD	E	Employee weakness	Not significant
			E	Employee strengths	Positive
			S	Costumer weakness	Not significant
			S	Costumer strengths	Not significant
			S	Community weakness	Not significant
			S	Community strengths	Not significant
			G	Governance weakness	Not significant
			G	Governance strengths	Positive
			E	Environment weakness	Not significant
			E	Environment strengths	Not significant
			S	Diversity weakness	Negative
			S	Diversity strengths	Positive
			S	Human rights weakness	Negative
			S	Human rights strengths	Positive
Moura-Leite et al. (2014)	ROA	MSCI, KLD	E, S	Primary stakeholder management	Positive
(B) Dependent variable: *Tobin's Q*					
Griffin et al. (2021)	Tobin's Q	Thomson Reuters Asset4	E, S	E/S	Positive
Bátae et al. (2021)	TQ	Refinitiv	E	Env_RU (*resource use efficiency*)	Not significant
			E	Env_EM (*emission and waste reductions*)	Not significant
			E	Env_IN (*environmental innovation*)	Not significant
			S	Soc_WF (*workforce*)	Not significant
			S	Soc_HRights (*human rights*)	Not significant
			S	Soc_COM (*community*)	Not significant
			S	Soc_PRD (*product responsibility*)	Not significant
			G	Gov_MN (*management and oversight*)	Not significant
			G	Gov_SH (*shareholder rights*)	Not significant
			G	Gov_CSR (*CSR strategy*)	Not significant
Kuzey et al. (2021)	Tobin's Q	Refinitiv	Overall	ESGs	Mixed
			Overall	ΔESGs	Not significant
			Overall	CSRcom (*CSR committee*)	(Partially) Negative
			Overall	ESGs × CSRcom	(Partially) Positive
			Overall	ΔESGs × CSRcom	Not significant
Dai et al. (2021)	Market-to-book	MSCI, KLD, Thomson Reuters Asset4	Overall	$CSR^c \times CSR^s$ _Supplier controls	Positive
			Overall	CSR^s _Supplier controls	Positive
			Overall	CSR^c _Supplier controls	Negative

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
			Overall	$CSR^c \times CSR^s$ _Customer controls	Positive
			Overall	CSR^s _Customer controls	Negative
			Overall	CSR^c _Customer controls	Not significant
Bu et al. (2021)	Tobin's Q	MSCI, KLD	Overall	CSR_PRE (*a variable denoting the optimal level of CSR activities*)	Positive
			Overall	TID (*talented inside directors*) × CSR_PRE	Positive
			Overall	CSR_RES (*excessive level of CSR activities, calculated as CSR minus CSR_PRE*)	Negative
			Overall	TID × CSR_RES	(Partially) Positive
			Overall	CSR	Positive
			Overall	TID × CSR	Positive
Ertugrul and Marciukaityte (2021)	log (Tobin's q)	MSCI, KLD	Overall	CSR net (*CSR strengths – CSR concerns*)	(Partially) Positive
			Overall	Unionization × CSR net	Negative
Lu et al. (2021)	Tobin's Q	MSCI, KLD	Overall	CSR	Not significant
			Overall	CSR × financial risk	Positive
			Overall, E	CSR × environmental risk	Positive
			Overall	CSR × earnings stability	Positive
			Overall	CSR × sales growth	Not significant
Hannah et al. (2021)	Tobin's Q	KLD, MSCI, Bloomberg	Overall	CSR	Positive
			Overall	CSR^2	Partially) Positive
			Overall	CSR	Positive
			Overall	CSR^2	(Partially) Positive
			Overall	CSR × SalesGR	(Partially) Positive
			Overall	CSR × AssetGR	(Partially) Positive
Atif et al. (2021)	Tobin's q	Bloomberg	G	WOBP (*% of women on the board*) × REN/TC (*Total renewable energy consumption as a percentage of total energy use*)	Positive
Jia and Li (2020)	TobinsQ	Thomson Reuters Asset4	Overall	CSPD (*above the sample median of sustainability performance*)	Positive
			Overall	ECC (*exposure to climate change*) × CSPD	Positive
			Overall	CSPD	(Partially) Positive
			Overall	EPU (*economic policy uncertainty*) × CSPD	Positive
			Overall	CSPD	(Partially) Positive
			Overall	POLI (*political instability*) × CSPD	Positive
Bardos et al. (2020)	Tobin's Q	MSCI, KLD	E, S	Community/environmental CSR	(Partially) Positive
Brower and Dacin (2020)	Tobin's Q	MSCI, KLD	Overall	Overall CSR activities (lag)	Positive
			Overall	Primary CSR activities (lag) (*primary CSP level is calculated as the sum of the firm's CSP strength scores for governance, employee relations, and product strengths for each firm-year observation in the data*)	Positive
			Overall	Secondary CSR activities (lag) (*secondary CSP level is calculated as the sum of each firm's CSP strength scores for environmental impact, community involvement, and diversity strengths for each firm year observation in the data*)	Positive

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
Liu et al. (2020)	Tobin's Q	MSCI, KLD	Overall	CSR	Positive
Nguyen et al. (2020)	ln (market-to-book)	MSCI, KLD	Overall	CSR proxy	Positive
			Overall	Long-term investor ownership × CSR proxy	Positive
Alareeni and Hamdan (2020)	Tobin's Q	Bloomberg	Overall	ESG index	Positive
			E	EVN index	Positive
			Overall	CSR index	Positive
			G	CG index	Positive
Saleem et al. (2021)	Tobin's-Q	Bloomberg	G	Gdev-index-index (*the governance deviance index*)	Positive
Boubakri et al. (2019)	Tobin's Q	Thomson Reuters Asset4	Overall	PCSR (*predicted CSR intensity*)	(Mostly) Positive
			Overall	STATE (*percentage of shares held by a government*) × PCSR	Positive
Albuquerque et al. (2019)	Tobin's Q	MSCI, KLD	Overall	lagged CSR1 variable	Positive
			Overall	lagged CSR2 variable	Positive
Luffarelli et al. (2019)	Tobin's q	MSCI, KLD	Overall	CSP	Not significant
			Overall	CSP × PMP (*product-market profile*)	Positive
Zolotoy et al. (2019)	Tobin's Q	MSCI, KLD	Overall	CSR	(Mostly) Positive
			Overall	CSR × religious adherence	(Mostly) Negative
Xie et al. (2019)	Market value (Tobin's Q)	Bloomberg	E	Verification type	Positive
			E	Green building policy	Not significant
			E	Sustainable packaging	Positive
			E	Environmental quality management policy	Negative
			E	Environmental supply chain management	Not significant
			E	Climate change policy	Not significant
			E	Climate change opportunities discussed	Not significant
			E	Risks of climate change discussed	Not significant
			E	Emissions reduction initiatives	Not significant
			E	New products climate change	Not significant
			E	Energy efficiency policy	Not significant
			S	Equal opportunity policy	Positive
			S	Human rights policy	Positive
			S	Training policy	Positive
			S	Employee CSR training	Not significant
			S	Health and safety policy	Not significant
			S	Fair remuneration policy	Not significant
Fauver et al. (2018)	Tobin's Q	Thomson Reuters Asset4	S	EF (*employee-friendliness*) index	(Mostly) Positive
Byun and Oh (2018)	log (Tobin's q)	MSCI, KLD	S	Net CSR coverage (*positive CSR-related news articles minus negative CSR-related news articles. its articles covering topics in community, diversity, and employee relations*)	Positive
			Overall	KLD index	(Partially) Positive

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
Buchanan et al. (2018)	Tobin's Q	Bloomberg	Overall	CSR	Positive
			Overall	CSR × crisis (2008Q3 - 2009Q1)	Negative
Taylor et al. (2018)	Tobin's Q	Bloomberg	Overall	ESG, social, environmental, governance disclosure	(Partially) Positive
			Overall	ADSALE (*advertising expenditures to sales*) × ESG, social, environmental, governance	(Partially) Positive
			Overall	CSR firm	Negative
Yu et al. (2018)	Industry-adjusted Tobin's Q	Bloomberg	Overall	ESG disclosure (industry-adjusted)	(Partially) Negative
			Overall	$(\text{ESG disclosure})^2$	(Partially) Positive
Shahzad and Sharfman (2017)	Tobin's q	MSCI, KLD	Overall	CSP	Mixed
Wang and Sarkis (2017)	Tobin's Q	Bloomberg	G	CSRGOV	Not significant
			E	CSRENV	Not significant
Hawn and Ioannou (2016)	Log Tobin's q	Thomson Reuters Asset4	Overall	Internal $(\text{CSR})_{t-1}$ /assets	Not significant
			Overall	External $(\text{CSR})_{t}$ /assets	Positive
Cornett et al. (2016)	Tobin's Q	MSCI, KLD	Overall	ESG index	(Partially) Positive
			Overall	ESG index × small	(Partially) Positive
Ferrell et al. (2016)	Tobin's Q	MSCI, KLD	Overall	CSR	(Partially) Positive
			Overall	CSR × entrenchment index	Positive
Cahan et al. (2015)	Tobin's Q	MSCI, KLD	Overall	CSR	Not significant
			Overall	CSR × media	Not significant
			Overall	CSR × H-H	Positive
			Overall	CSR × media × H-H	Positive
Gao and Zhang (2015)	Tobin's Q	MSCI, KLD	Overall	CSR score	Not significant
			Overall	CSR × DAS (*discretionary accrual smoothing*)	Positive
Jha and Cox (2015)	Tobin's Q	MSCI, KLD	Overall	CSR_S (*it is the sum of CSR_STRENGTHS and CSR_CONCERNS. the detailed descriptions of how CSR_STRENGTHS and CSR_CONCERNS are calculated are described later in this table. a higher number indicates greater social responsibility*)	Positive
			Overall	CSR_S	Positive
			Overall	CSR_S × HIGH SOCIAL CAPITAL	Not significant
Nguyen and Nguyen (2015)	Tobin's Q	MSCI, KLD	Overall	Aggregate strengths	Positive
			Overall	Aggregate concerns	Positive
Vomberg et al. (2015)	Tobin's Q	MSCI, KLD	S	Human capital	Not significant
			S	Brand equity × human capital	Positive
			S	Human capital × FMCG (*fast moving consumer goods*)	Negative
			S	Human capital × consumer durables	Negative
			S	Human capital × retail	Negative
Moura-Leite et al. (2014)	Tobin's Q	MSCI, KLD	E, S	Primary stakeholder management	Positive

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
(C) Dependent variable: *Stock return, contains CAR, AR*					
Ding et al. (2021)	Weekly stock return	Thomson Reuters	Overall	CSR score × COVID19	Positive
	Abnormal return	Asset4	Overall	CSR score × COVID19	Positive
Garel and Petit-Romec (2021)	Stock returns (*Feb. 20–Mar. 20*)	Thomson Reuters Asset4	E	Environmental score	(Mostly) Positive
Bose et al. (2021)	CAR	Refinitiv	Overall	HIGH_CSR	(Mostly) Positive
			E, Overall	LNEMISSION × HIGH_CSR	(Mostly) Negative
Bolton and Kacperczyk (2021)	Stock returns (RET)	MSCI, KLD,	E	SCOPE 1	(Partially) Positive
		Thomson Reuters	E	SCOPE 2	(Partially) Positive
		Asset4, Bloomberg	E	SCOPE 3	(Partially) Positive
Bae et al. (2021)	Raw_firm-level stock returns	MSCI, KLD	Overall	CSR_MSCI	Not significant
		Refinitiv, Thomson Reuters Asset4	Overall	CSR_REFINITIV	Not significant
	Mkt-adj_firm-level stock returns	MSCI, KLD	Overall	CSR_MSCI	Not significant
		Refinitiv, Thomson Reuters Asset4	Overall	CSR_REFINITIV	(Partially) Positive
Avramov et al. (2021)	Excess return	MSCI, KLD, MSCI IVA, Bloomberg, Asset4 (Refinitiv), Sustainalytics, and RobecoSAM	Overall	ESG	Not significant
			Overall	ESG × low ESG uncertainty	Negative
			Overall	Low ESG uncertainty	(Partially) Positive
	CAPM-adjusted return		Overall	ESG	Not significant
			Overall	ESG × low ESG uncertainty	Negative
			Overall	Low ESG uncertainty	(Partially) Positive
Doukas and Zhang (2021)	CAR (-3, +3)	MSCI, KLD	Overall	Adjusted CSR (*compute the total strengths and total concerns for each category and then divide the scores for each category by the respective maximum numbers of strength and concern scores to obtain adjusted strength and concern scores for each dimension. Finally, take the net difference between the total adjusted strength and total adjusted concern scores*)	(Partially) Negative
			Overall	Adjusted CSR × MA (*managerial ability*) -Score	Positive
	One-year BHAR		Overall	Adjusted CSR	Negative
			Overall	Adjusted CSR × MA-score	Positive
Liang et al. (2020)	Acquirer CAR [−1, +1]	Thomson Reuters Asset4	S	Acquirer employment quality (*domestic*)	Positive
			S	Acquirer employment quality (*cross-border*)	Negative
Boone and Uysal (2020)	CAR (−5, +5)	MSCI, KLD	E	Different reputation dummy (*takes a value of one if an acquirer and its target do not fall into the same environmental grouping*)	Negative
			E	Green firm dummy	Not significant

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
			E	Ratio of green firms	Not significant
			E	Toxic firm dummy	Negative
			E	Green firm dummy	Not significant
P.-A. Nguyen et al. (2020)	Excess stock returns	MSCI, KLD	Overall	CSR proxy	Not significant
			Overall	Long-term investor ownership × CSR proxy	(Partially) Negative
Tong et al. (2020)	Acquirer announcement return	MSCI, KLD	Overall	Target CSR	(Mostly) Positive
Zolotoy et al. (2019)	Market model AR_during 2008–2009	MSCI, KLD	Overall	CSR	Positive
			Overall	CSR × religious adherence	Negative
	Fama–French–Carhart model AR_during 2008–2009	MSCI, KLD	Overall	CSR	Positive
			Overall	CSR × religious adherence	Negative
Dutordoir et al. (2018)	CAR: *SEOs (seasoned equity offerings) announcements*	MSCI, KLD	Overall	AdjCSR (*sum of yearly adjusted community activities, diversity, employee relations, environmental record, human rights, and product quality and safety scores from KLD*)	Positive
Feng et al. (2018)	CAR_SEO (*seasoned equity offerings*)	MSCI, KLD	Overall	CSR	Positive
Choy et al. (2017)	CAR	Thomson Reuters Asset4	Overall	Corporate social responsibility	Not significant
Lins et al. (2017)	Raw return	MSCI, KLD	Overall	CSR	(Mostly) Positive
	Abnormal return		Overall	CSR	(Partially) Positive
Shiu and Yang (2017)	Abnormal returns	MSCI, KLD	Overall	Short-term CSR engagement	(Partially) Positive
			Overall	Long-term CSR engagement	Positive
	Cumulative abnormal returns		Overall	Short-term CSR engagement	Not significant
			Overall	Long-term CSR engagement	Positive
Gao and Zhang (2015)	Rett (*ex-dividend stock return during fiscal year t*)	MSCI, KLD	Overall	CSR × DAS	Positive
Borghesi et al. (2014)	Annual stock return premiums (*1992 to 2006*)	MSCI, KLD	Overall	CSR	Not significant
			Overall	Industry adjusted CSR	Not significant
Di Giuli and Kostovetsky (2014)	Stock returns	MSCI, KLD	Overall	KLD strengths	(Partially) Negative
	(monthly)		Overall	KLD concerns	Not significant
	Stock returns (monthly) Fama-MacBeth		Overall	KLD strengths	(Partially) Negative
			Overall	KLD concerns	Not significant
(D) Dependent variable: *Cost of equity, Other Risks*					

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
Tsang et al. (2021)	ROA_Volatility	Thomson Reuters Asset4	Overall	CSRContracting (*an indicator variable that equals 1 if senior executives' compensation is linked to CSR/H&S (Health and Safety)/sustainability targets (CSR contracting) in the year and 0 otherwise*)	Positive
			S, E	CSRPerf (*the average of Social performance score and Environmental performance score*)	Negative
	Stock_Return_Volatility		Overall	CSRContracting	Positive
			S, E	CSRPerf	Negative
Chen et al. (2021)	NSKEW (*the negative skewness of firm-specific weekly returns over the fiscal year period*)	MSCI, KLD	Overall	CSR	Not significant
	DUVOL (*the natural logarithm of the ratio of the standard deviation in the "down" weeks to the standard deviation in the "up" weeks*)		Overall	CSR	Not significant
Devie et al. (2020)	RISK	Bloomberg, other reliable sources	Overall	CSR	Negative
Boubakri et al. (2019)	Cost of equity capital	Thomson Reuters Asset4	Overall	PCSR (*predicted CSR intensity*)	Not significant
			Overall	STATE (*percentage of shares held by a government*) × PCSR	Negative
Albuquerque et al. (2019)	Firm Beta	MSCI, KLD	Overall	lagged CSR1 variable	Negative
			Overall	lagged CSR2 variable	Negative
Chang et al. (2019)	BETA (*systematic risk*)	MSCI, KLD	Overall	SD_CSR (*the standardized CSR score, which is equal to the sum of standardized CSR scores over six categories: environment, community, human rights, diversity, employee relations, and product*)	Negative
Albarrak et al. (2019)	Cost of equity	Bloomberg	E	ENV_COMMITEE (*environmental committee*)	(Partially) Negative
			E	ENV_SCORE	Not significant
			E	iCarbon × ENV_SCORE	Not significant
Lueg et al. (2019)	TRSK (*Total Risk*)	Bloomberg	Overall	ESG	Not significant
	BETA (*Systematic Risk*)		Overall	ESG	Negative
	IDIO (*Idiosyncratic Risk*)		Overall	ESG	Not significant
Breuer et al. (2018)	Implied cost of equity	Thomson Reuters Asset4	Overall	CSR	(Partially) Negative
	BETA		Overall	CSR	Not significant
	SIGMA		Overall	CSR	Not significant

(Continued)

TABLE 1 (Continued)

References	Dependent variable	ESG metrics	Independent variable		Conclusion
			Overall, E, S, G	Variable name	(Positive; Negative; Mixed)
Bhandari and Javakhadze (2017)	Alpha 3factor	MSCI, KLD	Overall	KLD	Negative
	Alpha 4factor		Overall	KLD	Negative
El Ghoul and Karoui (2017)	Alpha	MSCI, KLD	Overall	CSR (CSR score)	Negative
			Overall	CSR (Strengths)	Negative
			Overall	CSR (Concerns)	Not significant
Oh et al. (2017)	Idiosyncratic risk	MSCI, KLD	Overall	ADV (CSR)	Not significant
			Overall	ADV (CSR) × SIN stock	Positive
			Overall	Probability of KLD report	Not significant
L. Cai et al. (2016)	CAPM_BETA	MSCI, KLD	E	ENV	Negative
	FF_MKT_BETA		E	ENV	Negative
	DEVRET		E	ENV	Negative
Cheung (2016)	idio (*idiosyncratic risk*)	MSCI, KLD	Overall	CSR	Negative
	beta (*systematic risk*)		Overall	CSR	Negative
Becchetti et al. (2015)	Idiosyncratic volatility	MSCI, KLD	E, S	Stakeholder risk (*stakeholder risk as the relative sum of weaknesses (concerns) in corporate responsibility in the domains of community, diversity, employee relations, environment, human rights, and product quality according to official ratings of a primary (KLD) CSR rating agency*)	Negative
Cahan et al. (2015)	Cost of capital	MSCI, KLD	Overall	CSR	Not significant
			Overall	CSR × media	Not significant
			Overall	CSR × H-H	Not significant
			Overall	CSR × media × H-H	Negative
Ng and Rezaee (2015)	Cost of equity_IndEP	MSCI, KLD	E	ENV	Negative
			S	SOC	(Partially) Negative
			G	GOV	Negative
			Overall	KLD	Negative
	Cost of equity_GORDON		E	ENV	Negative
			S	SOC	(Partially) Negative
			G	GOV	Negative
			Overall	KLD	Negative
Kim et al. (2014)	NCSKEW (*the negative conditional skewness of firm-specific weekly returns over the fiscal year*)	MSCI, KLD	Overall	CSR_SCORE	Negative
	DUVOL (*the natural logarithm of the ratio of the standard deviation in the "down" weeks to the standard deviation in the "up" weeks*)		Overall	CSR_SCORE	Negative

The frame means that the enclosed variables are estimated using the same formula.

TABLE 2 Comparison of ESG score rating structure.

	Refinitiv	MSCI	Arabesque S-ray
Environmental	Resource use	Natural resources	Resource use
	Emissions	Climate change	Emissions
	Innovation	Environmental opportunities	Environmental solutions
		Pollution and waste	Waste
			Water
			Environmental stewardship
			Environmental management
Social	Workforce	Human capital	Employment quality
	Community	Social opportunities	Community relations
	Product responsibility	Product liability	Product quality and safety
	Human rights		Human rights
		Stakeholder opposition	
			Diversity
			Occupational health and safety
			Training and development
			Product access
			Labor rights
			Compensation
Corporate governance	Management	Corporate governance	Corporate governance
	Shareholders		
	CSR strategy		
		Corporate behavior	Business ethics

Here, the Bloomberg ESG database is omitted, since there are no ESG categorical topics.

almost all of them reported negative trends, except one study that reported a positive trend and another study that did not find statistical significance. This trend is also similar in the case of studies that used Bloomberg, Refinitiv, and Thomson Reuters Asset4 as ESG metrics, which generally found a negative direction, except one study that reported a mixed trend. Therefore, regardless of the type of ESG metrics employed, most recent studies have reported a negative effect, implying that engaging in ESG activities leads to robust and favorable results. If other conditions, such as free cash flow, are constant, the lower the value of the cost of capital, the greater the firm's value. Therefore, the robust trends observed in Tables 1B,D can be interpreted as an improvement in the firm's value assessment as a result of risk reduction due to ESG management.

Regarding social equity studies, based on articles selected from 2013 to 2021, we review the investigated issues, critical factors, geographical regions, and research fields, as presented in Table 3. Compared with qualitative theoretical analysis, quantitative analysis of social equity is a relatively new research area. The reviewed articles mostly appeared in the last 5 years. Social equity issues have been discussed worldwide, and quantitative analysis has been applied to a number of case studies in both developed and developing countries and regions.

In terms of the research field, social equity issues in transportation are the most studied topics. Accessibility of horizontal and vertical equity was used as an indicator to assess the extent to which residents can access the job market (El-Geneidy et al., 2016) and facilities (Yuan et al., 2017; Chen et al., 2018; Guo et al., 2020), as well as to discuss transportation design problems (Caggiani et al., 2017; Ruiz et al., 2017; Camporeale et al., 2019; Henke et al., 2020). Regarding transdisciplinary fields, social equity estimation is an essential part of the social sustainability index (Shaker and Sirodoev, 2016; Silva et al., 2018; Larimian and Sadeghi, 2021). Income gap has also been used as an indicator of social equity (Kangmennaang et al., 2017; Su et al., 2017). Regarding environmental issues, some studies have investigated the dissimilarity in costs or benefits and natural resources among different entities or protected areas (Halpern et al., 2013; Gurney et al., 2015) and constructed a social equity score that integrates energy issues (Chapman et al., 2021). Section Social equity will discuss the critical factors used in social equity evaluation in detail.

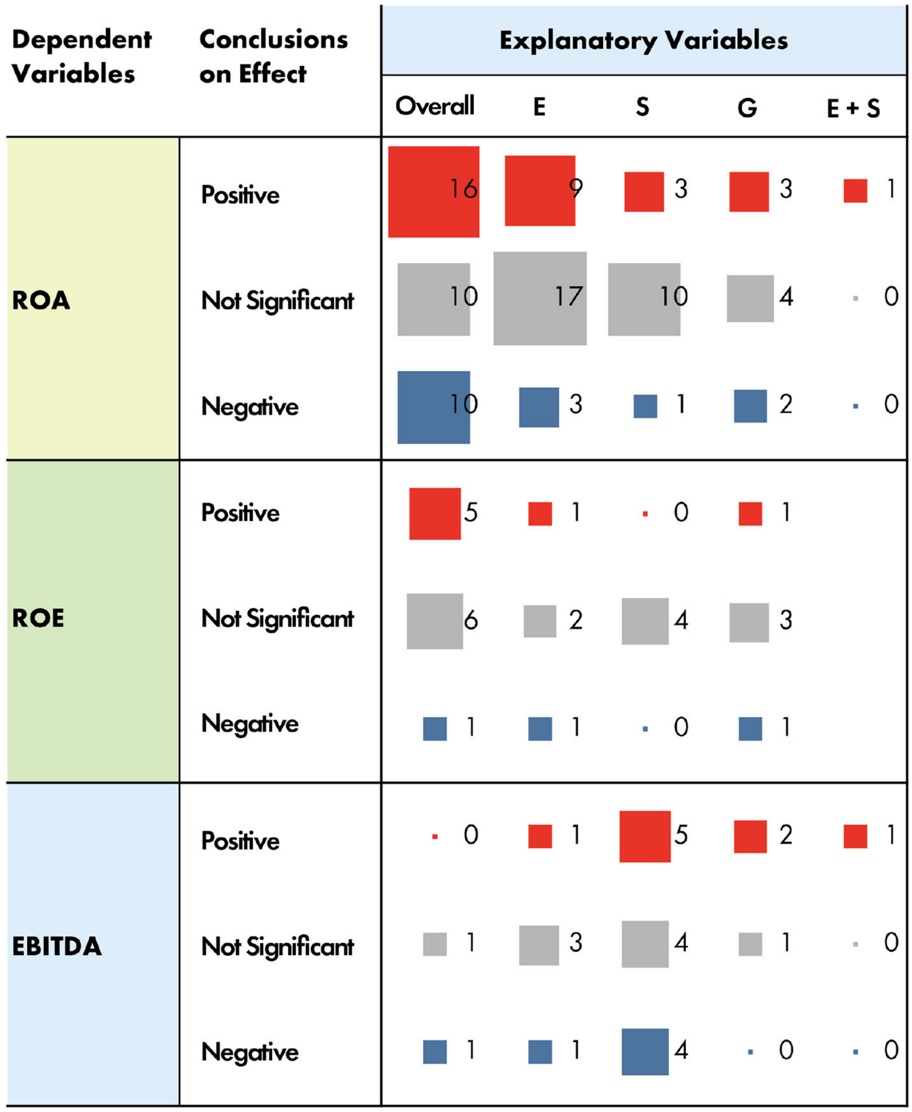


FIGURE 5
Heatmap of the results of the studies with ROA, ROE, EBITDA as dependent variables.

ESG metrics and social equity: A closer look at the methodology and elements assessed

This section first provides the details of the ESG metrics, with a brief background, methodology, and composition of the elements assessed. Then, the details of social equity evaluation studies are presented using the same procedure.

ESG metrics

Corporate sustainability reporting and rating, which are expected to impact individual corporations' behavior, surged in the last two decades (Scalet and Kelly, 2010). However, compared with financial reporting, which has a long history of evolution, it is still in its infancy (Tschopp and Huefner, 2015). Marlin and Marlin (2003) noted that the first phase of the corporate sustainability report in the 1970s and 1980s only focused on environmental management. Since then, CSR reporting has developed to involve multiple stakeholders and provide verifiable materials from the social auditor (the second phase is the 1990s) and has met third-party global reporting standards (the third phase is the 2000s) (Marlin and Marlin, 2003). Since then, various corporate sustainability reporting tools, such as frameworks (principles, initiatives, or guidelines) and standards, have been widely applied to evaluate corporations' efforts to achieve sustainability (Siew, 2015). Many

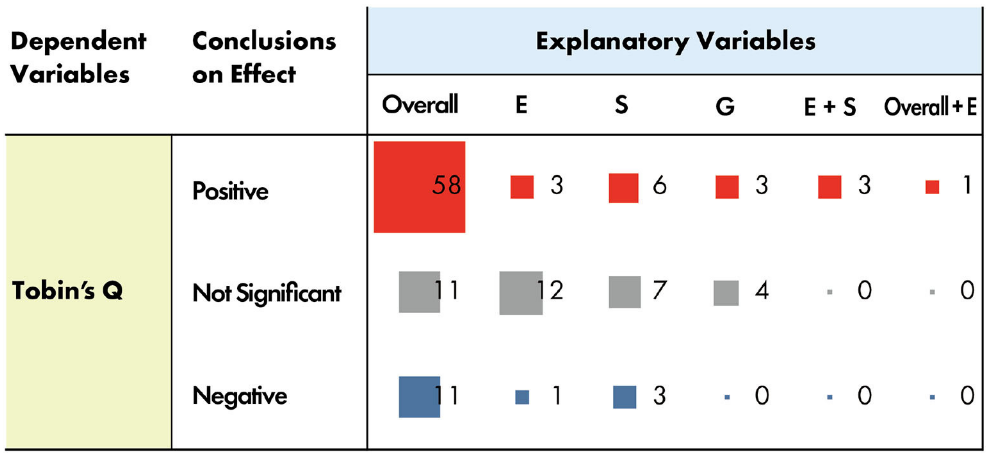


FIGURE 6
Heatmap of the results of the studies with Tobin's Q as dependent variables.

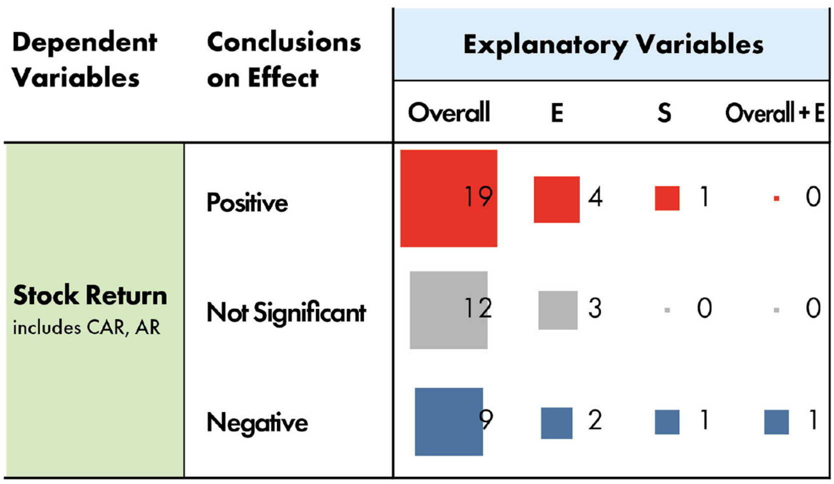


FIGURE 7
Heatmap of the results of the studies with Stock Return as dependent variables.

ESG rating agencies assess corporate sustainability based on the disclosed CSR reports and provide rating reports for multiple stakeholders. In the last 10 years, new criteria have been added to the assessment models, remarkably enhancing the accuracy and robustness of ESG ratings (Escrig-Olmedo et al., 2019).

Many studies have critiqued the low convergence of ESG ratings and called for being cautious about drawing conclusions based on these ratings (Siew, 2015; Chatterji et al., 2016; Berg et al., 2022). The main problems are the lack of common theorization (different definitions of good CSR) and commensurability (different measurements) (Chatterji et al., 2016). The scope, measurement, and weights contribute to the divergence of ESG ratings (Berg et al., 2022). Moreover, only a few ESG rating agencies disclose the details of their evaluating criteria and methods, leaving a black box in the ratings. Therefore, a universal ESG accounting standard with "dynamic materiality" is needed (Eccles and Mirchandani, 2022). Based on the accessible information about the rating methods and the elements assessed, we investigate four leading ESG databases widely used in academic research, investment, and business—Thomson Reuters' Refinitiv, MSCI, Bloomberg, and Arabesque S-Ray.

We first looked at how the ESG rating results correlate across the four databases and discuss the similarities and differences in the methodology and elements assessed in detail. Figure 8 depicts the distribution of ESG scores in each database and the correlation of scores based on the dataset in 2019. The ESG

TABLE 3 Summary of methodology across ESG metrics.

	Refinitiv	Bloomberg	MSCI	S-ray
Score range	0–100	0–100	0–10	0–100
Grade range	D- to A+ (12 grades)		CCC to AAA (7 grades)	
Other measures	Controversies		Controversies	Global compact score Preferences filter
Data sources	Company disclosure Media sources	Company disclosure	Company disclosure Media sources Specialized datasets (government databases, NGO, academic, etc.)	Company disclosure Media sources NGO
Coverage*	Around 9,000 firms	Around 12,000 firms	Around 8,500 firms	Around 8,000 firms
Update frequency	Monthly	Yearly	Monthly	Daily
Object of the Evaluation	Disclosure and performance	Disclosure	Performance (management capability) given both risks and opportunities	Performance given long and short-term risks and opportunities
Rating method	Full data-driven evaluation	Disclosure-based evaluation	Analysts' review	Semi data-driven evaluation and human oversight
Weight calculation	Data-based inner- and inter-industry adjustment	Industry adjustment	Industry adjustment Risk and opportunity exposure adjustment	Static review and data-based adjustment (sector- and industry-level, equal- and market cap-weighted monthly index returns)
Industry classification	TRBC	GICS	GICS	FactSet definition

* The number of firms assessed was counted in 2019.

scores and the scores of the components (E, S, and G) in the MSCI database have similar distributions, which are close to a normal distribution. However, the distribution in Bloomberg's ESG database varies, with a higher average G score and a lower average E score. As for Refinitiv's ESG ratings, the G score has a right-skewed distribution, whereas the others skew to the left to different extents. The scores of ESG components in S-Ray have right-skewed distributions. As noted in previous studies, the four investigated databases have low correlations. Most of the correlation coefficients are <0.5, ranging from −0.012 to 0.670. The correlations of integrated ESG scores range from 0.318 (MSCI and Bloomberg) to 0.549 (Refinitiv and Bloomberg). Regarding the ratings of ESG components, it is hard to find strong correlations between these databases. Regarding the E and S scores, MSCI ratings have the lowest correlation with the other three databases. Compared with those of E and S scores, the correlations of G scores are weaker and even insignificant between MSCI and Bloomberg, S-Ray and Reginitiv, as well as S-Ray and Bloomberg. We assume that the inconsistency between these ratings is due to the different methodologies and elements assessed, which will be discussed in the following parts.

In Table 2, we summarize the methodology of the four ESG databases. The final ESG ratings range from 0 to 100 points, except for MSCI, which uses 10 points as the maximum. In addition to ESG scores, Refinitiv and MSCI also provide concise and explicit ESG grade evaluations. The assessments are usually based on information individual firms disclose. Except Bloomberg, the other raters utilize media sources to construct controversies to adjust the final ESG ratings. All the four databases have a global coverage of at least 8,000 firms. Bloomberg's ESG scores are updated annually. Refinitiv and MSCI's ESG scores are updated monthly, and S-Ray's ESG score is updated daily.

Here, we follow the framework in previous studies to discuss the purpose of the evaluation (the scopes) and the rating procedure (the measures), including the method and weight (Chatterji et al., 2016; Berg et al., 2022). The purpose of the evaluation reflects how the rater defines good CSR, i.e., the theorization or scope of the assessment. In the four databases, there are mainly two directions in determining what is good CSR—information transparency and CSR performance. Bloomberg's disclosure score treats the transparency of ESG information as the most vital factor of CSR. Thus, for Bloomberg's ESG scores, higher information disclosure leads to higher rating results, without accounting for performance. Both MSCI and S-Ray's ESG ratings aim to assess performance in terms of ESG issues but from different perspectives. MSCI's ESG scores tend to evaluate the management's capability in handling both risks and opportunities. S-Ray's ESG scores account for

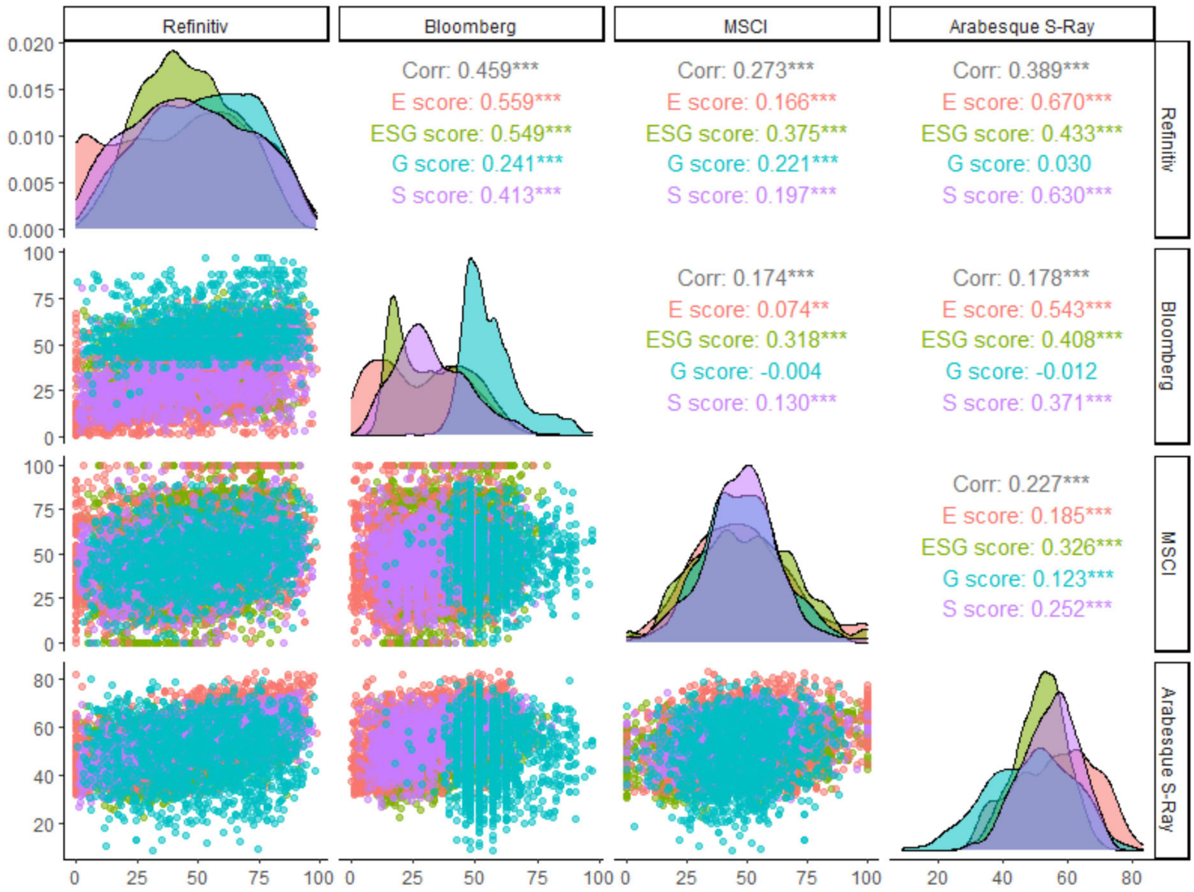


FIGURE 8
Correlation between ESG scores across different databases. This figure is based on the four databases in 2019.

performance, considering both long-term and short-term risks and opportunities, and the evaluation is conducted daily. Refinitiv's ESG scores evaluate both disclosure rate and relative performance among peers.

Regarding the rating process, there is no fully disclosed methodology information in these databases. Based on the accessible materials on methodology, we summarize some of the features as follows. Refinitiv has the advantage of a clear and verifiable method, as the assessment is entirely data-driven without any human intervention. Bloomberg's ESG scores only consider the degree of information disclosure, which makes the rating easy to understand and straightforward. However, MSCI's ESG scores reflect a more subjective assessment by specialists and analysts, which involves highly professional views, but the assessment is not easily understandable by users. S-Ray's ratings aim to seek a balance by combining a semi data-driven evaluation and human intervention. Regarding the weights, Refinitiv's ratings conduct inter- and intra-industry adjustment, which is a fully data-driven process that takes time. The industry classification is based on The Refinitiv Business Classifications. MSCI and Bloomberg's ratings are adjusted based on the Global Industry Classification Standard. Notably, MSCI's ratings are adjusted for risk or opportunities in each element assessed. Thus, the ratings in MSCI are not only a relative peer comparison but also an evaluation of firms' management capability in handling potential risks and opportunities.

Before comparing the detailed elements assessed, it is essential to note that the structures of ESG scores also vary across each database, as presented in Table 4. The main difference is categorizing the elements under the E, S, and G pillars. In the E pillar, the common categories are resource, emission, and innovation. Although it is given different names, all the three databases have these categories. In addition, pollution and waste are also evaluated in MSCI and S-Ray's ESG database. S-Ray also provides "environmental stewardship" and "environmental management" scores. In the S pillar, categories of human resource, community, and product responsibility are common among the three databases. Refinitiv and S-Ray have a category of "human rights," whereas MSCI provides another category of "stakeholder opposition." S-Ray's ESG scores provide detailed

TABLE 4 List of quantitative studies of social equity.

References	Indicators out of socioeconomic factors	Social (vulnerability) indicator	Sector or research field	Target area	Country or regions
Valizadeh and Hayati (2021)		Education Economic factors Right to quality of life Capacity development Fair pricing and TRN contracts Employment relations Child labor Non-discrimination and sup. Vuln. People Health cover, access, medic care	Agricultural and biological sciences (miscellaneous)	Fars province	Iran
Larimian and Sadeghi (2021)		Access to essential facilities Access to recreational facilities Access to educational facilities Access to transportation facilities	Urban studies	Dunedin city	New Zealand
A. Chapman et al. (2021)	Ratio of renewable energy to the total electricity Electricity access PM 2.5 explosure Environmental improvement indicator Energy poverty indicator	Income distribution GDP per capita Unemployment ratio	Renewable energy, sustainability and the environment		99 countries
Emrich et al. (2020)		Housing tenure Fianncial capital Race Language proficiency Housing quality Age Employment	Management, monitoring, policy and law	South Carolina floods 2015	United States
Henke et al. (2020)	Travel time	Total number of employees in traffic zone	Transportation	Puglia	Italy
Karakoc et al. (2020)		Population over the age of 65 Population under the age of 5 Population that is Hispanic Single-female based households Households that are in poverty	Urban studies	Shelby County	United States
Bennett et al. (2020)		Recognitional equity (4 items as below) (Rights Livelihoods	Nature and landscape conservation	6 countries on the Mediterranean Sea	6 countries on the Mediterranean Sea

(Continued)

TABLE 4 (Continued)

References	Indicators out of socioeconomic factors	Social (vulnerability) indicator	Sector or research field	Target area	Country or regions
		Traditional Knowledge Culture) Procidural equity (8 items) Distributional equity (8 items)			
Shigetomi et al. (2020)	GHG emissions Primary PM 2.5 Blue and green water mining risk for neodymium Industrial waste	Household bracket based on cumulative share of consumption	Renewable energy, sustainability and the environment		Japan
Guo et al. (2020)	Park accessibility	Elderly population	Transportation	Harbin city	China
Camporeale et al. (2019)	Number of bus trips	Unemployed population Young (<19 years old) Old (more than 65 years old)	Transportation	Molfetta	Italy
Chen et al. (2018)	Service area ratio Service density Service frequency Route diversity Accessibility within a statistical area Accessibility across statistical area	Percentage of senior population	Transportation	Edmonton	Canada
Silva et al. (2018)		Poverty Households with income below poverty line (%) Population living in extreme poverty (%) Average monthly income (ln) Gender Equality Ratio between average wages for women and men	Social sciences (miscellaneous)	State of Ceara	Brazil
Su et al. (2017)		Urban-rural income gap	Urban studies	Megaregion around Hangzhou Bay	China
Ruiz et al. (2017)	Bus Service Level by districts	Population by districts Dependent population rate Immigrant population rate Female population rate Level of economic activity	Transportation	Palma	Spain
Kangmennaang et al. (2017)	Firm pay gap		Agricultural and biological sciences (miscellaneous)		China

(Continued)

10.3389/frsus.2022.920955

10.3389/frsus.2022.920955

TABLE 4 (Continued)

References	Indicators out of socioeconomic factors	Social (vulnerability) indicator	Sector or research field	Target area	Country or regions
Yuan et al. (2017)	Public parks accessibility	Total population Rate of the Female population Rate of the population aged 0–19 Rate of the population aged 60 and over Rate of ethnic minority population Rate of the illiterate population Rate of the laid-off population Rate of the unemployed population	Transportation	Changting	China
Caggiani et al. (2017)		Residing population Workers Number of employees Residing disadvantaged population Young Unemployed population Low-income population	Transportation	Molfetta	Italy
El-Geneidy et al. (2016)	Travel time Transit fares	Household income Percentage of recent immigrants (since 2006) Percentage of workforce that is unemployed Percentage of residents with education at the level of only a high school diploma (25–64 years old)	Transportation	Montreal	Canada
Shaker and Sirodoev (2016)	Type of cooking fuel Computers, mobile phone, microwave, and DVD/VCR Access to improved water source Access and type of sanitation facility Number of household members outside the country Head of household education level		Social sciences (miscellaneous)		Moldova

(Continued)

TABLE 4 (Continued)

References	Indicators out of socioeconomic factors	Social (vulnerability) indicator	Sector or research field	Target area	Country or regions
Oswald Beiler and Mohammed (2016)	Public transit Access School proximity Network connectivity Mixed land uses Flood hazard Crash rates Truck volume Intermodal facilities	Race Limited English proficiency Age Disability Economic development Vehicles per household Household income Single parent household Cost of living Travel time	Transportation	Sullivan County	United States
Gurney et al. (2015)	Inverse of the Gini coefficient in terms of the percentage of retained catch per unit effort (CPUE)		Nature and landscape conservation	Kubulau	Fiji
Farber et al. (2014)	Ridership percentage Trip generations Distance traveled (miles)	Household income Hispanic Race Age Employment Education Licensed Limited mobility Home ownership Years of residence Place type Residence type	Transportation	Wasatch Front, Utah	United States
Di Ciommo and Lucas (2014)	Travel times Transport costs	Income	Transportation	Madrid	Spain
Halpern et al. (2013)	Fraction of fishing value lost inside marine reserves		Nature and landscape conservation	California	United States
	Fraction of community fishing grounds lost inside marine reserves		Nature and landscape conservation	Misool, Raja Ampat	Indonesia
	Fraction of money spent; fraction of area placed into marine reserves		Nature and landscape conservation	Coral Triangle	Southeast Asia

10.3389/frsus.2022.920955

topics in the S pillar, including diversity, occupational health and safety, training and development, product access, labor rights, and compensation. In terms of the G pillar, a similar category among the three databases is corporate governance or management. Moreover, categories of shareholders, CSR strategy, corporate behavior, and business ethics are provided across the databases.

Based on accessible methodology materials of the four ESG ratings, we collect the elements assessed in each ESG database. The total number of elements assessed in all the four databases is 842. The elements assessed have a significant divergence across the four databases. The Venn diagram of Figure 9 compares all the elements assessed, in which only four elements are common among all the four ESG ratings. The ratios of exclusive elements are 37.3% in Refinitiv's ESG scores, 38.1% in MSCI's ESG scores, 4.4% in S-Ray's ESG scores, and 7.1% in Bloomberg's ESG scores. Regarding the social aspect, there are 281 total elements across the four databases, which is 33.4% of all the ESG elements. Surprisingly, there are no common items in all the databases. The number of social elements in Bloomberg and S-Ray is much less than that in Refinitiv and MSCI. The observed significant divergence in the assessed elements across the four databases emphasizes the importance of developing a universal ESG accounting standard with "dynamic materiality", elaborated by Eccles and Mirchandani (2022).

Social equity

Here, when we refer to social equity, we are referring to a metric used to evaluate the equitability of various energy and environment issues, as well as the social aspects of sustainability. Energy-related social equity has its roots in the energy justice movement, which is based on environmental justice (Pettit, 2004) and climate justice (Bulkeley et al., 2013), focusing on energy issues and environmental benefits (Jenkins, 2018). Energy justice focuses on three key tenets—distributional justice (the distribution of costs and benefits), recognition justice (identifying who benefits or is burdened), and procedural justice [open access and engagement in policy decision-making processes (Jenkins et al., 2016)]. Social equity evaluations have been used for energy policy, energy emissions, energy law, energy finance, climate policy, and, most recently, energy transitions (McCauley and Heffron, 2018; Chapman et al., 2019).

Regarding energy-related sustainability evaluations, most of the studies focused on the more easily quantified environmental and economic aspects, or when considering social equity, they place more emphasis on qualitative rather than quantitative factors (Evans et al., 2009; Chapman et al., 2016).

In studies that evaluate sustainability and social equity quantitatively, a number of common factors come to the fore. Among them is the concept of employment, including relations (Valizadeh and Hayati, 2021), unemployment ratios and age groups (Farber et al., 2014; El-Geneidy et al., 2016; Caggiani et al., 2017; Yuan et al., 2017; Camporeale et al., 2019; Emrich et al., 2020; Chapman et al., 2021), and how different policies and technological shift affect employment outcomes. Moreover, income and education, which are often closely correlated, are also considered important in the literature (Farber et al., 2014; Oswald Beiler and Mohammed, 2016; Silva et al., 2018; Shigetomi et al., 2020; Chapman et al., 2021). Access is another factor of social equity that can be quantified, including access to energy, transportation, and essential facilities (Farber et al., 2014; Shigetomi et al., 2020; Chapman et al., 2021; Larimian and Sadeghi, 2021). Furthermore, recognizing that there is often a gap not only between nations but also within nations, some studies also investigate the urban–rural divide and the gaps between genders and different age groups (Farber et al., 2014; Oswald Beiler and Mohammed, 2016; Ruiz et al., 2017; Su et al., 2017; Yuan et al., 2017; Chen et al., 2018; Silva et al., 2018; Guo et al., 2020; Karakoc et al., 2020). Generally, the literature includes distributive, recognition, and procedural aspects of social equity evaluations, with some focusing on this aspect quantitatively (Bennett et al., 2020). As the concept of social equity is still in its nascent phase, the overlap of concepts is not consistent, and to enable quantitative evaluations of social aspects, economic and environmental factors are co-opted when considered socially important.

From a regional perspective, Europe is strongly represented in the literature along with the United States and with some case studies on Southeast Asia. Global studies are only beginning to emerge in the most recent literature, which is largely due to the relatively recent emergence of concepts and data limitations.

Discussion and conclusion

As the ESG market is expanding rapidly, with total global ESG investments of $35.3 trillion in 2020, ESG rating providers play an increasingly important role in the investment process through their assessments of companies across various ESG metrics. However, the lack of common theorization (different definitions of good CSR) and commensurability (different measurements), which has been pointed out in various studies (Chatterji et al., 2016) and examined in detail in this paper, highlights the improvements required in the field of ESG assessment to provide clear and transparent information to investors and to reduce confusion among companies that are trying to enhance their ESG performance.

Considering the effect of ESG factors on corporate performance, in summary, we find that the overall trend of the short-term effect on profitability is unclear. The effect on ROA or ROE is still far from conclusive. In terms of stock return, the results vary, as they utilize different ESG metrics. Multiple factors, such as samples from different markets and periods, could also be a reason for the inconsistent results.

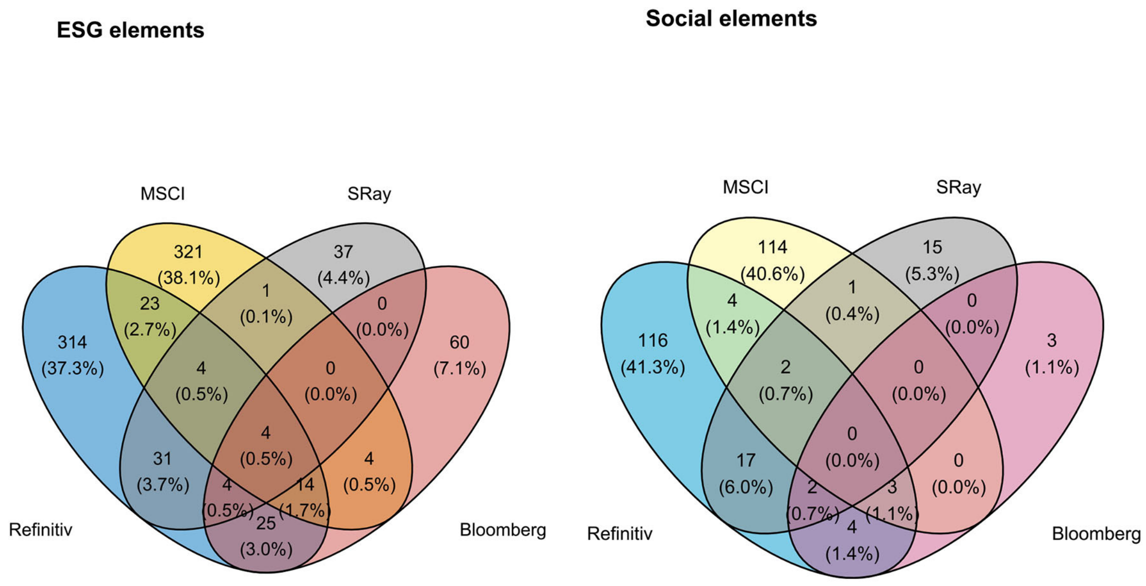


FIGURE 9
Comparison of evaluated factors across ESG database.

However, most of the current ESG evaluations do not reflect the financial impact, accounting measures, and short-term market returns sufficiently. The trend is robust and favorable when testing the effect on Tobin's Q, cost of equity, or other risks, which indicate the nature of ESG activities, thereby enhancing corporate sustainability in the long term. Although it is out of scope of this paper, future research could focus on the difference between ESG metrics and conduct an in-depth analysis of the metrics that impact upon financial outcomes.

To have a closer look at the ESG metrics, we first investigate how the results of the ESG ratings correlate across the four widely used databases (Thomson Reuters' Refinitiv, MSCI, Bloomberg, and Arabesque S-Ray). The results reveal that the four investigated databases have low correlations, with the correlations of integrated ESG scores ranging from 0.318 (MSCI and Bloomberg) to 0.549 (Refinitiv and Bloomberg). Moreover, regarding the ratings of the ESG components, it is also difficult to find strong correlations between these databases. Based on the accessible methodology materials of the four ESG ratings, we also collect the elements assessed in each ESG database and present the significant divergence of the elements assessed across the four databases. The ratios of exclusive elements are 37.3% in Refinitiv's ESG scores, 38.1% in MSCI's ESG scores, 4.4% in S-Ray's ESG scores, and 7.1% in Bloomberg's ESG scores. Regarding the social aspect, there are 281 elements in all the four databases, which is 33.4% of all the ESG elements. There are no common items in all the databases. The number of social elements in Bloomberg and S-Ray is much lower than that of Refinitiv and MSCI. Although the ESG metrics and the investment market are evolving rapidly, with investors, corporations, and the public giving more priority to the "S" in ESG, which includes social equity issues, such as diversity, income inequality, workers' safety, systemic racism, and companies' broader role in society. There is significant divergence among the different ESG databases in the elements assessed under the social category.

To provide a suitable yardstick for the assessment of social aspects, we investigated existing approaches used for social equity evaluations through a systematic review and closely examined the key elements assessed in these studies. Some of the common factors that we find in the studies that evaluated sustainability and social equity quantitatively are the concept of employment, such as relations, unemployment ratios and age groups, as well as income and education, which were also found to be important elements in ESG metrics (e.g., gender balance, salary, and training). In social equity studies, access is a factor which is considered important and which can be quantified, including access to energy, transportation, and essential facilities, whereas quantifying access is rarely observed in the major ESG metrics. Due to the influence of ESG metrics, the differences in the rating methodologies and the level of transparency in the rating decisions, which also incorporate qualitative judgments, are critical to understanding the resilience of the ESG financial intermediation chain. The results of this paper contribute to advancing the research community's and practitioners' knowledge by providing a detailed examination of commensurability of major ESG metrics, and whether these ESG metrics capture critical social elements. Furthermore, the results of this paper reveals the importance of promoting the transparency and comparability of scoring methodologies of established ESG rating providers and indices, as well as highlighting the importance of investigating studies and practices that quantitatively assess sustainability and social equity issues to ensure the overall veracity and quality of ESG metrics, as well as providing some evidence for their future expansion and improvement.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

AK, AC, KY, JX, JI, and ST carried out the analyses and wrote the manuscript with support from SM. AK and SM supervised the project. All authors contributed to the article and approved the submitted version.

Funding

This research is supported by JSPS KAKENHI Grant Number JP20H00648 and the Environment Research and Technology Development Fund (JPMEERF20201001) of the Environmental Restoration and Conservation Agency of Japan.

Acknowledgments

We would like to thank Mr. Okita and Mr. Ike from Vector Group for providing their insights and expertise on ESG investment market during number of discussions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

References

Alareeni, B. A., and Hamdan, A. (2020). ESG impact on performance of US S&P 500-listed firms. *Corp. Governance Int. J. Bus. Soc.* 20, 1409–1428. doi: 10.1108/CG-06-2020-0258

Albarrak, M. S., Elnahass, M., and Salama, A. (2019). The effect of carbon dissemination on cost of equity. *Bus. Strategy Environ.* 28, 1179–1198. doi: 10.1002/bse.2310

Albuquerque, R., Koskinen, Y., and Zhang, C. (2019). Corporate social responsibility and firm risk: theory and empirical evidence. *Manage. Sci.* 65, 4451–4469. doi: 10.1287/mnsc.2018.3043

Atif, M., Hossain, M., Alam, M. S., and Goergen, M. (2021). Does board gender diversity affect renewable energy consumption? *J. Corp. Financ.* 66, 101665. doi: 10.1016/j.jcorpfin.2020.101665

Avramov, D., Cheng, S., Lioui, A., and Tarelli, A. (2021). Sustainable investing with ESG rating uncertainty. *J. Financ. Econ.* 145, 642–664 doi: 10.1016/j.jfineco.2021.09.009

Bae, K.-H., El Ghoul, S., Gong, Z., and Guedhami, O. (2021). Does CSR matter in times of crisis? Evidence from the COVID-19 pandemic. *J. Corp. Financ.* 67, 101876. doi: 10.1016/j.jcorpfin.2020.101876

Bardos, K. S., Ertugrul, M., and Gao, L. S. (2020). Corporate social responsibility, product market perception, and firm value. *J. Corp. Financ.* 62, 101588. doi: 10.1016/j.jcorpfin.2020.101588

Bátae, O. M., Dragomir, V. D., and Feleagá, L. (2021). The relationship between environmental, social, and financial performance in the banking sector: a European study. *J. Clean. Prod.* 290, 125791. doi: 10.1016/j.jclepro.2021.125791

Becchetti, L., Ciciretti, R., and Hasan, I. (2015). Corporate social responsibility, stakeholder risk, and idiosyncratic volatility. *J. Corp. Financ.* 35, 297–309. doi: 10.1016/j.jcorpfin.2015.09.007

Bennett, N. J., Calò, A., Di Franco, A., Niccolini, F., Marzo, D., Domina, I., et al. (2020). Social equity and marine protected areas: perceptions of small-scale fishermen in the Mediterranean Sea. *Biol. Conserv.* 244, 108531. doi: 10.1016/j.biocon.2020.108531

Berg, F., Koelbel, J., and Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance.* 2022:rfac033. doi: 10.1093/rof/rfac033

Bhandari, A., and Javakhadze, D. (2017). Corporate social responsibility and capital allocation efficiency. *J. Corp. Financ.* 43, 354–377. doi: 10.1016/j.jcorpfin.2017.01.012

Boesso, G., Favotto, F., and Michelon, G. (2015). Stakeholder prioritization, strategic corporate social responsibility and company performance: further evidence. *Corp. Soc. Respons. Environ. Manage.* 22, 424–440. doi: 10.1002/csr.1356

Bolton, P., and Kacperczyk, M. (2021). Do investors care about carbon risk? *J. Financ. Econ.* 142, 517–549. doi: 10.1016/j.jfineco.2021.05.008

Boone, A., and Uysal, V. B. (2020). Reputational concerns in the market for corporate control. *J. Corp. Financ.* 61, 101399. doi: 10.1016/j.jcorpfin.2018.08.010

Borghesi, R., Houston, J. F., and Naranjo, A. (2014). Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests. *J. Corp. Financ.* 26, 164–181. doi: 10.1016/j.jcorpfin.2014.03.008

Bose, S., Minnick, K., and Shams, S. (2021). Does carbon risk matter for corporate acquisition decisions? *J. Corp. Financ.* 70, 102058. doi: 10.1016/j.jcorpfin.2021.102058

Boubakri, N., Guedhami, O., Kwok, C. C. Y., and Wang, H. (2019). Is privatization a socially responsible reform? *J. Corp. Financ.* 56, 129–151. doi: 10.1016/j.jcorpfin.2018.12.005

Breuer, W., Müller, T., Rosenbach, D., and Salzmann, A. (2018). Corporate social responsibility, investor protection, and cost of equity: a cross-country comparison. *J. Bank. Financ.* 96, 34–55. doi: 10.1016/j.jbankfin.2018.07.018

Brogi, M., and Lagasio, V. (2019). Environmental, social, and governance and company profitability: Are financial intermediaries different? *Corp. Soc. Respons. Environ. Manage.* 26, 576–587. doi: 10.1002/csr.1704

Brower, J., and Dacin, P. A. (2020). An institutional theory approach to the evolution of the corporate social performance – corporate financial performance relationship. *J. Manage. Stud.* 57, 805–836. doi: 10.1111/joms.12550

Bu, L., Chan, K. C., Choi, A., and Zhou, G. (2021). Talented inside directors and corporate social responsibility: a tale of two roles. *J. Corp. Financ.* 70, 102044. doi: 10.1016/j.jcorpfin.2021.102044

Buchanan, B., Cao, C. X., and Chen, C. (2018). Corporate social responsibility, firm value, and influential institutional ownership. *J. Corp. Financ.* 52, 73–95. doi: 10.1016/j.jcorpfin.2018.07.004

Bulkeley, H., Carmin, J. A., Castán Broto, V., Edwards, G. A. S., and Fuller, S. (2013). Climate justice and global cities: mapping the emerging discourses. *Global Environ. Change* 23, 914–925. doi: 10.1016/j.gloenvcha.2013.05.010

Byun, S. K., and Oh, J.-M. (2018). Local corporate social responsibility, media coverage, and shareholder value. *J. Bank. Financ.* 87, 68–86. doi: 10.1016/j.jbankfin.2017.09.010

Caggiani, L., Camporeale, R., and Ottomanelli, M. (2017). Planning and design of equitable free-floating bike-sharing systems implementing a road pricing strategy. *J. Adv. Transport.* 2017, 1–18. doi: 10.1155/2017/3182387

Cahan, S. F., Chen, C., Chen, L., and Nguyen, N. H. (2015). Corporate social responsibility and media coverage. *J. Bank. Financ.* 59, 409–422. doi: 10.1016/j.jbankfin.2015.07.004

Cai, L., Cui, J., and Jo, H. (2016). Corporate environmental responsibility and firm risk. *J. Bus. Ethics* 139, 563–594. doi: 10.1007/s10551-015-2630-4

Cai, X., Gao, N., Garrett, I., and Xu, Y. (2020). Are CEOs judged on their companies' social reputation? *J. Corp. Financ.* 64, 101621. doi: 10.1016/j.jcorpfin.2020.101621

Camilleri, M. A. (2020). The market for socially responsible investing: a review of the developments. *Soc. Respons. J.* 17, 412–428. doi: 10.1108/SRJ-06-2019-0194

Camporeale, R., Caggiani, L., and Ottomanelli, M. (2019). Modeling horizontal and vertical equity in the public transport design problem: a case study. *Transport. Res. Part A Policy Pract.* 125, 184–206. doi: 10.1016/j.tra.2018.04.006

Chang, C.-H., Chen, S.-S., Chen, Y.-S., and Peng, S.-C. (2019). Commitment to build trust by socially responsible firms: evidence from cash holdings. *J. Corp. Financ.* 56, 364–387. doi: 10.1016/j.jcorpfin.2019.03.004

Chapman, A., Fraser, T., and Dennis, M. (2019). Investigating ties between energy policy and social equity research: a citation network analysis. *Soc. Sci.* 8, 135. doi: 10.3390/socsci8050135

Chapman, A., Shigetomi, Y., Ohno, H., McLellan, B., and Shinozaki, A. (2021). Evaluating the global impact of low-carbon energy transitions on social equity. *Environ. Innovat. Societal Trans.* 40, 332–347. doi: 10.1016/j.eist.2021.09.002

Chapman, A. J., McLellan, B., and Tezuka, T. (2016). Proposing an evaluation framework for energy policy making incorporating equity: applications in Australia. *Energy Res. Soc. Sci.* 21, 54–69. doi: 10.1016/j.erss.2016.06.021

Chatterji, A. K., Durand, R., Levine, D. I., and Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Manage. J.* 37, 1597–1614. doi: 10.1002/smj.2407

Chen, Y, Bouferguene, A., Li, H. X., Liu, H., Shen, Y., and Al-Hussein, M. (2018). Spatial gaps in urban public transport supply and demand from the perspective of sustainability. *J. Clean. Prod.* 195, 1237–1248. doi: 10.1016/j.jclepro.2018.06.021

Chen, Y., Fan, Q., Yang, X., and Zolotoy, L. (2021). CEO early-life disaster experience and stock price crash risk. *J. Corp. Financ.* 68, 101928. doi: 10.1016/j.jcorpfin.2021.101928

Cheung, A. (2016). Corporate social responsibility and corporate cash holdings. *J. Corp. Financ.* 37, 412–430. doi: 10.1016/j.jcorpfin.2016.01.008

Choy, S. K., Lai, T.-K., and Ng, T. (2017). Do tax havens create firm value? *J. Corp. Financ.* 42, 198–220. doi: 10.1016/j.jcorpfin.2016.10.016

Cornett, M. M., Erhemjamts, O., and Tehranian, H. (2016). Greed or good deeds: An examination of the relation between corporate social responsibility and the financial performance of U.S. commercial banks around the financial crisis. *J. Bank. Financ.* 70, 137–159. doi: 10.1016/j.jbankfin.2016.04.024

Dai, R., Liang, H., and Ng, L. (2021). Socially responsible corporate customers. *J. Financ. Econ.* 142, 598–626. doi: 10.1016/j.jfineco.2020.01.003

Devie, D., Liman, L. P., Tarigan, J., and Jie, F. (2020). Corporate social responsibility, financial performance and risk in Indonesian natural resources industry. *Soc. Respons. J.* 16, 73–90. doi: 10.1108/SRJ-06-2018-0155

Di Ciommo, F., and Lucas, K. (2014). Evaluating the equity effects of road-pricing in the European urban context – The Madrid Metropolitan Area. *Appl. Geograp.* 54, 74–82. doi: 10.1016/j.apgeog.2014.07.015

Di Giuli, A., and Kostovetsky, L. (2014). Are red or blue companies more likely to go green? Politics and corporate social responsibility. *J. Financ. Econ.* 111, 158–180. doi: 10.1016/j.jfineco.2013.10.002

Ding, W., Levine, R., Lin, C., and Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *J. Financ. Econ.* 141, 802–830. doi: 10.1016/j.jfineco.2021.03.005

Doukas, J. A., and Zhang, R. (2021). Managerial ability, corporate social culture, and M&As. *J. Corp. Financ.* 68, 101942. doi: 10.1016/j.jcorpfin.2021.101942

Duque-Grisales, E., and Aguilera-Caracuel, J. (2021). Environmental, social and governance (ESG) scores and financial performance of multilatinas: moderating effects of geographic international diversification and financial slack. *J. Bus. Ethics* 168, 315–334. doi: 10.1007/s10551-019-04177-w

Dutordoir, M., Strong, N. C., and Sun, P. (2018). Corporate social responsibility and seasoned equity offerings. *J. Corp. Financ.* 50, 158–179. doi: 10.1016/j.jcorpfin.2018.03.005

Eccles, R. G., and Mirchandani, B. (2022). We Need Universal ESG Accounting Standards. *Harvard Business Review*. Available online at: https://hbr.org/2022/02/we-need-universal-esg-accounting-standards (accessed March 15, 2022).

El Ghoul, S., and Karoui, A. (2017). Does corporate social responsibility affect mutual fund performance and flows? *J. Bank. Financ.* 77, 53–63. doi: 10.1016/j.jbankfin.2016.10.009

El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., and Loong, C. (2016). The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transport. Res. Part A Policy Pract.* 91, 302–316. doi: 10.1016/j.tra.2016.07.003

Emrich, C. T., Tate, E., Larson, S. E., and Zhou, Y. (2020). Measuring social equity in flood recovery funding. *Environ. Hazards* 19, 228–250. doi: 10.1080/17477891.2019.1675578

Ertugrul, M., and Marciukaityte, D. (2021). Labor unions and corporate social responsibility. *J. Bank. Financ.* 125, 106061. doi: 10.1016/j.jbankfin.2021.106061

Escrig-Olmedo, E., Fernández-Izquierdo, M., Ferrero-Ferrero, I., Rivera-Lirio, J., and Muñoz-Torres, M. (2019). Rating the raters: evaluating how ESG rating agencies integrate sustainability principles. *Sustainability* 11, 915. doi: 10.3390/su11030915

Evans, A., Strezov, V., and Evans, T. J. (2009). Assessment of sustainability indicators for renewable energy technologies. *Renewable Sustain. Energy Rev.* 13, 1082–1088. doi: 10.1016/j.rser.2008.03.008

Farber, S., Bartholomew, K., Li, X., Páez, A., and Nurul Habib, K. M. (2014). Assessing social equity in distance based transit fares using a model of travel behavior. *Transport. Res. Part A Policy Pract.* 67, 291–303. doi: 10.1016/j.tra.2014.07.013

Fauver, L., McDonald, M. B., and Taboada, A. G. (2018). Does it pay to treat employees well? International evidence on the value of employee-friendly culture. *J. Corp. Financ.* 50, 84–108. doi: 10.1016/j.jcorpfin.2018.02.003

Feng, Z.-Y., Chen, C. R., and Tseng, Y.-J. (2018). Do capital markets value corporate social responsibility? Evidence from seasoned equity offerings. *J. Bank. Financ.* 94, 54–74. doi: 10.1016/j.jbankfin.2018.06.015

Ferrell, A., Liang, H., and Renneboog, L. (2016). Socially responsible firms. *J. Financ. Econ.* 122, 585–606. doi: 10.1016/j.jfineco.2015.12.003

Gao, L., and Zhang, J. H. (2015). Firms' earnings smoothing, corporate social responsibility, and valuation. *J. Corp. Financ.* 32, 108–127. doi: 10.1016/j.jcorpfin.2015.03.004

Garel, A., and Petit-Romec, A. (2021). Investor rewards to environmental responsibility: evidence from the COVID-19 crisis. *J. Corp. Financ.* 68, 101948. doi: 10.1016/j.jcorpfin.2021.101948

Global Sustainable Investment Alliance. (2021). *Global Sustainable Investment Review 2020*. Available online at: http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf

Griffin, D., Guedhami, O., Li, K., and Lu, G. (2021). National culture and the value implications of corporate environmental and social performance. *J. Corp. Financ.* 71, 102123. doi: 10.1016/j.jcorpfin.2021.102123

Guo, M., Liu, B., Tian, Y., and Xu, D. (2020). Equity to urban parks for elderly residents: perspectives of balance between supply and demand. *Int. J. Environ. Res. Public Health* 17, 8506. doi: 10.3390/ijerph17228506

Gurney, G. G., Pressey, R. L., Ban, N. C., Álvarez-Romero, J. G., Jupiter, S., and Adams, V. M. (2015). Efficient and equitable design of marine protected areas in Fiji through inclusion of stakeholder-specific objectives in conservation planning. *Conserv. Biol.* 29, 1378–1389. doi: 10.1111/cobi.12514

Halpern, B. S., Klein, C. J., Brown, C. J., Beger, M., Grantham, H. S., Mangubhai, S., et al. (2013). Achieving the triple bottom line in the face of inherent trade-offs among social equity, economic return, and conservation. *Proc. Nat. Acad. Sci.* 110, 6229–6234. doi: 10.1073/pnas.1217689110

Hannah, S. T., Sayari, N., Harris, F. H., de, B., and Cain, C. L. (2021). The direct and moderating effects of endogenous corporate social responsibility on firm valuation: theoretical and empirical evidence from the global financial crisis. *J. Manage. Stud.* 58, 421–456. doi: 10.1111/joms.12586

Harrison, J. S., and Berman, S. L. (2016). Corporate social performance and economic cycles. *J. Bus. Ethics* 138, 279–294. doi: 10.1007/s10551-015-2646-9

Hawn, O., and Ioannou, I. (2016). Mind the gap: the interplay between external and internal actions in the case of corporate social responsibility. *Strategic Manage. J.* 37, 2569–2588. doi: 10.1002/smj.2464

Henke, I., Carten,ì, A., Molitierno, C., and Errico, A. (2020). Decision-making in the transport sector: a sustainable evaluation method for road infrastructure. *Sustainability* 12, 764. doi: 10.3390/su12030764

Hoang, T., Przychodzen, W., Przychodzen, J., and Segbotangni, E. A. (2020). Does it pay to be green? A disaggregated analysis of U.S. firms with green patents. *Bus. Strategy Environ.* 29, 1331–1361. doi: 10.1002/bse.2437

Hoi, C. K., Wu, Q., and Zhang, H. (2018). Community social capital and corporate social responsibility. *J. Bus. Ethics* 152, 647–665. doi: 10.1007/s10551-016-3335-z

Jenkins, K. (2018). Setting energy justice apart from the crowd: lessons from environmental and climate justice. *Energy Res. Soc. Sci.* 39, 117–121. doi: 10.1016/j.erss.2017.11.015

Jenkins, K., McCauley, D., Heffron, R., Stephan, H., and Rehner, R. (2016). Energy justice: a conceptual review. *Energy Res. Soc. Sci.* 11, 174–182. doi: 10.1016/j.erss.2015.10.004

Jha, A., and Cox, J. (2015). Corporate social responsibility and social capital. *J. Bank. Financ.* 60, 252–270. doi: 10.1016/j.jbankfin.2015.08.003

Jia, J., and Li, Z. (2020). Does external uncertainty matter in corporate sustainability performance? *J. Corp. Financ.* 65, 101743. doi: 10.1016/j.jcorpfin.2020.101743

Kangmennaang, J., Kerr, R. B., Lupafya, E., Dakishoni, L., Katundu, M., and Luginaah, I. (2017). Impact of a participatory agroecological development project on household wealth and food security in Malawi. *Food Secur.* 9, 561–576. doi: 10.1007/s12571-017-0669-z

Karakoc, D. B., Barker, K., Zobel, C. W., and Almoghathawi, Y. (2020). Social vulnerability and equity perspectives on interdependent infrastructure network component importance. *Sustain. Cities Soc.* 57, 102072. doi: 10.1016/j.scs.2020.102072

Kim, Y., Li, H., and Li, S. (2014). Corporate social responsibility and stock price crash risk. *J. Bank. Financ.* 43, 1–13. doi: 10.1016/j.jbankfin.2014.02.013

Kong, Y., Antwi-Adjei, A., and Bawuah, J. (2020). A systematic review of the business case for corporate social responsibility and firm performance. *Corp. Soc. Respon. Environ. Manage.* 27, 444–454. doi: 10.1002/csr.1838

Kumar, K., Boesso, G., and Michelon, G. (2016). How do strengths and weaknesses in corporate social performance across different stakeholder domains affect company performance? *Bus. Strategy Environ.* 25, 277–292. doi: 10.1002/bse.1874

Kuzey, C., Uyar, A., Nizaeva, M., and Karaman, A. S. (2021). CSR performance and firm performance in the tourism, healthcare, and financial sectors: do metrics and CSR committees matter? *J. Clean. Prod.* 319, 128802. doi: 10.1016/j.jclepro.2021.128802

Larimian, T., and Sadeghi, A. (2021). Measuring urban social sustainability: scale development and validation. *Environ. Plan. B Urban Anal. City Sci.* 48, 621–637. doi: 10.1177/2399808319882950

Liang, H., Renneboog, L., and Vansteenkiste, C. (2020). Cross-border acquisitions and employment policies. *J. Corp. Financ.* 62, 101575. doi: 10.1016/j.jcorpfin.2020.101575

Lins, K. V, Servaes, H., Tamayo, A., Begley, T., Clubb, C., Cocco, J., Cooper, M., et al. (2017). Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. *J. Financ.* 72, 1785–1824. doi: 10.1111/jofi.12505

Liu, Y., Lei, L., and Buttner, E. H. (2020). Establishing the boundary conditions for female board directors' influence on firm performance through CSR. *J. Bus. Res.* 121, 112–120. doi: 10.1016/j.jbusres.2020.08.026

Lu, H., Oh, W.-Y., Kleffner, A., and Chang, Y. K. (2021). How do investors value corporate social responsibility? Market valuation and the firm specific contexts. *J. Bus. Res.* 125, 14–25. doi: 10.1016/j.jbusres.2020.11.063

Lueg, K., Krastev, B., and Lueg, R. (2019). Bidirectional effects between organizational sustainability disclosure and risk. *J. Clean. Prod.* 229, 268–277. doi: 10.1016/j.jclepro.2019.04.379

Luffarelli, J., Markou, P., Stamatogiannakis, A., and Gonçalves, D. (2019). The effect of corporate social performance on the financial performance of business-to-business and business-to-consumer firms. *Corp. Soc. Respons. Environ. Manage.* 26, 1333–1350. doi: 10.1002/csr.1750

Lys, T., Naughton, J. P., and Wang, C. (2015). Signaling through corporate accountability reporting. *J. Account. Econ.* 60, 56–72. doi: 10.1016/j.jacceco.2015.03.001

Marlin, A., and Marlin, J. T. (2003). A brief history of social reporting. *Business Respect*. Available online at: http://www.cityeconomist.com/images/A_brief_history_of_social_reporting-2003.doc

McCauley, D., and Heffron, R. (2018). Just transition: Integrating climate, energy and environmental justice. *Energy Policy* 119, 1–7. doi: 10.1016/j.enpol.2018.04.014

Moura-Leite, R. C., Padgett, R. C., and Galán, J. I. (2014). Stakeholder management and nonparticipation in controversial business. *Bus. Soc.* 53, 45–70. doi: 10.1177/0007650310395547

Naseem, T., Shahzad, F., Asim, G. A., Rehman, I. U., and Nawaz, F. (2020). Corporate social responsibility engagement and firm performance in Asia Pacific: the role of enterprise risk management. *Corp. Soc. Respons. Environ. Manage.* 27, 501–513. doi: 10.1002/csr.1815

Ng, A. C., and Rezaee, Z. (2015). Business sustainability performance and cost of equity capital. *J. Corp. Financ.* 34, 128–149. doi: 10.1016/j.jcorpfin.2015.08.003

Nguyen, P., and Nguyen, A. (2015). The effect of corporate social responsibility on firm risk. *Soc. Respons. J.* 11, 324–339. doi: 10.1108/SRJ-08-2013-0093

Nguyen, P.-A., Kecskés, A., and Mansi, S. (2020). Does corporate social responsibility create shareholder value? The importance of long-term investors. *J. Bank. Financ.* 112, 105217. doi: 10.1016/j.jbankfin.2017.09.013

Oh, H., Bae, J., and Kim, S.-J. (2017). Can sinful firms benefit from advertising their CSR efforts? Adverse effect of advertising sinful firms' CSR engagements on firm performance. *J. Bus. Ethics* 143, 643–663. doi: 10.1007/s10551-016-3072-3

Oswald Beiler, M., and Mohammed, M. (2016). Exploring transportation equity: development and application of a transportation justice framework. *Transport. Res. Part D Transport Environ.* 47, 285–298. doi: 10.1016/j.trd.2016.06.007

Pettit, J. (2004). Climate justice: a new social movement for atmospheric rights. *IDS Bull.* 35, 102–106. doi: 10.1111/j.1759-5436.2004.tb00142.x

Ruiz, M., Segui-Pons, J. M., and Mateu-LLad,ó, J. (2017). Improving bus service levels and social equity through bus frequency modelling. *J. Transport Geogr.* 58, 220–233. doi: 10.1016/j.jtrangeo.2016.12.005

Saleem, I., Khan, M. N. A., Hasan, R., and Ashfaq, M. (2021). Corporate board for innovative managerial control: implications of corporate governance deviance perspective. *Corp. Governance Int. J. Bus. Soc.* 21, 450–462. doi: 10.1108/CG-04-2020-0151

Scalet, S., and Kelly, T. F. (2010). CSR rating agencies: what is their global impact? *J. Bus. Ethics* 94, 69–88. doi: 10.1007/s10551-009-0250-6

Shahzad, A. M., and Sharfman, M. P. (2017). Corporate social performance and financial performance: sample-selection issues. *Bus. Soc.* 56, 889–918. doi: 10.1177/0007650315590399

Shaker, R. R., and Sirodoev, I. G. (2016). Assessing sustainable development across Moldova using household and property composition indicators. *Habitat Int.* 55, 192–204. doi: 10.1016/j.habitatint.2016.03.005

Shigetomi, Y., Chapman, A., Nansai, K., Matsumoto, K., and Tohno, S. (2020). Quantifying lifestyle based social equity implications for national sustainable development policy. *Environ. Res. Lett.* 15, 084044. doi: 10.1088/1748-9326/ab9142

Shiu, Y. M., and Yang, S. L. (2017). Does engagement in corporate social responsibility provide strategic insurance-like effects? *Strategic Manage. J.* 38, 455–470. doi: 10.1002/smj.2494

Siew, R. Y. J. (2015). A review of corporate sustainability reporting tools (SRTs). *J. Environ. Manage.* 164, 180–195. doi: 10.1016/j.jenvman.2015.09.010

Silva, J. F. B. A., Rebouças, S. M. D. P., Abreu, M. C. S., de, and Ribeiro, M., da C. R. (2018). Construção de um índice de desenvolvimento sustentável e análise espacial das desigualdades nos municípios cearenses. *Revista de Administração Pública* 52, 149–168. doi: 10.1590/0034-7612163114

Su, S., Liu, Z., Xu, Y., Li, J., Pi, J., and Weng, M. (2017). China's megaregion policy: performance evaluation framework, empirical findings and implications for spatial polycentric governance. *Land Use Policy* 63, 1–19. doi: 10.1016/j.landusepol.2017.01.014

Taylor, J., Vithayathil, J., and Yim, D. (2018). Are corporate social responsibility (CSR) initiatives such as sustainable development and environmental policies value enhancing or window dressing? *Corp. Soc. Respons. Environ. Manage.* 25, 971–980. doi: 10.1002/csr.1513

Tebini, H., M'Zali, B., Lang, P., and Perez-Gladish, B. (2016). The economic impact of environmentally responsible practices. *Corp. Soc. Respons. Environ. Manage.* 23, 333–344. doi: 10.1002/csr.1383

Tong, L., Wang, H., and Xia, J. (2020). Stakeholder preservation or appropriation? The influence of target CSR on market reactions to acquisition announcements. *Acad. Manage. J.* 63, 1535–1560. doi: 10.5465/amj.2018.0229

Tsang, A., Wang, K. T., Liu, S., and Yu, L. (2021). Integrating corporate social responsibility criteria into executive compensation and firm innovation: international evidence. *J. Corp. Financ.* 70, 102070. doi: 10.1016/j.jcorpfin.2021.102070

Tschopp, D., and Huefner, R. J. (2015). Comparing the evolution of CSR reporting to that of financial reporting. *J. Bus. Ethics* 127, 565–577. doi: 10.1007/s10551-014-2054-6

Valizadeh, N., and Hayati, D. (2021). Development and validation of an index to measure agricultural sustainability. *J. Clean. Prod.* 280, 123797. doi: 10.1016/j.jclepro.2020.123797

Vomberg, A., Homburg, C., and Bornemann, T. (2015). Talented people and strong brands: the contribution of human capital and brand equity to firm value. *Strategic Manage. J.* 36, 2122–2131. doi: 10.1002/smj.2328

Wang, Z., and Sarkis, J. (2017). Corporate social responsibility governance, outcomes, and financial performance. *J. Clean. Prod.* 162, 1607–1616. doi: 10.1016/j.jclepro.2017.06.142

Widyawati, L. (2020). A systematic literature review of socially responsible investment and environmental social governance metrics. *Bus. Strategy Environ.* 29, 619–637. doi: 10.1002/bse.2393

Xie, J., Nozawa, W., Yagi, M., Fujii, H., and Managi, S. (2019). Do environmental, social, and governance activities improve corporate financial performance? *Bus. Strategy Environ.* 28, 286–300. doi: 10.1002/bse.2224

Yu, E. P., Guo, C. Q., and Luu, B., Van. (2018). Environmental, social and governance transparency and firm value. *Bus. Strategy Environ.* 27, 987–1004. doi: 10.1002/bse.2047

Yuan, Y., Xu, J., and Wang, Z. (2017). Spatial equity measure on urban ecological space layout based on accessibility of socially vulnerable groups—A case study of changting, China. *Sustainability* 9, 1552. doi: 10.3390/su9091552

Zolotoy, L., O'Sullivan, D., and Chen, Y. (2019). Local religious norms, corporate social responsibility, and firm value. *J. Bank. Financ.* 100, 218–233. doi: 10.1016/j.jbankfin.2019.01.015

frontiers | Frontiers in Sustainability

TYPE Systematic Review
PUBLISHED 03 October 2022
DOI 10.3389/frsus.2022.943145

Check for updates

OPEN ACCESS

EDITED BY
Yogendra Shastri,
Indian Institute of Technology Bombay, India

REVIEWED BY
Obste Therasme,
SUNY College of Environmental Science and Forestry, United States
Xinqiang YOU,
Fuzhou University, China

*CORRESPONDENCE
Shunsuke Managi
managi@doc.kyushu-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 13 May 2022
ACCEPTED 01 September 2022
PUBLISHED 03 October 2022

CITATION
Zulfhazli, Keeley AR, Takeda S and Managi S (2022) A systematic review of the techno-economic assessment of various hydrogen production methods of power generation.
Front. Sustain. 3:943145.
doi: 10.3389/frsus.2022.943145

A systematic review of the techno-economic assessment of various hydrogen production methods of power generation

Zulfhazli[1,2], Alexander Ryota Keeley[1,3], Shutaro Takeda[3] and Shunsuke Managi[1,3]*

[1]Department of Civil Engineering, School of Engineering, Kyushu University, Fukuoka, Japan, [2]Department of Civil Engineering, Universitas Malikussaleh, Lhokseumawe, Indonesia, [3]Urban Institute, Kyushu University, Fukuoka, Japan

Hydrogen is a low or zero-carbon energy source that is considered the most promising and potential energy carrier of the future. In this study, the energy sources, feedstocks, and various methods of hydrogen production from power generation are comparatively investigated in detail. In addition, this study presents an economic assessment to evaluate cost-effectiveness based on different economic indicators, including sensitivity analysis and uncertainty analysis. Proton exchange membrane fuel cell (PEMFCs) technology has the most potential to be developed compared to several other technologies. PEMFCs have been widely used in various fields and have advantages (i.e., start-up, zero-emissions, high power density). Among the various sources of uncertainty in the sensitivity analysis, the cost estimation method shows inflationary deviations from the proposed cost of capital. This is due to the selection process and untested technology. In addition, the cost of electricity and raw materials, as the main factors that are unpredictable.

KEYWORDS

hydrogen, energy primary source, hydrogen production technology, power generation, techno-economic assessment

Introduction

Hydrogen is a low or zero-carbon energy source that is considered the most promising and potential energy carrier of the future (Hanley et al., 2018). The current global demand for pure hydrogen is estimated to be around 70 million tons (Bourne, 2012), whereas the global hydrogen demand is expected to reach more than 300 million tons in 2050 (International Energy Agency., 2015). It has been forecasted that hydrogen will be a leading change in the global energy system toward a sustainable energy system (Staffell et al., 2019). Hydrogen can be produced from renewables, such as hydro, wind, wave, solar, biomass, and geothermal, as well as non-renewables such as coal, natural gas, and nuclear energy sources. Due to its energy carrier's nature, hydrogen offers high flexibility because it is easily converted to electricity in fuel cells for power generation, transportation, etc., (Hosseini and Wahid, 2016). In addition, hydrogen has the potential to deliver economically viable, monetarily, socially, and energy-efficient solutions to challenges related to the rising global energy demand, such as global warming (Dutta, 2014).

Power generation from renewable energy sources has been discovered and studied for decades and has been implemented on a large scale in many countries (IEA., 2016). Renewable energy is the fastest-growing source of electricity generation, and it has been predicted that its share will increase to 39% by 2050. Economic considerations are vital to evaluate the feasibility of an energy system while providing clear and cost-effective criteria. Techno-economic feasibility assessment of a particular technology considers several aspects such as technological appropriateness, economic viability, and financial incentives (Jamil et al., 2012; Rajendran and Murthy, 2019). In specific, techno-economic and sensitivity analysis of the hydrogen production methods is needed to improve the economic aspects of hydrogen. Among others will substantially impact future hydrogen production project designs and the development of innovative approaches to cut total production costs to make the fuel more affordable (Yukesh Kannah et al., 2021).

Although there have been several recent systematic reviews of the techno-economic assessment of hydrogen production, they have not principally been in the context of the techno-economic assessment of various hydrogen production methods of power generation. First (Yukesh Kannah et al., 2021), reviewed the sensitivity of various hydrogen production processes, such as (i) thermochemical conversion (e.g., pyrolysis, gasification, and steam reforming of natural gas), (ii) electrolysis water, (iii) renewable liquid reforming, and (iv) biochemical conversion. In terms of economics, steam reforming of natural gas is an economical and effective method for hydrogen production, as it has low operational (70 to 80%), feedstock (0.3 USD/kg H_2), and production (1.25 to 3.50 USD/kg H_2) costs. Second (El-Emam and Özcan, 2019), highlighted the techno-economic of hydrogen production and the environmental aspect of selected routes. The study found that geothermal, biomass, and nuclear-driven electrolysis and thermochemical technologies may replace conventional methods for hydrogen generation. Third Abe et al. (2019), viewed hydrogen as an appropriate long-term energy carrier for the economy. Solid-stage storage systems based on metal hybrids are a promising alternative to storing hydrogen in a hydrogen-powered system. Metal hybrids cannot store large quantities of hydrogen and are unable to release hydrogen at low temperatures.

Therefore, the main focus of our review discusses various hydrogen production methods, including their techno-economic aspects, sensitivity analysis, and uncertainty analysis. In addition, the current study addresses the following research question: "What are the economic performance indicators of the hydrogen energy systems for power generation?" This study aims to determine the economic performance indicators of hydrogen energy systems for power generation.

The remainder of this paper is organized as follows. Section Methods—systematic review of the literature introduces the method systematic of the literature review. Section Results presents the main results, including feedstock, hydrogen production methods, techno-economic performance, sensitivity, and uncertainty analyses. Finally, Section Discussion contains the discussion.

Methods—systematic review of the literature

The general systematic review of the literature is carried out based on the method suggested by Tranfield et al. (2003), Thürer et al. (2018) for retrieving and selecting published data sets from Scopus and the Web of Science (WoS). The primary goal is to find and choose articles that describe hydrogen production methods, power generation, and techno-economic performance indicators. The articles are gathered by conducting a thorough search of Scopus and the WoS. The selection of articles is based on the title, keywords, abstracts, highlights, and type of document. This study uses different keywords for the search, such as "economic simulation AND hydrogen OR cost energy" in Scopus and "hydrogen OR H_2 AND economic simulation OR energy cost" in the WoS. The document type is restricted to articles and reviews, excluding conference papers and books. In addition, the publication year is restricted to 2000–2020.

The selected articles are those that are relevant to the topic of this review and are grouped based on the quality of the research, that is, whether the article answers a series of questions related to the research and describes the facts based on real research scenarios. The analysis is carried out based on energy sources, feedstock, various hydrogen production processes, technique production, power generation, techno-economics in commercialization, and the economics of various hydrogen production processes. The following research questions are added based on the various researches and the analysis of articles:

- What is the source or primary energy of hydrogen production?
- What is the feedstock of hydrogen production?
- What are the types of hydrogen production methods?
- What is the technique of hydrogen conversion?
- What are the types of hydrogen production methods for power generation?
- What are the economic performance indicators of the hydrogen energy system for power generation?

The original sample of 901 articles comprises 392 articles in Scopus and 509 articles in the WoS (eight articles were removed because they were duplicates). After excluding apparently unrelated articles, that is, articles that are not related to hydrogen production and techno-economic, the number was reduced to 741 articles. The high number of unrelated articles is due to the use of the common keyword hydrogen. The articles were further reduced by 152 after screening them based on title and abstract. Finally, the total number articles that are used for the analysis is 52 (Figure 1).

Results

The review results of the techno-economic assessment of various hydrogen production are obtained from 116 case studies, which were mostly about countries in Europe (48/166) (Figures 2a,b). The country-wise distribution is as follows: Italy (8/48), Greece (8/48), Germany (8/48), Romania (8/48), Norway (5/48), France (2/48), Finland (2/48), Turkey (2/48), Spain (2/48), Switzerland (1/48), and Serbia (1/48). The country-wise distribution of the techno-economic assessment of various hydrogen production studies on Asia (42/116) is as follows: China (21/42), Iran (5/42), Thailand (5/42), Republic of Korea (3/42), Saudi Arabia (3/42), Pakistan (2/47), and UEA (2/42), and that of the American continent is 19/116, including Canada (11/19), USA (4/19), Brazil (3/19), and Mexico (1/19). However, the contribution from Australia (5/116) and Africa—represented by Morocco (2/116)—are comparatively small. The primary energy sources in Europe are dominated by photovoltaic (15/48), followed by wind (13/48), unspecified renewable energy sources (4/48), coal (2/48), biodiesel (1/48), hydropower (1/48), and unspecified sources (4/48). The primary energy sources in Asia are photovoltaic (11/42), wind (9/42), biomass (7/42), methanol (5/42), photovoltaic/wind (4/42), natural gas (2/42), as well as coal (1/42), algae nuclear energy (1/42), and renewable energy sources (1/42). Finally, the primary energy sources in Australia comprise wind (3/5) and photovoltaic (2/5), and the primary energy source in Africa is photovoltaic (2/2) (Figure 3).

Feedstock

Hydrogen is not a source of energy, but it is a pure form that functions as an energy carrier or as an industrial raw material (Ozbilen et al., 2011). Hydrogen can be combined with other materials to produce hydrogen-based fuels (Bourne, 2012). Hydrogen feedstocks can be produced from sources such as natural gas, coal, water, biomass, and fossil fuels and can be readily used in engines or turbines (Donaldson et al., 2012; Ren et al., 2013; Yao et al., 2017; Li et al., 2018; Nurdiawati et al., 2019) (Figure 4). Figure 4 reveals that water is the most widely used hydrogen feedstock in countries such as China, Canada, Italy, Brazil, the USA, and the Republic of Korea, followed by other feedstocks, such as coal, coal plus biomass (soil waste), biomass, and natural gas.

Hydrogen production methods

Hydrogen elements can be found abundantly in nature, such as freshwater, seawater, biomass, hydrogen sulfide, and fossil fuels. However, to produce hydrogen with zero or low environmental impact, it must be extracted from fossil fuels. In general, the process of extracting hydrogen from natural resources can be classified into four categories—thermal, electrical, photonic, and biochemical. Thermal and electrical energy can be produced from renewable energy (such as solar, wind, geothermal, hydro, and biomass), fossil energy, or nuclear energy. Photonic energy can be obtained from solar radiation only. Biochemical energy reserved in organic matter can be processed by microorganisms that produce hydrogen from sundry substrates, or it can be chemically transferred to thermal energy (Dincer, 2012; Dincer and Acar, 2014). Previous studies grouped all the case studies into the following four categories based on the classification of various hydrogen production methods: electrochemical, thermochemical, biochemical, and thermal-electrochemical (Dincer, 2012; Dincer and Acar, 2014).

First, the hydrogen production methods in the electrochemical category include electrolysis technologies, such as alkaline electrolyzer (AEL) and proton exchange membrane electrolyzer (PEMEL). AEL and PEMEL are mature and commercially available. AEL is the world's oldest and most widely utilized technology for large-scale systems. PEMEL are generally used for hydrogen production on a modest scale. While PEMEL offers some advantages compared to AEL, including high current densities, voltage efficiency, and quick system response when working dynamically (David et al., 2019) (Yodwong et al., 2020). Electrolysis is the process through which electricity is used to split water into its components (i.e., oxygen and hydrogen). Hydrogen production processes through nuclear-based thermochemical cycles and renewable energy base electrolysis have much lower effects on the environment than steam reforming (Ozbilen et al., 2011). Water is infiltrated into the proton exchange membrane electrolysis cell; then, hydrogen ions are absorbed by the membrane, and this recombining process forms the hydrogen molecules. Proton exchange membrane electrolyzes are considered an alternative to producing hydrogen from renewable energy sources (Silva et al., 2010). Plasma decomposition of natural gas was included in the electrochemical category. In the reviewed studies, electrochemical technology was the most extensively used method in hydrogen production, accounting for approximately 74.14% of all the case studies, followed by thermochemical technology (22.41%) and thermal-electrochemical technology (3.45%) (Figure 5a). Furthermore, electrolysis was the most frequently used method in the electrochemical category, accounting for 84%, and alkaline electrolysis accounts for 3% of all the case studies. Second, the thermochemical category includes aqueous stage reforming, auto thermal reforming, steam reforming, gasification (coal or biomass),

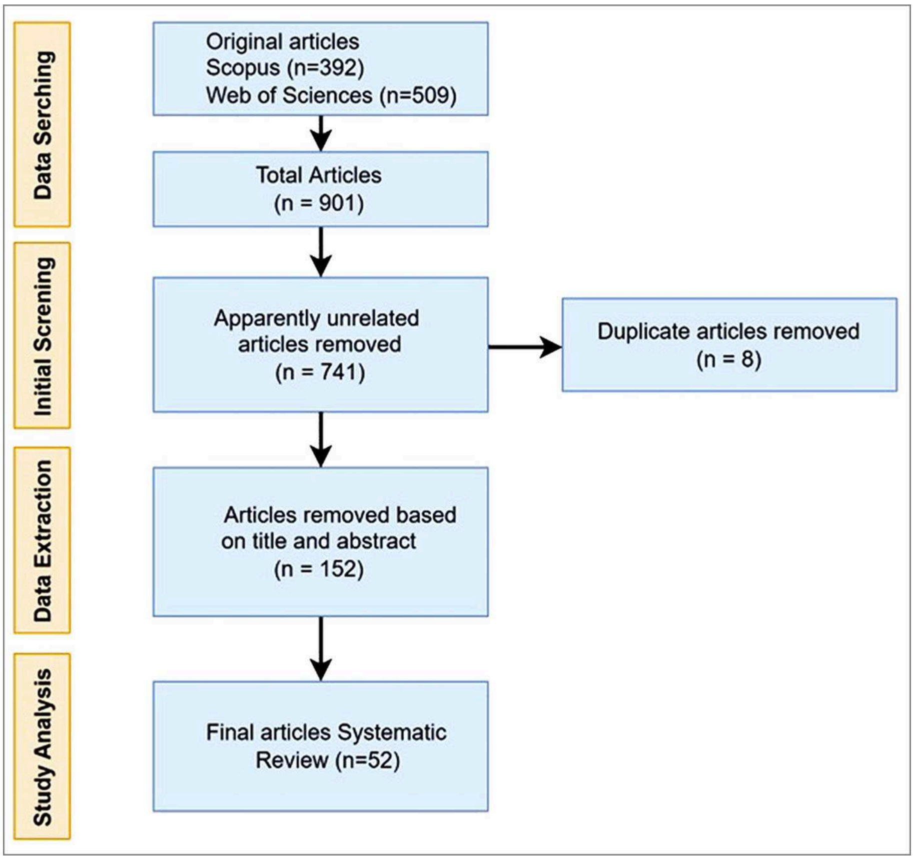


FIGURE 1
The systematic literature review process.

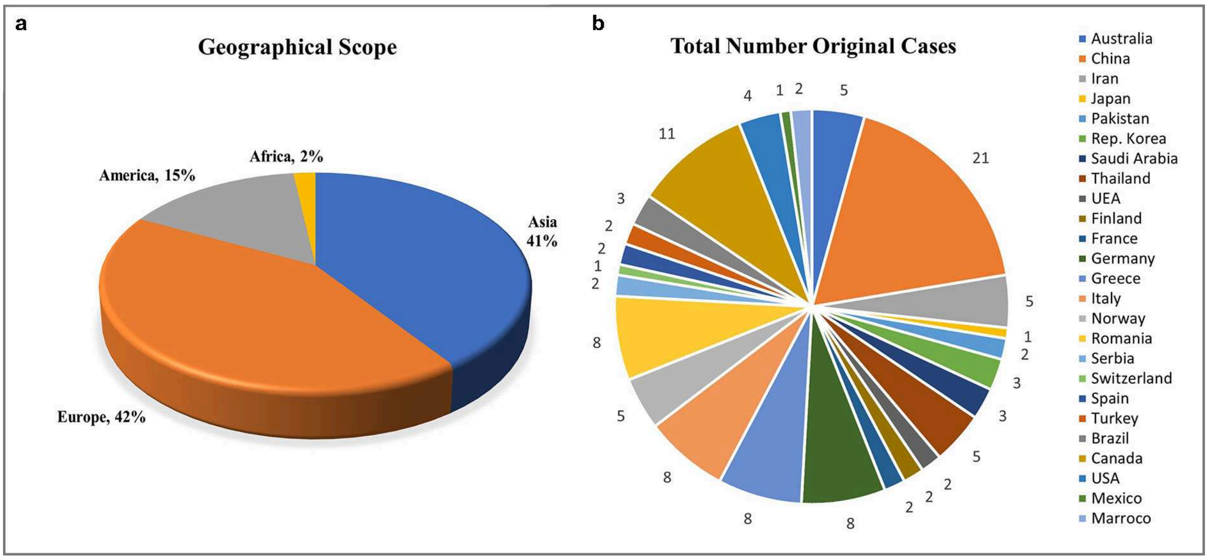


FIGURE 2
(a) Geographical scope and **(b)** the number of original case studies.

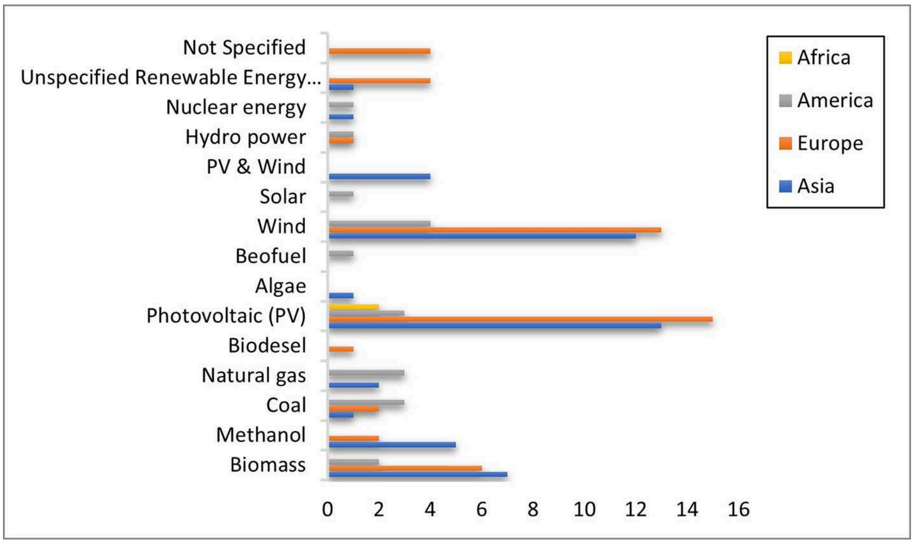


FIGURE 3
Geographical primary energy sources.

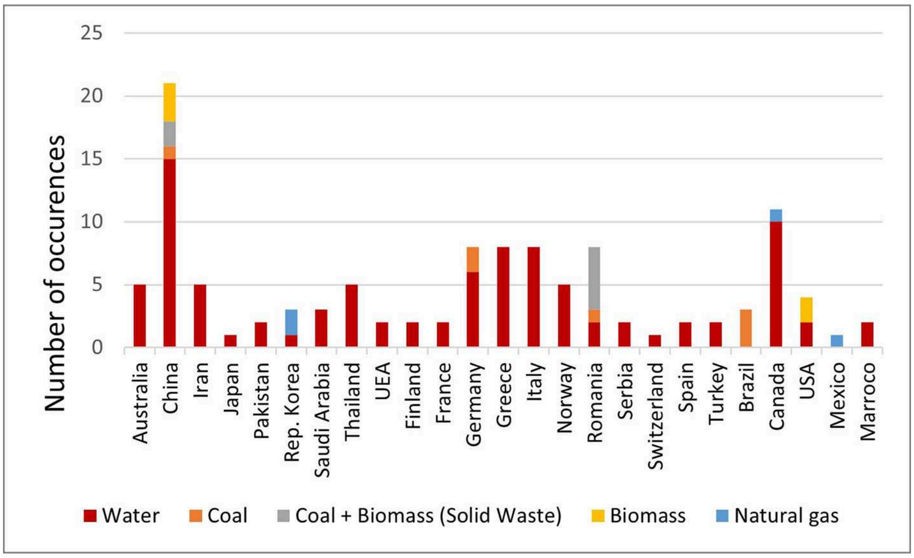


FIGURE 4
Hydrogen feedstock of various countries.

thermal cracking of fossil hydrocarbons, and water splitting. Thermochemical is the process of separating water using a heater to obtain hydrogen. The thermochemical hydrogen production process is an immature technology that must be refined over time. Gasification and reforming based on thermochemical account for 21.55 and 2.59% of all the case studies (Figure 5b). The thermochemical cycle normally does not require catalysts as a driver of chemical reactions. Chemical materials involved in the process are recycled and are the material source from which hydrogen is derived. The water-splitting thermochemical cycle is as follows: (i) it does not require hydrogen–oxygen separation membranes; (ii) it does not require overestimating thermal energy source (600–1,200 k); (iii) it does not require extra electrical energy to drive the process (Dincer, 2012). Third, the biochemical category includes fermentation and dark fermentation. The Biochemical category includes photolytic (direct water separation), photosynthetic bacteria (solar-assisted organic decomposition), dark fermentation

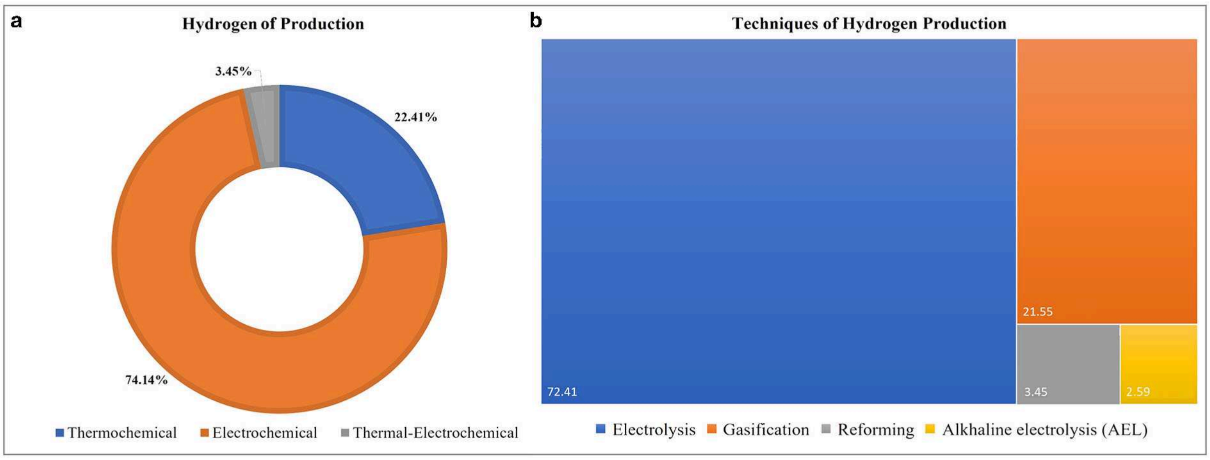


FIGURE 5
(a) Hydrogen production methods and **(b)** technique of hydrogen production.

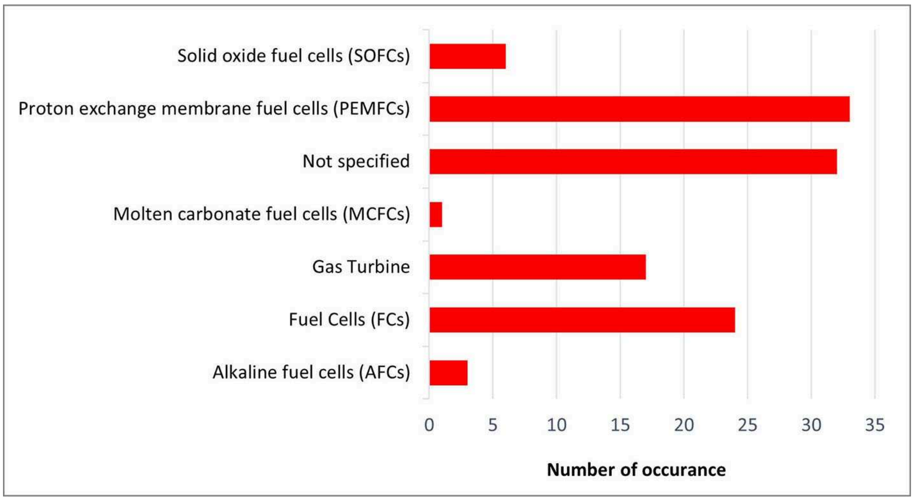


FIGURE 6
Hydrogen power generation.

(organic decomposition), and microbial-assisted electrolysis (electrical-assisted organic decomposition).

Power generation

The development of renewable energy sources (RES) is important for the sustainable growth of any nation due to the depletion of fossil fuels, the rising cost of fossil fuels worldwide, and the need to reduce emission levels. The selection and deployment of hydrogen-based power generation conversion technology are mostly governed by the electricity-requiring application. Technologies that use hydrogen as a fuel cell for power generation must provide flexible energy to ensure stability and resilience. According to the type of electrolyte used, fuel cells (FCs) can be categorized as phosphoric acid fuel cells (PAFCs), molten carbonate fuel cells (MCFCs), solid oxide fuel cells (SOFCs), and proton exchange membrane fuel cells (PEMFCs), molten carbonate fuel cells (MCFCs). The hydrogen that reacts with oxygen in the fuel to supply electrical energy consists of a piston engine and a gas turbine (Figure 6). A fuel cell is a device that converts the chemical energy of a fuel into electrical energy through an electrochemical reaction. Fuel cells are a flexible power generation technology with 50–60% electrical efficiency. Fuel cell stacks have a shorter technical lifetime (10.000 to 40.000 operating hours). However, compared to PEMFCs, and hydrogen is available, PEMFCs have the most potential for development (Bourne, 2012; Wang et al., 2017). Now-a-days, PEMFCs are applied extensively in numerous fields. PEMFCs provide the advantages of practically zero emissions, high power density, high efficiency, and low operating temperature compared to other fuel cell types. In addition, PEMFC providing short start and response times at the stack level appears to be the optimal technology for application drives.

Hydrogen gas turbine power generation technology is designed for large-scale power generation. Incorporating hydrogen is a potential pathway for gas turbine decarbonization by replacing natural gas with hydrogen. Each gas turbine model has a specific capability for hydrogen combustion, mainly determined by the combustion system. Gas turbine technology has three main components: a compressor, a combustion chamber, and a turbine stage. The central part of the energy is the turbine stage, which drives the compressor and gives the generator the power to run and generate electricity (Wang et al., 2021). Showed that the output of the introduction cycle is composed of a wind turbine, solar energies, and AFCs was 10.5 kW of electricity, and the electrical efficiency was 56.9%. In addition, the electrolyzer uses 9.9 kW of electricity to produce 221.3 grams of hydrogen fuel.

Techno-economic performance indicator

This section discusses the techno-economic analysis, including profitability, sensitivity, and uncertainty analysis, using various simulation results, such as Monte Carlo simulation, Aspen HYSYS (Kim et al., 2018), Aspen Plus, MATLAB, and HOMER simulation. The Aspen HYSYS simulation model is used to determine the effect of various operating conditions on the performance of the packed-bed reactor and membrane reactor (Kim et al., 2018). Moreover, it is used to determine the future risk and uncertainty in prediction (Zahid et al., 2020). Techno-economic assessment is a methodological framework for examining the technical and economic performance of a process, product, or service and includes the study of the economic impact of technology. A techno-economic assessment (TEA) is a cost-benefit comparison that considers technological and economic factors. An economic summary of hydrogen production is presented in Table 2, where each cost component is presented including capital expenditure (CAPEX), operating expenditure (OPEX), and other variables.

The CAPEX and OPEX are the main costs in a techno-economic assessment. The key issue is to minimize the CAPEX and OPEX of various hydrogen generation systems while simultaneously increasing production volume. This allows for a reduction in the cost of producing hydrogen from several energy sources. The expenditures involved with building a new facility are referred to as CAPEX. Fixed-capital investment (FCI) is the funds used to finance a facility. FCI in the first and second years is 60% and 40% of total FCI, respectively, while working capital cost is 15% of total FCI (Lee et al., 2020). Likewise, OPEX represents the various day-to-day expenses required to maintain sustainable business operations. It can also be said that they refer to the enormous costs involved in maintaining plant operations. OPEX consists of the costs of raw materials, operating labor, maintenance, and utilities. Annual expenses are considered based on items related to operating expenses and general and administrative expenses. In many cases, CAPEX values are estimated using software such as Aspen Plus or Aspen Hysys to simulate processes and perform economic analysis. The influence of plant size and capacity on CAPEX is substantial. The larger the facilities, the more the CAPEX, but the lower the production expenses. In addition, environmental influences such as integrated carbon capture and storage systems have a major impact on total production costs, resulting in an increase in CAPEX due to the use of additional equipment.

Sensitivity analysis

The main objective of the sensitivity analysis (SA) is to obtain the effect of various economic factors on the cost of a unit of hydrogen produced and determine some influential factors, including ensuring the surroundings and conditions of any operating plant after investment (Kim et al., 2018). The SA can provide information on the factor that is most sensitive and has a significant impact, including making decisions before investing. Generally, the sensitivity indicators to consider include sensitivity to capital cost, sensitivity to feedstock, and sensitivity to the internal rate of return (Khunathorncharoenwong et al., 2020; Yukesh Kannah et al., 2021). In some instances that renewables, such as wind, were used as the electricity source, several variable inputs were estimated, such as plant parameters (e.g., capacity and storage capacity), capital expenditure (CAPEX) parameters (e.g., hydrogen storage, electrolysis, and methanation), operation expenditure (OPEX) parameters (e.g., standby cost), and operating parameters (e.g., restart a level and restart time) (Rivera-Tinoco et al., 2016; Gorre et al., 2020).

The crucial parameter to perform sensitivity analysis depends on the hydrogen production process. Many studies consider capital cost, operating cost, replacement, operation & maintenance, and net present value (NPV) for process electrolysis. Other studies consider the parameters of hydrogen cost, sales price, consumption, operation expenditure, fuel, and savagery, including taxes. Reforming process parameters consist of hydrogen production costs: reactor, membrane module, compressor, pressure swing adsorption (PSA), supplement, reactants, PSA OPEX, electricity, labor, natural gas, membrane replacement, maintenance, and other costs were considered for sensitivity analysis. Finally, the gasification process only considers the NPV (Table 1).

The wind is the basic concept of SA in the application of power to gas technology to convert renewable electricity into molecular form. Electrolysis costs are reduced by 54%, and gas

TABLE 1 Summary sensitivity analysis parameter.

Article	Hydrogen production	Sensitivity analysis parameter	Key performance indicator
Gorre et al. (2020)	Electrolysis	Plant parameter, CAPEX, OPEX	-
Khunathorncharoenwong et al. (2020)	Electrolysis	Hydrogen cost, sales price, and consumption	Net present value
Hamayun et al. (2019)	Electrolysis	Operation expenditure	-
Kim et al. (2018)	Reforming	Hydrogen production cost: reactor, membrane module, compressor, pressure swing adsorption (PSA), supplement, reactants, PSA OPEX, electricity, labor, natural gas, membrane replacement, maintenance, and other costs were considered for SA	Net production cost
Rivera-Tinoco et al. (2016)	Electrolysis	Electricity price, lifespan, investment, maintenance, electrolyze, low-cost power electrolyzes, and high equipment lifespan	-
König et al. (2015)	Electrolysis	NPC: capital cost, wind power, carbon dioxide cost, Oxygen revenue, and cavern capital cost	-
Guinot et al. (2015a)	Electrolysis	Capital cost, fuel cost, operation & maintenance, interest rate, and availability factor	-
Donaldson et al. (2012)	Gasification	NPV: sunflower residue, activated carbon, and hydro price	-
Tzamalis et al. (2011)	Electrolysis	NPV: capital cost, replacement, O&M, fuel, and salvage	-
Tsatsaronis et al. (2008)	Gasification	Capital cost, cost of heat, cost of coal, and currency (current and constant)	-
Shaner et al. (2016)	Electrolysis	Capital cost, operating expenses, replacement cost, and tax	-

production costs are reduced by 40% implying a lower average price for hydrogen, thus allowing for reduced equipment costs. A reduced methanation CAPEX can reduce the amount of hydrogen that is not converted into synthetic natural gas. Thus, synthetic natural gas (SNG) production costs are more sensitive to CAPEX electrolyzed than CAPEX methanation (Gorre et al., 2020). Hydrogen price is the most sensitive parameter and is more economical in the conventional process than low-pressure steam consumption (Khunathorncharoenwong et al., 2020). A heavier load on the electrolysis section results in higher power plate CAPEX and OPEX. However, the system efficiency can impact the high cost reduction process of all systems because of the areal dependencies of most of the components (Shaner et al., 2016; Hamayun et al., 2019). When the overloaded functionality of the installed capacity is 5.0%, the cost can be reduced, leading to a capital cost reduction of 3.6%. NPC was reduced by 0.9% due to the high cost of the electricity component. The output electricity cost is highly sensitive to the efficiency of the power plant (Zahid et al., 2020). Furthermore, lesser by-product yield is substantial from an economical perspective.

An NPV is considered one of the indicators to decide the feasibility of the target technology (Lee et al., 2020). When the NPV is zero, the project is not expected to generate significant profits or losses. Therefore, a project with a positive NPV is considered profitable and acceptable, while a project with a negative NPV means that this technology needs to be developed to obtain economic gains. The NPV decreases as the price of renewable electricity or the rate of degradation increases in relation to the cost of the system. On the other hand, the internal rate of return (IRR), is the discount rate that corresponds to an NPV equal to zero. IRR is a financial risk indicator used to assess the profitability of an investment. Where IRR involves

TABLE 2 Summary economic of hydrogen production.

References	Energy source	Technology of H2 production	H2 for power generation	Capital expenditure	Operational expenditure	Interest rate	Project lifetime (years)	H2 production capacity	Plant efficiency (%)	Electricity cost ($/kWh)	H2 cost
Lee et al. (2020)	Unspecified renewable energy	Electrolysis	Not specified	n/a	n/a	-	10	700 Nm 3 h-1	-	-	3.88–9.30
Zahid et al. (2020)	Nuclear energy	Electrolysis	PEMFCs	2,291.4 $/kW	-	4	-	266 MW	-	-	-
Schnuelle et al. (2020)	Photovoltaic (PV)	Alkaline Electrolysis	FCs	n/a	n/a	-	-	770–1,324 €/kW	-	n/a	n/a
Liu et al. (2020)	Photovoltaic, wind	Electrolysis	Gas turbine	-	2.374–2.379	-	-	-	-	-	-
Gorre et al. (2020)	Wind	Electrolysis	FCs	650 €/kWel	n/a	n/a	20	-	-	n/a	-
Khunathorncharoenwong et al. (2020)	Not specified	Electrolysis	Gas turbine, FCs	2.8–3.4 m	-	-	-	-	-	-	4.020 $/kg
Wang et al. (2019)	Not specified	Gasification	Gas turbine	n/a	n/a	-	15	-	-	-	-
Nurdiawati et al. (2019)	Algae	Gasification	PEMFCs	-	n/a	-	-	-	-	0.030 $/kg	n/a
Jiang et al. (2019)	Wind	Gasification	Not specified	n/a	n/a	-	20	-	0.47–1	-	4.34€/kg
Hamayun et al. (2019)	Photovoltaic, wind	Electrolysis	PEMFCs	21.288.900 $	7.645.920$	-	-	5 MW	n/a	-	-
Nieminen et al. (2019)	Wind	Electrolysis	PEMFCs	n/a	n/a	-	20	30 MW	-	624–625	2.90–3.40$/kg
Martínez-Salazar et al. (2019)	Natural gas	Reforming	Not specified	n/a	n/a	-	40	-	-	-	-
Jamshidi and Askarzadeh (2019)	Photovoltaic	Electrolysis	Gas turbine	-	-	-	20	-	-	-	-
Touili et al. (2018)	Photovoltaic	Electrolysis	PEMFCs, SOFCs	n/a	n/a	-	-	-	-	-	-

(Continued)

10.3389/frsus.2022.943145

10.3389/frsus.2022.943145

TABLE 2 (Continued)

References	Energy source	Technology of H2 production	H2 for power generation	Capital expenditure	Operational expenditure	Interest rate	Project lifetime (years)	H2 production capacity	Plant efficiency (%)	Electricity cost ($/kWh)	H2 cost
Duman and Güler (2018)	Wind	Electrolysis	PEMFCs, FCs	n/a	n/a	-	20	-	0.1694	-	-
Li et al. (2018)	Coal	Gasification	No specified	n/a	n/a	10	25	-	-	-	120 CNY/kg
Kim et al. (2018)	Natural gas	Reforming	FCs	n/a	n/a	-	-	-	-	-	n/a
Haghi et al. (2018)	Natural gas, Beofule, Wind, Solar	Electrolysis	Not specified	n/a	n/a	-	20	-	-	-	n/a
Al-Sharafi et al. (2017)	Photovoltaic, Wind	Electrolysis	Gas turbine, PEMFCs, SOFCs	2,000 $/kW	-	-	25	-	-	-	-
Aziz (2017)	Photovoltaic, Wind	Electrolysis	FCs	9,500 $	250$/year	-	15	-	-	-	-
Yao et al. (2017)	Biomass	Gasification, Reforming and Alkaline electrolysis	PEMFCs, FCs, SOFCs	n/a	n/a	-	25	-	-	-	90 kg h-1
Ye et al. (2017)	Photovoltaic, Wind	Electrolysis	PEMFCs	n/a	n/a	-	20–25	-	-	-	-
Schlachtberger et al. (2017)	Photovoltaic, Wind	Electrolysis	Not Specified	n/a	n/a	-	25–80	-	-	-	-
Walker et al. (2016)	Natural Gas	Electrolysis	Not Specified	-	n/a	-	-	-	-	n/a	-
Martin et al. (2016)	Biodiesel	Electrolysis	Not Specified	-	-	7	-	-	-	-	-
Brka et al. (2016)	Wind	Electrolysis	PEMFCs	-	-	-	25	-	-	-	-
Rivera-Tinoco et al. (2016)	Methanol	Electrolysis	PEMFCs, SOFCs	n/a	n/a	-	-	-	-	n/a	-
Rivarolo et al. (2016)	Photovoltaic, Wind	Electrolysis	Not Specified	-	-	-	-	-	-	-	-
Stojković and Bakić (2016)	Photovoltaic, Wind	Electrolysis	FCs	n/a	n/a	-	20	-	-	-	-
König et al. (2015)	Wind	Electrolysis	PEMFCs	n/a	n/a	-	-	-	-	-	-
Cormos (2015)	Not Specified	Electrolysis	PEMFCs	n/a	n/a	-	25	-	-	n/a	-

(Continued)

TABLE 2 (Continued)

References	Energy source	Technology of H2 production	H2 for power generation	Capital expenditure	Operational expenditure	Interest rate	Project lifetime (years)	H2 production capacity	Plant efficiency (%)	Electricity cost ($/kWh)	H2 cost
Guinot et al. (2015b)	Not Specified	Electrolysis	PEMFCs	n/a	n/a	-	-	-	-	-	-
Guinot et al. (2015a)	Photovoltaic	Electrolysis	PEMFCs	n/a	n/a	-	20	-	-	n/a	-
Olateju et al. (2014)	Wind	Electrolysis	Gas turbine	n/a	n/a	-	20	563 MW	-	-	-
Sarkar and Bhattacharyya (2012)	Photovoltaic, Wind	Electrolysis	Gas turbine	n/a	n/a	-	-	-	-	n/a	-
Cormos (2014)	Biomass	Gasification	Not specified	n/a	n/a	-	-	400–425 MW	-	-	-
Shiroudi et al. (2013)	Photovoltaic	Electrolysis	PEMFCs	n/a	n/a	-	25	-	0.7	-	-
Tzamalis et al. (2013)	Wind	Electrolysis	FCs	n/a	n/a	-	-	-	-	-	-
Banerjee et al. (2013)	Biomass	Gasification	SOFCs	n/a	n/a	-	20	2,000	-	-	-
Donaldson et al. (2012)	Biomass	Gasification	Not specified	8.6 $M	n/a	0.06	-	-	-	0.12 $/kWh	-
Carapellucci and Giordano (2012)	Unspecified renewable energy	Electrolysis	PEMFCs	n/a	n/a	-	-	-	-	-	-
Shabani and Andrews (2011)	Photovoltaic	Electrolysis	FCs	n/a	-	-	20	-	-	-	-
Tzamalis et al. (2011)	Photovoltaic	Electrolysis	FCs	n/a	n/a	-	-	-	-	-	-
Tsatsaronis et al. (2008)	Coal	Gasification	AFCs	n/a	n/a	-	-	-	-	-	-
Greiner et al. (2007)	Photovoltaic	Electrolysis	GT, FCs	n/a	n/a	-	25	-	-	-	-
Zoulias and Lymberopoulos (2007)	Photovoltaic	Electrolysis	FCs	n/a	n/a	-	20	-	-	-	-
Santarelli and Macagno (2004)	Photovoltaic	Electrolysis	GR, MCFCs	n/a	-	-	-	-	-	-	-
Scherer et al. (1999)	Not specified	Electrolysis	AFCs	n/a	n/a	-	-	-	-	-	-
Shaner et al. (2016)	Photovoltaic	Electrolysis	PEMFCs	n/a	n/a	-	-	-	0.61	-	-

10.3389/frsus.2022.943145

comparing more than one potential project, the level of internal investment indicates the one that is most profitable, regardless of project size and technology. According to established practice, an internal rate of return (IRR) of 10% is assumed, consisting of the interest rate for own capital and credit capital. Hydrogen production costs are calculated iteratively using a plant cash flow analysis that includes total annual expenses and revenues.

Uncertainty analysis

The evaluations built on assumptions and estimates inevitability produce uncertainty in results. TEA describes uncertainty caused explicitly by errors in data input, the tension in the model itself, and the characteristic of the context in which the analysis is carried out. In the initial step of uncertainty analysis, it is always important to systematically identify the variables that generate uncertainty. The second step is determining the number of computations necessary to confirm compliance with the acceptance criteria and standard tolerance limit. In addition, TEA performs an uncertainty analysis to evaluate the parameters that most influence the project's economic performance. For example, the sensitivity parameter might vary by up to 20% relative to the baseline value (van der Spek et al., 2020). Sensitivity analysis assesses the influence of a single parameter at a time. In the meantime, a Monte Carlo simulation was conducted to examine the combined effect of numerous parameters on the economic performance of an investment. This simulation forecasts economic indicator uncertainty by randomly generating parameter values within the ranges above. In addition, the simulation examines the process's uncertainty and calculates the chance that the developed system will be profitable. Here, (Lee et al., 2020) conducted an uncertainty analysis to identify changes in the unit price of electricity and the selling price of H2 in the net present value range. The uncertainty analysis reveals that, economically, the selling price of hydrogen is more influential than renewable electricity prices, such as hydro and onshore wind energy, which is considered promising renewable power source for reducing the cost of producing hydrogen.

Discussion

This section examines the techno-economic assessment of various hydrogen productions for power generation studies included in this systematic review. Natural gas and coal are the two most crucial feedstock sources for hydrogen generation. The technology for producing hydrogen from these two feedstocks is highly developed, and there is a lot of experience operating these plants. The cost of hydrogen from various energy sources depends on the energy conversion and production costs. Most hydrogen production techniques require either thermal or electrical energy input from the energy source. Concurrently, this energy source is supplied by the energy conversion plant, representing the increase in energy's final cost. These expenses are typically the most significant contributors to the total cost of hydrogen.

Compared to other fuel cells (FCs), proton exchange membrane fuel cells (PEMFC) have become a power source for many applications and a possible option for reducing greenhouse gas emissions. PEMFCs combined with photovoltaics and batteries are now considered an excellent alternative to power generation. The application of the independent control mode can realize the optimal economical operation of the hybrid power generation system (HPGS) without a communication network. It can reduce marginal cost by up to 19.08% compared to traditional droop control. Furthermore, the cooperative control mode can achieve minimum generation costs and a difference in battery energy storage devices' charge balance state, even when the line resistance effect is quite significant (Yang et al., 2019; Okonkwo et al., 2021).

A techno-economic assessment is important now-a-days, but there are many different uncertainties in how to calculate it. The use of non-standard procedures, assumptions, and data of varying quality makes it difficult to compare the values of the literature with each other and draw rational conclusions. Several assumptions are made when calculating TEA, such as type of financing, cost and space of land acquisition, cost of raw materials, the yield of raw materials, factory life, construction time, labor costs, product costs, and utility costs. However, these assumptions do not reflect the actual reality. It affects the calculation. On the other hands, Sensitivity analysis helps determine the state and condition of each plant operating after the investment. It is more beneficial to decide before investing. The sensitivity to the cost of capital can be determined by calculating the return on investment. This is a key parameter to identify the technology running from start to finish and the return on investment at each stage of growth.

In closing, numerous researches focus on hydrogen production sources, systems, and distinct hydrogen storage alternatives. In addition, studies focusing on the social and environmental implications of the sources and systems necessary for hydrogen production are scarce. Another drawback of this study is that it concentrates on metrics that cannot be compared to others get more definitive results. Future research might include hydrogen end-use possibilities, such as various fuel cells, to improve the analysis of the long-term viability of hydrogen-based energy systems.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number (s) can be found in the article/supplementary material.

Author contributions

Z carried out the experiment with support from AK and ST. Z and AK wrote the manuscript. SM helped supervise the project. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by JSPS KAKENHI Grant Number JP20H00648 and the Environment Research and Technology Development Fund (JPMEERF20201001) of the Environmental Restoration and Conservation Agency of Japan.

Acknowledgments

Z would like to thank Malikussaleh University, Asian Development Bank (ADB) and The Ministry of Research, and Technology and Higher Education of the Republic of Indonesia for the full support through the Doctoral Degree Scholarship Program (Grant No. L3749-INO).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

References

Abe, J. O., Popoola, A. P. I., Ajenifuja, E., and Popoola, O. M. (2019). Hydrogen energy, economy and storage: review and recommendation. *Int. J. Hydrog. Energy* 44, 15072–15086. doi: 10.1016/j.ijhydene.2019.04.068

Al-Sharafi, A., Sahin, A. Z., Ayar, T., and Yilbas, B. S. (2017). Techno-economic analysis and optimization of solar and wind energy systems for power generation and hydrogen production in Saudi Arabia. *Renew. Sust. Energ. Rev. 69*, 33–49. doi: 10.1016/j.rser.2016.11.157

Aziz, A. S. (2017). Techno-economic analysis using different types of hybrid energy generation for desert safari camps in UAE. *Turk. J. Electr. Eng. Comput. Sci.* 25, 2122–2135. doi: 10.3906/elk-1602-159

Banerjee, S., Tiarks, J. A., Lukawski, M., Kong, S. C., and Brown, R. C. (2013). Technoeconomic analysis of biofuel production and biorefinery operation utilizing geothermal energy. *Energy Fuels* 27, 1381–1390. doi: 10.1021/ef301898n

Bourne, S. (2012). The future of fuel: the future of hydrogen. *Fuel Cells Bull.* 2012, 12–15. doi: 10.1016/S1464-2859(12)70027-5

Brka, A., Al-Abdeli, Y. M., and Kothapalli, G. (2016). Predictive power management strategies for stand-alone hydrogen systems: operational impact. *Int. J. Hydrog. Energy* 41, 6685–6698. doi: 10.1016/j.ijhydene.2016.03.085

Carapellucci, R., and Giordano, L. (2012). Modeling and optimization of an energy generation island based on renewable technologies and hydrogen storage systems. *Int. J. Hydrog. Energy* 37, 2081–2093. doi: 10.1016/j.ijhydene.2011.10.073

Cormos, C. C. (2014). Techno-economic and environmental analysis of hydrogen and power co-generation based on co-gasification of coal and biomass / solid wastes with carbon capture. *Chem. Eng. Trans.* 37, 139–144. doi: 10.3303/CET1437024

Cormos, C. C. (2015). Biomass direct chemical looping for hydrogen and power co-production: process configuration, simulation, thermal integration and techno-economic assessment. *Fuel Process. Technol.* 137, 16–23. doi: 10.1016/j.fuproc.2015.04.001

David, M., Ocampo-Martínez, C., and Sánchez-Peña, R. (2019). Advances in alkaline water electrolyzers: a review. *J. Energy Storage 23*, 392–403. doi: 10.1016/j.est.2019.03.001

Dincer, I. (2012). Green methods for hydrogen production. *Int. J. Hydrog. Energy* 37, 1954–1971. doi: 10.1016/j.ijhydene.2011.03.173

Dincer, I., and Acar, C. (2014). Review and evaluation of hydrogen production methods for better sustainability. *Int. J. Hydrog. Energy* 40, 11094–11111. doi: 10.1016/j.ijhydene.2014.12.035

Donaldson, A. A., Kadakia, P., Gupta, M., and Zhang, Z. (2012). Production of energy and activated carbon from agri-residue: sunflower seed example. *Appl. Biochem. Biotechnol.* 168, 154–162. doi: 10.1007/s12010-011-9358-0

Duman, A. C., and Güler, Ö. (2018). Techno-economic analysis of off-grid PV/wind/fuel cell hybrid system combinations with a comparison of regularly and seasonally occupied households. *Sustain. Cities Soc.* 42, 107–126. doi: 10.1016/j.scs.2018.06.029

Dutta, S. (2014). A review on production, storage of hydrogen, and its utilization as an energy resource. *J. Ind. Eng. Chem.* 20, 1148–1156. doi: 10.1016/j.jiec.2013.07.037

El-Emam, R. S., and Özcan, H. (2019). A comprehensive review on the techno-economics of sustainable large-scale clean hydrogen production. *J. Clean. Prod.* 220, 593–609. doi: 10.1016/j.jclepro.2019.01.309

Gorre, J., Ruoss, F., Karjunen, H., Schaffert, J., and Tynjäl,ä, T. (2020). Cost benefits of optimizing hydrogen storage and methanation capacities for Power-to-Gas plants in dynamic operation. *Appl. Energy* 257, 113967. doi: 10.1016/j.apenergy.2019.113967

Greiner, C. J., KorpÅs, M., and Holen, A. T. (2007). A Norwegian case study on the production of hydrogen from wind power. *Int. J. Hydrog. Energy*. 32, 1500–1507. doi: 10.1016/j.ijhydene.2006.10.030

Guinot, B., Champel, B., Montignac, F., Lemaire, E., Vannucci, D., Sailler, S., et al. (2015a). Techno-economic study of a PV-hydrogen-battery hybrid system for off-grid power supply: impact of performances' aging on optimal system sizing and competitiveness. *Int. J. Hydrog. Energy* 40, 623–632. doi: 10.1016/j.ijhydene.2014.11.007

Guinot, B., Montignac, F., Champel, B., and Vannucci, D. (2015b). Profitability of an electrolysis based hydrogen production plant providing grid balancing services. *Int. J. Hydrog. Energy* 40, 8778–8787. doi: 10.1016/j.ijhydene.2015.05.033

Haghi, E., Raahemifar, K., and Fowler, M. (2018). Investigating the effect of renewable energy incentives and hydrogen storage on advantages of stakeholders in a microgrid. *Energy Policy* 113, 206–222. doi: 10.1016/j.enpol.2017.10.045

Hamayun, M. H., Hussain, M., Maafa, I. M., and Aslam, R. (2019). Integration of hydrogenation and dehydrogenation system for hydrogen storage and electricity generation – simulation study. *Int. J. Hydrog. Energy* 44, 20213–20222. doi: 10.1016/j.ijhydene.2019.06.053

Hanley, E. S., Deane, J. P., and Gallachóir, B. P. Ó. (2018). The role of hydrogen in low carbon energy futures–a review of existing perspectives. *Renew. Sustain. Energy Rev.* 82, 3027–3045. doi: 10.1016/j.rser.2017.10.034

Hosseini, S. E., and Wahid, M. A. (2016). Hydrogen production from renewable and sustainable energy resources: promising green energy carrier for clean development. *Renew. Sustain. Energy Rev.* 57, 850–866. doi: 10.1016/j.rser.2015.12.112

IEA. (2016). *Medium-Term Market Report 2016 Market Analysis and Forecasts to 2021. 278*. Available online at: https://www.iea.org/publications/freepublications/publication/MTRMR2016.pdf (accessed June 28, 2021).

International Energy Agency. (2015). *Energy Technology Perspectives 2015. International Energy Agency*. Available online at: http://www.iea.org/termsandconditionsuseandcopyright/%5Cnhttp://www.oecd-ilibrary.org/energy/energy-technology-perspectives-2015_energy_tech-2015-en (accessed May 7, 2021).

Jamil, M., Kirmani, S., and Rizwan, M. (2012). Techno-economic feasibility analysis of solar photovoltaic power generation: a review. *Smart Grid Renew. Energy* 03, 266–274. doi: 10.4236/sgre.2012.34037

Jamshidi, M., and Askarzadeh, A. (2019). Techno-economic analysis and size optimization of an off-grid hybrid photovoltaic, fuel cell and diesel generator system. *Sustain. Cities Soc.* 44, 310–320. doi: 10.1016/j.scs.2018.10.021

Jiang, Y., Deng, Z., and You, S. (2019). Size optimization and economic analysis of a coupled wind-hydrogen system with curtailment decisions. *Int. J. Hydrog. Energy* 44, 19658–19666. doi: 10.1016/j.ijhydene.2019.06.035

Khunathorncharoenwong, N., Charoensuppanimit, P., Assabumrungrat, S., and Kim-Lohsoontorn, P. (2020). Techno-economic analysis of alternative processes for alcohol-assisted methanol synthesis from carbon dioxide and hydrogen. *Int. J. Hydrog. Energy* 46, 24591–24606. doi: 10.1016/j.ijhydene.2020.01.230

Kim, S., Ryi, S. K., and Lim, H. (2018). Techno-economic analysis (TEA) for CO2 reforming of methane in a membrane reactor for simultaneous CO2 utilization and ultra-pure H2 production. *Int. J. Hydrog. Energy* 43, 5881–5893. doi: 10.1016/j.ijhydene.2017.09.084

König, D. H., Freiberg, M., Dietrich, R. U., and Wörner, A. (2015). Techno-economic study of the storage of fluctuating renewable energy in liquid hydrocarbons. *Fuel* 159, 289–297. doi: 10.1016/j.fuel.2015.06.085

Lee, H., Lee, B., Byun, M., and Lim, H. (2020). Economic and environmental analysis for PEM water electrolysis based on replacement moment and renewable electricity resources. *Energy Convers. Manag.* 224, 113477. doi: 10.1016/j.enconman.2020.113477

Li, G., Liu, Z., Liu, T., Shan, J., Fang, Y., and Wang, Z. (2018). Techno-economic analysis of a coal to hydrogen process based on ash agglomerating fluidized bed gasification. *Energy Convers. Manag.* 164, 552–559. doi: 10.1016/j.enconman.2018.03.035

Liu, J., Sun, W., and Harrison, G. P. (2020). The economic and environmental impact of power to hydrogen/power to methane facilities on hybrid power-natural gas energy systems. *Int. J. Hydrog. Energy* 45, 20200–20209. doi: 10.1016/j.ijhydene.2019.11.177

Martin, S., Albrecht, F. G., Van Der Veer, P., Lieftink, D., and Dietrich, R. U. (2016). Evaluation of on-site hydrogen generation *via* steam reforming of biodiesel: Process optimization and heat integration. *Int. J. Hydrog. Energy* 41, 6640–6652. doi: 10.1016/j.ijhydene.2016.02.138

Martínez-Salazar, A. L., Melo-Banda, J. A., Coronel-García, M. A., García-Vite, P. M., Martínez-Salazar, I., and Domínguez-Esquivel, J. M. (2019). Technoeconomic analysis of hydrogen production *via* hydrogen sulfide methane reformation. *Int. J. Hydrog. Energy* 44, 12296–12302. doi: 10.1016/j.ijhydene.2018.11.023

Nieminen, H., Laari, A., and Koiranen, T. (2019). CO2 hydrogenation to methanol by a liquid-phase process with alcoholic solvents: A techno-economic analysis. *Processes* 7, 1–24. doi: 10.3390/pr7070405

Nurdiawati, A., Zaini, I. N., Irhamna, A. R., Sasongko, D., and Aziz, M. (2019). Novel configuration of supercritical water gasification and chemical looping for highly-efficient hydrogen production from microalgae. *Renew. Sustain. Energy Rev.* 112, 369–381. doi: 10.1016/j.rser.2019.05.054

Okonkwo, P. C., Ige, O. O., Barhoumi, E. M., Uzoma, P. C., Emori, W., Benamor, A., et al. (2021). Platinum degradation mechanisms in proton exchange membrane fuel cell (PEMFC) system: a review. *Int. J. Hydrog. Energy* 46, 15850–15865. doi: 10.1016/j.ijhydene.2021.02.078

Olateju, B., Monds, J., and Kumar, A. (2014). Large scale hydrogen production from wind energy for the upgrading of bitumen from oil sands. *Appl. Energy* 118, 48–56. doi: 10.1016/j.apenergy.2013.12.013

Ozbilen, A., Dincer, I., and Rosen, M. A. (2011). A comparative life cycle analysis of hydrogen production *via* thermochemical water splitting using a Cu-Cl cycle. *Int. J. Hydrog. Energy* 36, 11321–11327. doi: 10.1016/j.ijhydene.2010.12.035

Rajendran, K., and Murthy, G. S. (2019). Techno-economic and life cycle assessments of anaerobic digestion – a review. *Biocatal. Agric. Biotechnol.* 20, 101207. doi: 10.1016/j.bcab.2019.101207

Ren, J., Manzardo, A., Toniolo, S., and Scipioni, A. (2013). Sustainability of hydrogen supply chain. Part I: Identification of critical criteria and cause-effect analysis for enhancing the sustainability using DEMATEL. *Int. J. Hydrog. Energy* 38, 14159–14171. doi: 10.1016/j.ijhydene.2013.08.126

Rivarolo, M., Bellotti, D., Magistri, L., and Massardo, A. F. (2016). Feasibility study of methanol production from different renewable sources and thermo-economic analysis. *Int. J. Hydrog. Energy* 41, 2105–2116. doi: 10.1016/j.ijhydene.2015.12.128

Rivera-Tinoco, R., Farran, M., Bouallou, C., Auprêtre, F., Valentin, S., Millet, P., et al. (2016). Investigation of power-to-methanol processes coupling electrolytic hydrogen production and catalytic CO2 reduction. *Int. J. Hydrog. Energy* 41, 4546–4559. doi: 10.1016/j.ijhydene.2016.01.059

Santarelli, M., and Macagno, S. (2004). A thermoeconomic analysis of a PV-hydrogen system feeding the energy requests of a residential building in an isolated valley of the Alps. *Energy Convers. Manag.* 45, 427–451. doi: 10.1016/S0196-8904(03)00156-0

Sarkar, J., and Bhattacharyya, S. (2012). Application of graphene and graphene-based materials in clean energy-related devices Minghui. *Arch. Thermodyn.* 33, 23–40. doi: 10.2478/v10173-012-0026-8

Scherer, G. W. H., Newson, E., and Wokaun, A. (1999). Economic analysis of the seasonal storage of electricity with liquid organic hydrides. *Int. J. Hydrog. Energy* 24, 1157–1169. doi: 10.1016/S0360-3199(98)00177-3

Schlachtberger, D. P., Brown, T., Schramm, S., and Greiner, M. (2017). The benefits of cooperation in a highly renewable European electricity network. *Energy* 134, 469–481. doi: 10.1016/j.energy.2017.06.004

Schnuelle, C., Wassermann, T., Fuhrlaender, D., and Zondervan, E. (2020). Dynamic hydrogen production from PV and wind direct electricity supply – Modeling and techno-economic assessment. *Int. J. Hydrog. Energy* 45, 29938–29952. doi: 10.1016/j.ijhydene.2020.08.044

Shabani, B., and Andrews, J. (2011). An experimental investigation of a PEM fuel cell to supply both heat and power in a solar-hydrogen RAPS system. *Int. J. Hydrog. Energy* 36, 5442–5452. doi: 10.1016/j.ijhydene.2011.02.003

Shaner, M. R., Atwater, H. A., Lewis, N. S., and McFarland, E. W. (2016). A comparative technoeconomic analysis of renewable hydrogen production using solar energy. *Energy Environ. Sci.* 9, 2354–2371. doi: 10.1039/C5EE02573G

Shiroudi, A., Taklimi, S. R. H., Mousavifar, S. A., and Taghipour, P. (2013). Stand-alone PV-hydrogen energy system in Taleghan-Iran using HOMER software: optimization and techno-economic analysis. *Environ. Dev. Sustain.* 15, 1389–1402. doi: 10.1007/s10668-013-9443-3

Silva, S. B., de Oliveira Marco, M. A. G., and Severino, M. M. (2010). Economic evaluation and optimization of a photovoltaic-fuel cell-batteries hybrid system for use in the Brazilian Amazon. *Energy Policy* 38, 6713–6723. doi: 10.1016/j.enpol.2010.06.041

Staffell, I., Scamman, D., Velazquez Abad, A., Balcombe, P., Dodds, P. E., Ekins, P., et al. (2019). The role of hydrogen and fuel cells in the global energy system. *Energy Environ. Sci.* 12, 463–491. doi: 10.1039/C8EE01157E

Stojković, S. M., and Bakić, V. V. (2016). Techno-economic analysis of stand-alone photovoltaic/wind/battery/hydrogen systems for very Small-Scale applications. *Therm. Sci.* 20, S261–S273. doi: 10.2298/TSCI15030 8195S

Thürer, M., Tomaševi,ć, I., Stevenson, M., Qu, T., and Huisingh, D. (2018). A systematic review of the literature on integrating sustainability into engineering curricula. *J. Clean. Prod.* 181, 608–617. doi: 10.1016/j.jclepro.2017.12.130

Touili, S., Alami Merrouni, A., Azouzoute, A., El Hassouani, Y., and Amrani, A.-I. (2018). A technical and economical assessment of hydrogen production potential from solar energy in Morocco. *Int. J. Hydrog. Energy* 43, 22777–22796. doi: 10.1016/j.ijhydene.2018.10.136

Tranfield, D., Denyer, D., and Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review* introduction: the need for an evidence- informed approach. *Br. J. Manag.* 14, 207–222. doi: 10.1111/1467-8551.00375

Tsatsaronis, G., Kapanke, K., and María Blanco Marigorta, A. (2008). Exergoeconomic estimates for a novel zero-emission process generating hydrogen and electric power. *Energy* 33, 321–330. doi: 10.1016/j.energy.2007.10.007

Tzamalis, G., Zoulias, E. I., Stamatakis, E., Parissis, O. S., Stubos, A., and Lois, E. (2013). Techno-economic analysis of RES and hydrogen technologies integration in remote island power system. *Int. J. Hydrog. Energy* 38, 11646–11654. doi: 10.1016/j.ijhydene.2013.03.084

Tzamalis, G., Zoulias, E. I., Stamatakis, E., Varkaraki, E., Lois, E., and Zannikos, F. (2011). Techno-economic analysis of an autonomous power system integrating hydrogen technology as energy storage medium. *Renew. Energy* 36, 118–124. doi: 10.1016/j.renene.2010.06.006

van der Spek, M., Fout, T., Garcia, M., Kuncheekanna, V. N., Matuszewski, M., McCoy, S., et al. (2020). Uncertainty analysis in the techno-economic assessment of CO2 capture and storage technologies. Critical review and guidelines for use. *Int. J. Greenh. Gas Control* 100, 103113. doi: 10.1016/j.ijggc.2020.103113

Walker, S. B., Van Lanen, D., Fowler, M., and Mukherjee, U. (2016). Economic analysis with respect to Power-to-Gas energy storage with consideration of various market mechanisms. *Int. J. Hydrog. Energy* 41, 7754–7765. doi: 10.1016/j.ijhydene.2015.12.214

Wang, T., Li, Q., Chen, W., and Liu, T. (2017). "Application of energy management strategy based on state machine in fuel cell hybrid power system," in *2017 IEEE Transportation Electrification Conference and Expo, Asia-Pacific, ITEC Asia-Pacific*, Harbin.

Wang, Y., Li, G., Liu, Z., Cui, P., Zhu, Z., and Yang, S. (2019). Techno-economic analysis of biomass-to-hydrogen process in comparison with coal-to-hydrogen process. *Energy* 185, 1063–1075. doi: 10.1016/j.energy.2019.07.119

Wang, Z., Zhang, X., and Rezazadeh, A. (2021). Hydrogen fuel and electricity generation from a new hybrid energy system based on wind and solar energies and alkaline fuel cell. *Energy Rep.* 7, 2594–2604. doi: 10.1016/j.egyr.2021.04.060

Yang, H., Li, Q., Wang, T., Qiu, Y., and Chen, W. (2019). A dual mode distributed economic control for a fuel cell– photovoltaic-battery hybrid power generation system based on marginal cost. *Int. J. Hydrog. Energy* 44, 25229–25239. doi: 10.1016/j.ijhydene.2019.02.180

Yao, J., Kraussler, M., Benedikt, F., and Hofbauer, H. (2017). Techno-economic assessment of hydrogen production based on dual fluidized bed biomass steam gasification, biogas steam reforming, and alkaline water electrolysis processes. *Energy Convers. Manag.* 145, 278–292. doi: 10.1016/j.enconman.2017.04.084

Ye, B., Zhang, K., Jiang, J. J., Miao, L., and Li, J. (2017). Towards a 90% renewable energy future: a case study of an island in the South China Sea. *Energy Convers. Manag.* 142, 28–41. doi: 10.1016/j.enconman.2017.03.038

Yodwong, B., Guilbert, D., Phattanasak, M., Kaewmanee, W., Hinaje, M., and Vitale, G. (2020). Proton exchange membrane electrolyzer modeling for power electronics control: a short review. *C — J. Carbon Res.* 6, 29. doi: 10.3390/c6020029

Yukesh Kannah, R., Kavitha, S., Preethi, Parthiba Karthikeyan, O., Kumar, G., Dai-Viet, N. V., et al. (2021). Techno-economic assessment of various hydrogen production methods – a review. *Bioresour. Technol.* 319, 124175. doi: 10.1016/j.biortech.2020.124175

Zahid, A. H., Amin, N., Nisar, F., and Saghir, S. (2020). Analysis of MTH-System (Methylcyclohexane-Toluene-Hydrogen-System) for hydrogen production as fuel for power plants. *Int. J. Hydrog. Energy* 45, 32234–32242. doi: 10.1016/j.ijhydene.2020.08.248

Zoulias, E. I., and Lymberopoulos, N. (2007). Techno-economic analysis of the integration of hydrogen energy technologies in renewable energy-based stand-alone power systems. *Renew. Energy* 32, 680–696. doi: 10.1016/j.renene.2006.02.005

frontiers | Frontiers in Sustainability

REVIEW
published: 03 October 2022
doi: 10.3389/frsus.2022.910924

Satellite Data Applications for Sustainable Energy Transitions

Morgan R. Edwards[1,2]*, Tracey Holloway[2,3], R. Bradley Pierce[3,4], Lew Blank[1], Madison Broddle[5], Eric Choi[6], Bryan N. Duncan[7], Ángel Esparza[6], Giacomo Falchetta[8,9,10], Meredith Fritz[11], Holly K. Gibbs[2,12,13], Henry Hundt[1], Tyler Lark[2,13], Amy Leibrand[11,14], Fei Liu[7,15], Becca Madsen[16], Tanya Maslak[11], Bhartendu Pandey[17], Karen C. Seto[18] and Paul W. Stackhouse Jr.[19]

[1] La Follette School of Public Affairs, University of Wisconsin–Madison, Madison, WI, United States, [2] Nelson Institute Center for Sustainability and the Global Environment, University of Wisconsin–Madison, Madison, WI, United States, [3] Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, WI, United States, [4] Space Science and Engineering Center, University of Wisconsin–Madison, Madison, WI, United States, [5] Booz Allen Hamilton, Inc., at NASA Langley Research Center, Hampton, VA, United States, [6] GHGSat Inc., Houston, TX, United States, [7] Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, United States, [8] International Institute for Applied Systems Analysis, Laxenburg, Austria, [9] Centro Euro-Mediterraneo sui Cambiamenti Climatici, Università Ca'Foscari, Venice, Italy, [10] RFF-CMCC European Institute on Economics and the Environment, Venice, Italy, [11] Battelle Memorial Institute, Columbus, OH, United States, [12] Department of Geography, University of Wisconsin–Madison, Madison, WI, United States, [13] DOE Great Lakes Bioenergy Research Center, University of Wisconsin–Madison, Madison, WI, United States, [14] Science and Technology Corporation, NASA Goddard Space Flight Center, Greenbelt, MD, United States, [15] Goddard Earth Sciences Technology and Research (GESTAR) II, Morgan State University, Baltimore, MD, United States, [16] Electric Power Research Institute, Palo Alto, CA, United States, [17] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, United States, [18] Yale School of the Environment, Yale University, New Haven, CT, United States, [19] NASA Langley Research Center, Hampton, VA, United States

OPEN ACCESS

Edited by:
Shutaro Takeda,
Kyoto University, Japan

Reviewed by:
Deepak Rajagopal,
University of California, Los Angeles, United States
Daikichi Seki,
Kyoto University, Japan

***Correspondence:**
Morgan R. Edwards
morgan.edwards@wisc.edu

Specialty section:
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

Received: 01 April 2022
Accepted: 21 June 2022
Published: 03 October 2022

Citation:
Edwards MR, Holloway T, Pierce RB, Blank L, Broddle M, Choi E, Duncan BN, Esparza Á, Falchetta G, Fritz M, Gibbs HK, Hundt H, Lark T, Leibrand A, Liu F, Madsen B, Maslak T, Pandey B, Seto KC and Stackhouse PW Jr (2022) Satellite Data Applications for Sustainable Energy Transitions. Front. Sustain. 3:910924. doi: 10.3389/frsus.2022.910924

Transitioning to a sustainable energy system poses a massive challenge to communities, nations, and the global economy in the next decade and beyond. A growing portfolio of satellite data products is available to support this transition. Satellite data complement other information sources to provide a more complete picture of the global energy system, often with continuous spatial coverage over targeted areas or even the entire Earth. We find that satellite data are already being applied to a wide range of energy issues with varying information needs, from planning and operation of renewable energy projects, to tracking changing patterns in energy access and use, to monitoring environmental impacts and verifying the effectiveness of emissions reduction efforts. While satellite data could play a larger role throughout the policy and planning lifecycle, there are technical, social, and structural barriers to their increased use. We conclude with a discussion of opportunities for satellite data applications to energy and recommendations for research to maximize the value of satellite data for sustainable energy transitions.

Keywords: energy, satellite, sustainability, decision-making, data

INTRODUCTION

Actors across the energy system – from local, state, and national governments to electric utilities, technology developers, and a wide variety of energy end users – are grappling with options to limit the rise in global temperature to well below 2°C (and preferably 1.5°C) and achieve net-zero carbon dioxide (CO_2) emissions targets (Hultman et al., 2020; Klemun et al., 2020). Meeting these ambitious goals will require far-reaching energy transitions in electricity, transportation, buildings, and industry (IPCC, 2018; Cui et al., 2019). Climate change is also shifting patterns in energy demand and increasing disruptions in energy access due to damage to infrastructure caused by extreme temperatures, floods, droughts, hurricanes, and other disasters. New sources of information are needed to support sustainable energy transitions and evaluate whether energy planning and policy decisions are effective and equitable (Carley and Konisky, 2020). By providing observations of Earth from space, satellite data hold new potential to address these global challenges.

Since the first satellite images were made publicly available in 1972, applications of satellite data have expanded significantly (Davis, 2007; Inman et al., 2013). Satellite data vary in spatial resolution (from tens of kilometers to less than a meter), frequency of observations (from weeks to minutes), and coverage (from continuous observations from geostationary satellites to global coverage from polar-orbiting satellites) (Medina-Lopez et al., 2021). There are trade-offs across these design features, with free, publicly available data from government sources tending toward global coverage and a growing number of private companies offering targeted observations of particular locations. Cloud-computing services further enhance the prospects for widespread use of satellite data by allowing broad user communities to process large amounts of data on the fly (Gorelick et al., 2017). Beyond the satellite technology itself, research has also advanced applications of satellite data to decision-making through comparisons with other data sources, integration with models, and case studies applying satellite data to particular contexts and examining barriers to use (Milford and Knight, 2017; Holloway et al., 2018)[1,2].

Decisions related to energy supply, demand, impacts, and resilience all stand to benefit from growing integration of satellite data. Satellite applications for energy supply include mapping renewable resource potential to support infrastructure siting, development, and maintenance. Applications for energy demand include assessing energy use patterns to predict future needs and identify locations with unserved demand, both on an ongoing basis and in the aftermath of power disruptions. Applications for energy impacts include monitoring the effects of energy use on climate, air quality, and water and land systems, as well as efforts to reduce these impacts. Existing information sources used in the past have often been limited in spatial coverage and accessibility for a diversity of stakeholders and decision-making needs. These stakeholders also frequently lack access to timely information needed to support cross-cutting reliability and resilience goals, as well as disaster response. Expanded use of satellite data can now help address these information gaps.

This paper reviews the current state of satellite data for energy applications and potential future directions for research. We focus specifically on satellite tools for remote sensing because of their broad scale and routine measurements, as well as their underutilized potential for energy policy and planning. Each section presents an overview of conceptual and practical applications of satellite data, drawing primarily from the peer-reviewed literature. Applications vary in their level of maturity, from well-established uses with strong links to decision frameworks to emerging areas where there are significant technical, social, and/or structural barriers to applying satellite data to decision-making. While previous work examines satellite data for various energy applications in isolation, there is significant potential to increase the value of satellite data for energy decision needs by bridging insights across energy issues. Understanding the value and potential of satellite data to address energy-related challenges is particularly salient given the speed and scale of energy transitions required to mitigate and adapt to climate change.

Abbreviations: ABI, Advanced Baseline Imager; AHI, Advanced Himawari Imager; AIMM, Alternative Approved Instrument Monitoring Method; AMEL, Alternative Means of Emission Limitation; AMSR-E and AMSR2, Advanced Microwave Scanning Radiometers; ARLs, Application Readiness Levels; ASAR, Advanced Synthetic Aperture Radar; ASTER, Advanced Spaceborne Thermal Emission and Reflection Radiometer; CEMS, Continuous Emissions Monitoring System; CERES, Clouds and Earth's Radiant Energy System; CO, Carbon monoxide; CO_2, Carbon dioxide; CO2M, Copernicus Carbon Dioxide Monitoring mission; COP26, 2021 United Nations Climate Change Conference; DMSP-OLS, Defense Meteorological Satellite Program-Operational Linescan System; DNB, VIIRS Day Night Band; DOE, U.S. Department of Energy; EOSDIS, Earth Observing System Data and Information System; EPA, U.S. Environmental Protection Agency; EPRI, Electric Power Research Institute; ESA, European Space Agency; ETM+, The Enhanced Thematic Mapper Plus; EUMETSAT, European Organization for the Exploitation of Meteorological Satellites; FIRMS, NASA's Fire Information for Resource Management System; GEMS, Geostationary Environmental Monitoring Spectrometer; GeoCarb, Geostationary Carbon Cycle Observatory; GOES, Geostationary Operational Environmental Satellite; GOME-2, Global Ozone Monitoring Experiment-2; GOSAT, Greenhouse Gas Observing Satellite; IMEO, International Methane Emissions Observatory; InSAR, Interferometric Synthetic Aperture Radar; ISS, International Space Station; LANCE, Land, Atmosphere Near-real-time Capability for EOS; LDAR, Leak Detection and Repair; MERRA-2, Modern-Era Retrospective analysis for Research and Applications, Version 2; MODIS, Moderate Resolution Imaging Spectroradiometer; MSG, Meteosat Second Generation; MSI, European Sentinel-2 MultiSpectral Instrument; MTG, European Meteosat Third Generation; NASA, National Aeronautics and Space Administration; NMVOC, Methane and Non-methane Volatile Organic Compounds; NO, Nitric oxides; NO_2, Nitrogen dioxide; NOAA, National Oceanic and Atmospheric Administration; NOAA ASCAT, National Oceanic and Atmospheric Administration Advanced Scatterometer; NO_X, Nitrogen oxides; NREL, National Renewable Energy Laboratory; NSRDB, U.S. National Solar Radiation Data Base; NTL, Nighttime Light; OCO-2, Orbiting Carbon Observatory-2; OLI, Operational Land Imager; OMI, Ozone Monitoring Instrument; POWER, NASA Prediction Of Worldwide Energy Resources project; PV, Photovoltaic; ROW, right-of-way; RSPO, Roundtable for Sustainable Palm Oil; SAR, Synthetic Aperture Radar; SCIAMACHY, SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY; SMAP, Soil Moisture Active Passive; SMMR, Scanning Multichannel Microwave Radiometer; SRON, Netherlands Institute for Space Research; SSMI, Special Sensor Microwave Imager; Suomi NPP satellite, Suomi National Polar-orbiting Partnership; TEMPO, Tropospheric Emissions: Monitoring of Pollution; TRMM, Tropical Rainfall Measuring Mission; TROPOMI, Tropospheric Monitoring Instrument; UNEP, United Nations Environment Program; UVN, UV/Visible/Near-infrared; VCD, Vertical Column Density; VIIRS, Visible Infrared Imaging Radiometer Suite; VOC, Volatile Organic Compounds; WHO, World Health Organization.

[1]EPRI. Application of Image Processing Algorithms to Improve Predictive Reliability Assessments: Identifying Physical Threats Using GIS and Satellite Imagery. https://www.epri.com/research/products/000000003002018884.

[2]EPRI. Program on Technology Innovation: Using Hyperspectral Imagery and Artificial Intelligence (AI) to Detect Stressed and Dead Trees. https://www.epri.com/research/products/000000003002022770.

The author team represents experts in a wide variety of energy and satellite topics from academia, government organizations, research institutions, and private companies. Following the introduction, we discuss satellite data applications in energy supply, energy demand, energy impacts, and energy resilience. We then describe an example of a satellite data distribution platform for energy users. We conclude with a discussion of the potential and limitations of satellite data across energy applications and recommendations for research to enhance the usefulness of satellite data for energy stakeholders.

ENERGY SUPPLY

Many studies have quantified the enormous expansion in renewable energy needed to achieve global climate policy goals (IPCC, 2022). Satellite data can support the development, deployment, and forecasting of renewable energy sources such as bioenergy, hydropower, solar photovoltaics, wind turbines, and geothermal energy. Beyond assessing the potential for new systems, satellite data can also help optimize performance and track the rate of technology adoption.

Bioenergy Resources and Production

Satellite data are a leading source of information for policy and planning decisions related to bioenergy feedstock supply and productivity. Space-based data routinely inform assessments of biofuel feedstock availability and land use impacts, as well as potential competition with food production and impacts on other ecosystem goods and services. Productivity can be quantified from satellite observations of vegetation greenness and further constrained or refined using indirect satellite-based information on climate, soil conditions, and other co-determinants of productivity. Satellite-based estimates of land availability and supply have been used by industry, policymakers, and other bioenergy stakeholders in the evaluation and design of production systems and regulations.

Data on land cover have been used to identify abandoned agricultural lands with potential to support bioenergy feedstock production (Zumkehr and Campbell, 2013; Baxter and Calvert, 2017; Goga et al., 2019; Næss et al., 2021) and to screen for land that may be deemed as marginal for food production (Nalepa and Bauer, 2012; Kang et al., 2013; Khanna et al., 2021) due to economic instability (Jiang et al., 2021), environmental sensitivity (Wang et al., 2020), and biophysical limitations in climate, soils, or topography (Gelfand et al., 2013; Gu and Wylie, 2016). For example, satellite-based productivity thresholds on low-yielding lands have been used to identify marginal areas for second generation bioenergy production (Longato et al., 2019). From local to global scales, estimates of the maximum potential production of bioenergy can support energy planning and policy (Cai et al., 2011; Smith et al., 2012; Haberl et al., 2013). Bioenergy producers or investors can also use estimates of local feedstock supply (e.g., corn) to identify locations for siting future biorefineries.

Satellite-constrained estimates of total bioenergy production potential have also been used to project the contribution that bioenergy might make toward global climate policy goals or to meet national pledges to the Paris Agreement (IPCC, 2018; Creutzig et al., 2021). Policies such as the Low Carbon Fuel Standard and the Renewable Fuel Standard in the U.S. have used satellite-based estimates of land use change associated with bioenergy to measure and regulate greenhouse gas emissions intensity associated with different bioenergy systems, as well as to determine the eligibility of various fuels in each regulation (US EPA, 2010; Leland et al., 2018). Other work has used field-level remote sensing data to analyze changes in bioenergy feedstock supply caused by these policies, finding that the U.S. Renewable Fuel Standard, for example, led to an 8.7% increase in U.S. corn cultivation (Lark et al., 2022).

New data sources and advances in data science will open the door for highly detailed and precise ground-based data to complement data from satellites. For example, parcel-level data on land ownership and sales could enable a more refined understanding of how producers respond to policy and market incentives, and productivity measurements collected directly from agricultural equipment could significantly expand data availability. Nonetheless, satellite data will continue to provide irreplaceable information on bioenergy production that covers large geographic extents in a consistent manner over time, particularly with the increased availability of high-quality, high-resolution, and low-cost commercial and small-satellite platforms.

Hydropower and Water Supply

Satellite data are commonly used in water resource assessment for planning hydropower projects, monitoring reservoir size, and evaluating the environmental impacts of rerouting or damming water. Hydrological and hydrometeorological variables, such as precipitation, snow extent, soil moisture, runoff, and evapotranspiration, influence the availability of water resources to support power generation. Hydropower currently accounts for ∼60% of global renewable electricity production and is projected to play a major role in flexible power systems as the world transitions to cleaner energy sources (International Hydropower Association, 2021). Planned hydroelectric projects also dominate the renewable energy sector in sub-Saharan Africa, where significant untapped potential exists (Stiles and Murove, 2019), offering opportunities for new uses of satellite data products (Leibrand et al., 2019). Tracking and monitoring water resources is critical to ensuring and managing future water supply, especially given projected changes in water resources due to climate change (Fletcher et al., 2019).

Landsat and Terra satellites have been collecting environmental and climate data for several decades and provide a long historical record to help identify trends and spatial patterns in river flow, snow melt, land cover, and other variables that impact water availability, which is useful in decision-making for hydropower operations (see **Figure 1** for an example). Other satellites provide data in near real time, such as the National Aeronautics and Space Administration's (NASA's) Soil Moisture Active Passive (SMAP) mission, which measures global soil moisture in increments as short as 3 hours, with a latency of 24 hours and a revisit time of 2–3 days, thus reducing the need for field evaluation.

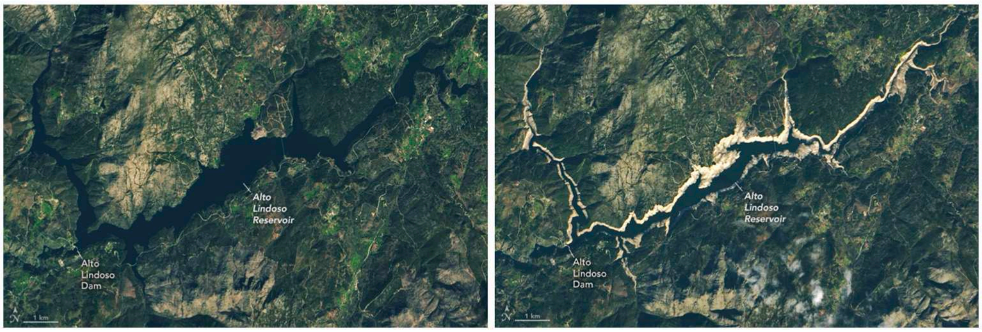


FIGURE 1 | Tracking the impact of drought on a hydropower reservoir at the Alto Lindoso Dam in Portugal from March 6, 2021 **(L)** to February 5, 2022 **(R)**, using Landsat 8 data [Credit: NASA Earth Observatory image by Lauren Dauphin, using Landsat data from the U.S. Geological Survey (NASA Earth Observatory, 2022)].

Satellite-based data on groundwater, surface water height and extent, and precipitation may be used to assess seasonal and historical changes in water storage. Freeze-thaw data derived from satellite microwave radiometry from NASA's Scanning Multichannel Microwave Radiometer (SMMR), Special Sensor Microwave/Imager (SSM/I), and Advanced Microwave Scanning Radiometers (AMSR-E and AMSR2) have been used to evaluate the dynamics of seasonal snow, ice melt, and soil thaw as a proxy for measuring water mobility over time (Kimball and McDonald, 2020). Taken together, satellite-derived hydrological and hydrometeorological data can identify trends in water availability, potential for flooding and drought, and other environmental aspects for improved decision-making in the hydropower sector.

Machine learning and data assimilation are advancing data analysis to improve observations for hydropower in areas where ground-based data are scarce. For example, machine learning has been combined with near-real-time rainfall data from NASA's Tropical Rainfall Measuring Mission (TRMM) and soil moisture data from the National Oceanic and Atmospheric Administration's (NOAA's) Advanced Scatterometer (ASCAT) to simulate streamflow in India (Kumar et al., 2021). Machine learning with various satellite-derived hydrometeorological variables has also been used to calculate streamflow in the Hanjiang River in China (He et al., 2021). Data assimilation, another approach to data fusion, has also improved land surface model predictions of water storage, particularly when multiple satellite data products are combined (Khaki et al., 2020).

There are new opportunities to use satellite data for hydropower planning and management (International Hydropower Association, 2020). NASA, NOAA, the European Space Agency (ESA), and other Earth observing organizations provide open-source data and offer training on how to apply data to real-world decisions, working to reduce barriers to use and accessibility. The value of these data is especially high in regions with gaps in ground-based data and with high climate variability, where uncertainties in water resources present challenges for hydropower planning and operations.

Solar Photovoltaic Systems

Satellite data have long been used to measure annual solar insolation in conjunction with ground-based pyranometer data (Perez et al., 2013). For example, the U.S. National Solar Radiation Data Base (NSRDB) from the National Renewable Energy Laboratory (NREL) uses data from the NOAA Geostationary Operational Environmental Satellite (GOES), NASA Moderate Resolution Imaging Spectroradiometer (MODIS) instrument, and NASA's Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) assimilation model to create a dataset that shows historical levels of solar energy resources in any location in the U.S. (Sengupta et al., 2018). The multiple source dataset goes back to 1998 at a temporal resolution of half an hour. Additionally, global solar radiation data are made available back to the early 1980's using fused geosynchronous and polar orbiting satellites, including data products available since 2020 from NASA's Clouds and Earth's Radiance Energy System (CERES) (Zhang et al., 2004; Rutan et al., 2015; Karlsson et al., 2017; Stackhouse et al., 2021).

With increased penetration of variable wind and solar power on the grid, there is a new focus on system performance and short-term wind and solar resource forcasting (Janjai et al., 2011; Pfenninger and Staffell, 2016; Peters et al., 2018). For example, machine learning has been used to predict cloud velocities to understand where drops in photovoltaic (PV) system production might occur (Cheng et al., 2022), and satellite-derived aerosol levels may be used to assess the impact of air pollution on PV arrays (see example in **Figure 2**) (Li et al., 2017). Local decision-makers can also use satellite-derived maps to inform cost-effective renewable energy project upkeep, such as vegetation management (Yu et al., 2018).

Satellite data can also be used to track renewable energy deployment, assess solar access disparities, and potentially support third party validation of renewable energy adoption

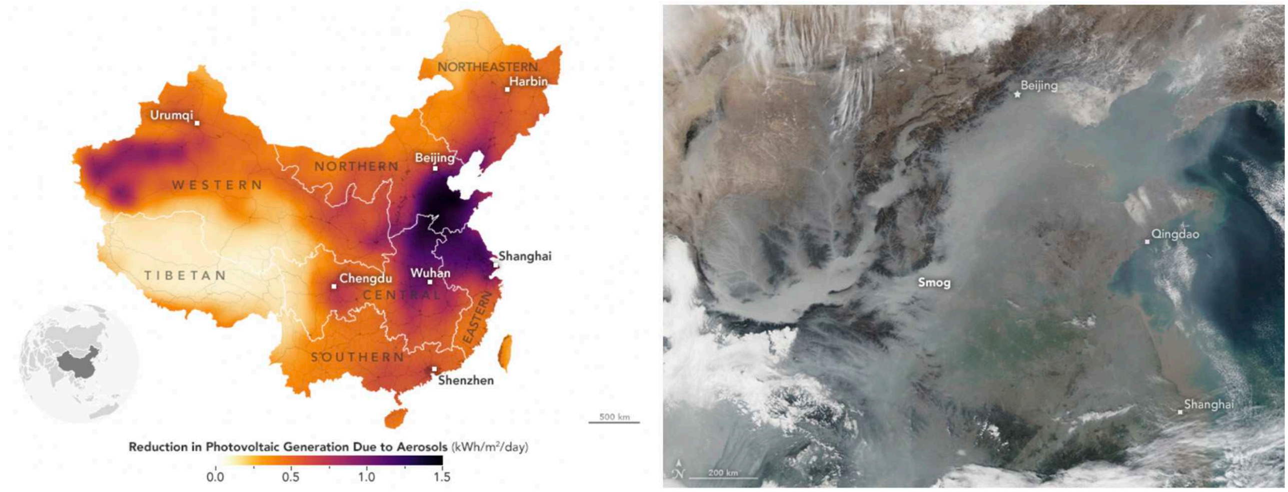


FIGURE 2 | (L) Impact of aerosols on the average amount of radiation reaching the land surface of China between 2003 and 2014 [Credit: Joshua Stevens, NASA Earth Observatory, using data from Li et al. (2017)]. **(R)** Natural-color image of haze over eastern China from the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar-orbiting Partnership (NPP) satellite on January 25, 2017. (Credit: Jeff Schmaltz, NASA, LANCE/EOSDIS).

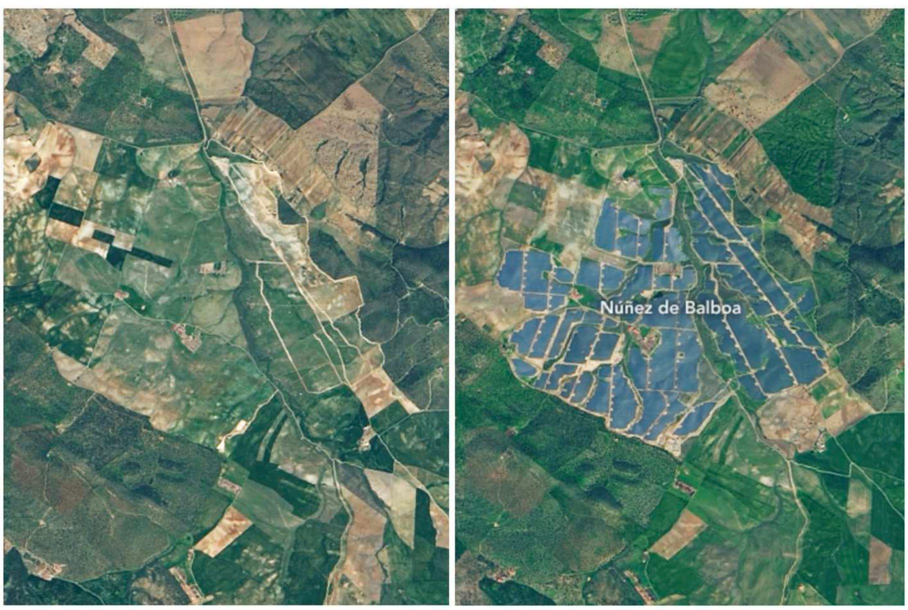


FIGURE 3 | Satellite images of a 500 MW solar power plant on the Iberian Peninsula. **(L)** shows imaging before installation in 2020, **(R)** shows imaging after installation (Credit: NASA Earth Observatory image by Lauren Dauphin, using Landsat data from the U.S. Geological Survey).

under climate agreements (see example in **Figure 3**)[3]. Standard solar PV accounting methods generally focus on limited regions and often miss smaller systems. Satellite image processing offers an efficient method for tracking growth in solar energy across large geographic areas (Kruitwagen et al., 2021), but smaller residential and microgrid systems are still difficult to track (Ishii et al., 2016).

Offshore Wind Projects

Using traditional *in-situ* measurements such as buoys to measure offshore wind resources is expensive, time consuming, and limited in its geographic coverage. As an alternative, synthetic aperture radar (SAR) data from satellites is being used to estimate wind power from wave heights and direction. Recent efforts

[3]United Nations. Net Zero Coalition. https://www.un.org/en/climatechange/net-zero-coalition.

have focused on improving the SAR method's accuracy. For example, calibrating satellite data on sea winds can help improve estimates of wind speeds (Soukissian and Papadopoulos, 2015), and advanced data analysis methods like machine learning can help predict wind energy production (Majidi Nezhad et al., 2021a). To estimate wind energy at actual wind hub heights (∼100 m), near sea-surface (∼10 m) wind readings from the ESA Envisat Advanced Synthetic Aperture Radar (ASAR) are used to extrapolate wind speeds at greater elevations (Badger et al., 2016). In areas impacted by wake effects, SAR data available from missions such as the Envisat ASAR and Sentinel 1 can also measure wind speeds (Ahsbahs et al., 2018).

Although satellites cannot "see" future winds, satellite data can be used to improve forecasts of wind resource availability for wind projects (Inman et al., 2013). In offshore applications, SAR data can be used to constrain short-term weather predictions and provide temporally and spatially expansive estimates of wind speeds and wave heights (Zen et al., 2021). Both measures are important for the design, planning, and operation phases of offshore wind projects, including efficiently screening for promising offshore wind resource areas and reducing uncertainty around installation weather windows. Future areas for research include improving the spatial resolution of wave and wind detection, as the current practice is to assume similar conditions across an entire wind farm based on a limited set of estimates (Medina-Lopez et al., 2021). Additionally, inter-hour offshore wind resource forecasting is becoming more critical as coastal power grids rely on greater penetration of offshore turbines, which recent satellite products, such as ESA's Aeolus mission, will help improve (Medina-Lopez et al., 2021).

Geothermal Energy

Satellite data is supporting the exploration and monitoring of geothermal energy sources, which have the potential to provide non-emitting baseload power (Vargas et al., 2022). Remotely sensed thermal infrared data has been used since the 1980's to detect geothermal activity and identify potential sites for geothermal plants, providing a less costly data source than field investigations (Majidi Nezhad et al., 2021b). Thermal infrared bands that are sensitive to surface temperatures are used to identify anomalies that are potentially the result of subsurface geothermal activity. Instruments that have been used for geothermal prospecting include MODIS, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Landsat's Enhanced Thematic Mapper Plus (ETM+), and SAR (Howari, 2015). For example, one recent study used ASTER data to map geothermal potential along a section of the East African Rift System, where previous mapping coverage was limited, using a combination of surface temperature estimates and indicator minerals (Hewson et al., 2020).

The coarse resolution of thermal sensors provides a means to target field activities but limits their usefulness to broader scale detection of geothermal anomalies. However, satellite data can be useful for studying geothermal potential and ground temperature recovery because of the ability to construct long-term datasets. A key tool that allows for this type of analysis is the Interferometric Synthetic Aperture Radar (InSAR) technique, which can map ground deformation through clouds and at night, providing expansive temporal and spatial coverage (Mellors et al., 2018; Majidi Nezhad et al., 2021b). For example, two years of Sentinel-1 SAR data was used to analyze Iceland's untapped geothermal energy, as well as pressure changes from geothermal fluid extraction for a new power plant (Receveur et al., 2019). Future research can look to relate satellite-derived prospecting with existing geothermal data (or exploratory drilling) to improve data relevance to future geothermal applications (Howari, 2015).

ENERGY DEMAND

Tracking energy demand, both temporally and spatially, is critical to a just and sustainable energy transition. Nighttime lights (NTL) data have been actively used to monitor energy use and electrification and identify gaps for further policy development. With 770 million people worldwide without access to electricity, and many others lacking reliable and affordable heat and power (Hernández, 2015; Reames, 2016; IEA, 2021c), NTL data may be the most important data product to inform decisions to support energy access and restoration.

Energy Use and Infrastructure

Nighttime lights are a widely used indicator of energy use and infrastructure (NASA Earthdata, 2021) and have been correlated with economic activity, urbanization, population density, and energy consumption and access (Falchetta and Noussan, 2019). There are two principal datasets that provide NTL. The first digital NTL dataset is available from 1992-2013 through the Defense Meteorological Satellite Program-Operational Linescan System (DMSP-OLS). However, each pixel in these images has only 64 potential values, a consequence of the 6-bit radiometric resolution of the satellite instrument. Due to this limited range, the data become saturated when NTL levels are high, especially in urban areas, limiting NTL applications to planning and policy at the city scale. Limited low-light detection also curtails NTL utility in dimly lit regions such as rural areas. The 2.7 km spatial resolution further limits energy-related applications at local scales.

The second and more recent NTL dataset is developed from the VIIRS Day Night Band (DNB) onboard the Suomi-NPP satellite, launched in 2011. VIIRS NTL is a significant improvement over DMSP-OLS NTL in two ways: the spatial resolution is much improved at 750 m, and the sensor has a larger dynamic range, with improved calibration that allows for accurate measurements of very low and high intensity nighttime lights. Recent advances to harmonize the DMSP and VIIRS NTL data have made them easier to access and integrate for wider applications (Li et al., 2020a).

For scientific studies, the most robust NTL dataset is Black Marble, which uses raw VIIRS data and corrects for atmospheric and radiometric issues (Romn et al., 2018). These data are calibrated across time, validated against ground-based data, and available at daily resolution. NASA scientists are currently working on a high-definition version of Black Marble, which will allow researchers to downscale NTL data at finer spatial resolutions by integrating Landsat and Sentinel Earth observations and street-level GIS data into the Black Marble product, and thereby improve NTL visualization in dense urban areas (NASA Goddard Space Flight Center, 2021). In 2021, the World Bank created the Light Every Night dataset, which is a complete archive of all NTL data collected over the past three decades (Min et al., 2021)[4]. Higher-resolution NTL images are available via photographs taken from the International Space Station (ISS) and the private company NOKTOsat (de Miguel et al., 2014; Noktosat, 2021). Additional sources of NTL data exist, but many are not publicly accessible (Li et al., 2019).

The combination of finer spatial, radiometric, and temporal resolutions, as well as integration of new data sources and processing techniques, can provide near-real-time estimates of energy use. To evaluate energy access, NTL data may be combined with on-the-ground information from utilities, GIS data, and local knowledge of energy access (Zhao et al., 2019). Fusing satellite data products with data from mobile phones can also support assessments of energy use, energy poverty, and disaster response (Steele et al., 2017), while NTL data combined with census data, national household surveys, or meter data can help users better understand and address inequities in energy infrastructure and access at scale (Mann et al., 2016; Pandey et al., 2022). Satellites can also identify changes in energy demand, such as those associated with COVID-19 or holidays from different cultures (Román and Stokes, 2015; Elvidge et al., 2021; Stokes and Román, 2022).

There are several important limitations to the use of NTL as an indicator of energy and other socioeconomic variables to inform policy. Broadly, NTL are an imperfect proxy for energy use and access. Current satellite data products cannot accurately measure energy use at smaller scales relevant for many policy questions, such as at household, street, or neighborhood levels, especially in high-density areas (Falchetta et al., 2020a). NTL data are less accurate for measuring electrification in areas where energy supply is intermittent, as conventional uses of NTL and other satellite observations are often binary (i.e., the lights are on or off) (Dugoua et al., 2017). Streetlights, car lights, and LED lighting may also make an area appear more or less electrified than it truly is (Zhao et al., 2019). Finally, NTL data may be more appropriate for estimating energy and other variables in some regions than others (Zhu et al., 2019a). For example, in areas with fires or oil and gas flaring, NTL may reflect these sources rather than electrification.

Global Energy Access

Over the past few decades, countries around the world have made large investments to support the goal of universal energy access and improve the reliability of electricity supply (Aklin et al., 2018), yet access to electricity and modern cooking fuels and technologies remains low in some regions (World Bank, 2019). The main gaps are found in sub-Saharan Africa (570 million lacking electricity), Central and Southern Asia (103 million), and Southeast Asia (40 million) (World Bank, 2019). While these regional statistics provide a general understanding of the existing gap, it is critical to develop tools to map the geographic distribution and temporal dynamics of these populations to provide a fine-grained, up-to-date understanding of electricity access across the world.

Tracking of energy poverty and access has generally been carried out through household surveys administered by national governments and international organizations. Satellite-based NTL data can serve as a proxy for electricity access to support electrification planning, complementing traditional survey methods (see example in **Figure 4**) (Min et al., 2013; Burlig and Preonas, 2016; Dugoua et al., 2017; Fobi et al., 2018; Avtar et al., 2019). These data are often combined with data on population density and other socioeconomic indicators (Stokes and Seto, 2019; Zhao et al., 2019; Falchetta et al., 2020b). NTL data have shown that lack of electrification is most pronounced in countries where a large proportion of the population lives in dispersed, rural settlements with few resources (Doll and Pachauri, 2010). However, these data also suggest that energy access can decline in urban areas that were once more reliably electrified as utilities struggle to keep pace with increasing energy needs associated with rapid urbanization, especially peri-urban areas and informal settlements (Falchetta et al., 2020b).

Nighttime lights have also been combined with utility data to inform renewable energy and microgrid infrastructure planning, as well as electrification of essential services such as healthcare facilities (Korkovelos et al., 2019; Moner-Girona et al., 2019, 2021). If utility data are unavailable (e.g., after a natural disaster or in rural or low-resource settings), NTL data can be used as a proxy to estimate energy access (Fragkias et al., 2017). The stability of NTL radiance over time has also allowed it to be used to evaluate supply reliability and to measure the impact of hydroelectricity disruptions due to drought events (Arderne et al., 2020; Falchetta et al., 2020b). These studies seek to go beyond the binary classification of energy access and lack of access, which is crucial as energy poverty is a multi-dimensional challenge (Pelz et al., 2018, 2021). Thus, despite limitations of NTL data, its usefulness for understanding energy access continues to grow.

Sub-Saharan Africa stands to particularly benefit from the use of NTL data for electrification planning. Lack of energy access and unreliable electricity have hampered economic growth, and policymakers across the region face the challenge of expanding energy access to almost half the continent (IEA, 2019). Using NTL, population, and settlement data, one study estimated that between 2014 and 2019, 115 million people in sub-Saharan Africa gained access to electricity. However, in some cases, energy access did not equate to energy use, and some countries that had made strides in expanding access saw limited use in newly electrified households (Falchetta et al., 2019, 2020a). These studies highlight that increases in access must be accompanied by increases in generation and grid infrastructure to improve the quality and reliability of electricity that is delivered to households (Falchetta et al., 2020b).

Improvements to NTL data, primarily via increases in resolution and reductions in uncertainty as instruments and algorithms advance, will enable broader data use by

[4] World Bank - Light Every Night. *Registry of Open Data on AWS* https://registry.opendata.aws/wb-light-every-night/.

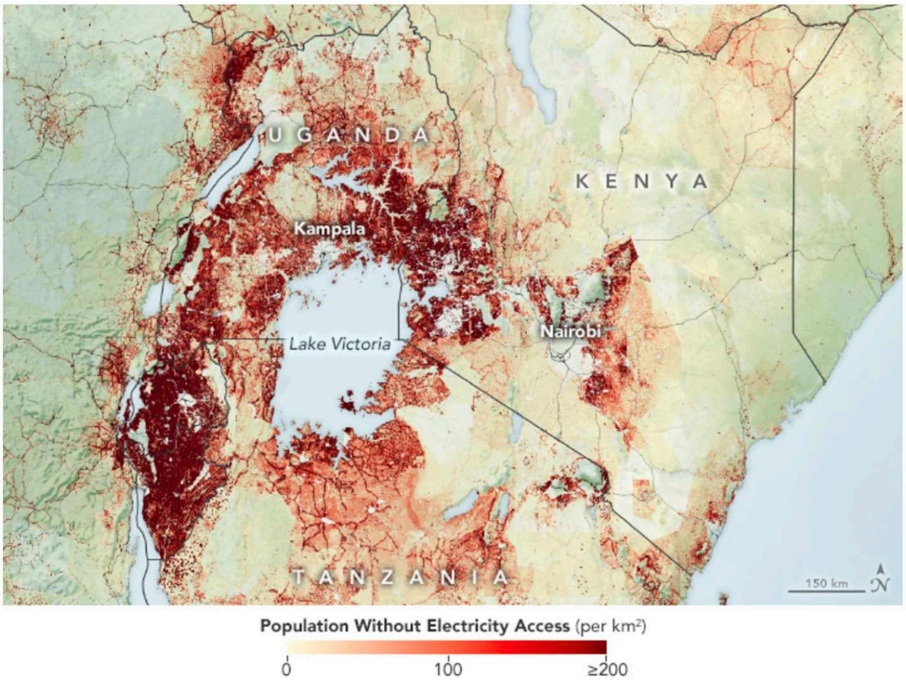


FIGURE 4 | Estimate of the number of people without access to electricity in East Africa using data from the VIIRS instrument on the NOAA-NASA Suomi NPP satellite, land cover type data from NASA's MODIS instrument, and data from Oak Ridge National Laboratory's LandScan. [Credit: NASA Earth Observatory image by Lauren Dauphin, using data from Falchetta et al. (2019, 2020a) (NASA Earth Observatory, 2021a)].

policymakers, utility managers, emergency response personnel, and other stakeholders. Usefulness of NTL data can be further improved with integration of GIS maps and geoprocessing tools (Dugoua et al., 2017). Data at finer resolutions will also increase the usability of NTL and expand applications in which it can be used. With higher spatial and radiometric resolution, and finer time scales of collection, researchers can start to examine a range of issues related to the quality and consistency of energy availability – not just whether energy infrastructure exists, but the frequency (reliability) of lighting and how quickly lighting is restored after a major disaster such as a hurricane (Romn et al., 2019), blackout, or conflict. Similarly, these data can be used to track the urban development process and to identify locations that have inadequate energy infrastructure (Stokes and Seto, 2019).

Urban Areas and Urbanization

Urban areas account for approximately 75% of global final energy use, and this demand is strongly correlated with urban form and structure (Seto et al., 2014). Therefore, characterizing urban areas can inform estimates of energy demand, even at the global scale, and can be useful in planning future energy investments to support sustainability and other goals. Urban expansion can lead to categorical changes in land cover, such as when agricultural areas become urban, as well as magnitude changes, such as urban intensification. The distinction between measurement of categorical vs. magnitude changes is important because the optimal methods and reliability of estimates differ between the two. Measuring categorical change is typically easier than measuring the magnitude of urban change.

The majority of published studies have focused on mapping two-dimensional urban expansion, or outward urban growth (Zhu et al., 2019b; Reba and Seto, 2020). It is only in the past decade that the research community began to examine volumetric growth of urban areas (see **Figure 5** for an example). Three-dimensional characterization of the built environment reveals more about urban form, structure, and resource use, such as the demand for reinforced steel and concrete or embodied and operational energy use. Backscatter data from the QuikSCAT SeaWinds scatterometer have been shown to be able to characterize urban volumetric infrastructure growth for large cities (Frolking et al., 2013; Creutzig et al., 2016; Mahtta et al., 2019; Li et al., 2020b). The recent development of a time series with ERS, QuikSCAT, and ASCAT backscatter data covering three decades will enable new studies of urban built structures and their energy implications (Frolking et al., 2022).

A recent review of algorithms to detect, characterize, and monitor urban land changes found that most methods have been developed and applied for only a few regions (e.g., the U.S. and China), with 75% of studies focused on high-income or upper-middle-income countries (Reba and Seto, 2020). Furthermore, while 11% of the world's urban population lives in cities with populations greater than 5 million, 41% of studies have focused on these very large cities, whereas most future urban growth

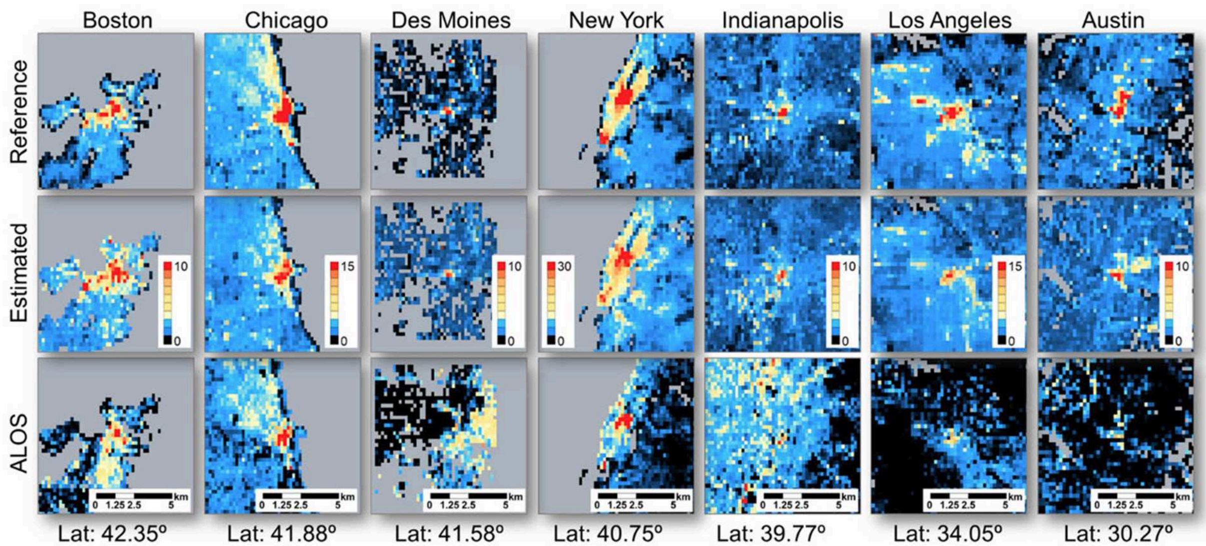


FIGURE 5 | Estimated spatial patterns in building heights in seven U.S. cities using Sentinel-1 data, compared to reference (non-satellite) measurements and Advanced Land Observing Satellite (ALOS) data (unit: m). Measurements are presented at 500 m resolution with a spatial extent of 20 × 20 km. Areas outside the study domain are colored as gray. [Credit: Li et al. (2020b)].

will occur in towns and cities with populations of less than 1 million (Seto et al., 2014). Applying satellite data to urban growth at smaller scales could support urban planning and policies related to these growing sources of urban energy demand and resource use.

ENERGY IMPACTS

Energy impacts on land, water, and air quality have long histories of regulation and management in environmental policy. Even for these well-established contexts, satellite data introduce new opportunities and challenges in connecting with decision frameworks. Extending the relevance of satellite data to greenhouse gas emissions is a growing area of research, recognizing the complexities in connecting space-based detection of gases with on-the-ground decision needs (Esparza and Gauthier, 2021).

Land and Water Impacts

Land use and water represent two of the largest impacts of crop-based bioenergy production as well as mining and other infrastructure for fossil fuels and nuclear energy. Satellites offer the potential to track and monitor energy-related impacts on land and water, which are key to successful resource management and disaster response. Because bioenergy systems rely on large amounts of biomass feedstocks, typically grown on land, they can result in particularly large land and water impacts. These impacts include direct changes in land use as well as indirect impacts via price effects that lead to expansion or contraction of crops used for biofuels or other purposes. Land use associated with bioenergy systems can also have ensuing consequences for biodiversity, water quality and use, and CO_2 emissions (Berndes et al., 2013; Popp et al., 2014). Satellite data can inform assessments of these impacts, as well as emissions from bioenergy and fossil fuel infrastructure, including refineries and power plants.

Satellite data have been instrumental in tracking patterns in land use and land cover change associated with existing bioenergy development. For example, the expansion of corn ethanol production in the U.S. has led to increases in corn cultivation, with satellite data being used to monitor resulting changes in crop rotations, land conversion, and participation in land conservation programs (Brown et al., 2014; Motamed et al., 2016; Wright et al., 2017). These changes may also contribute to shifts in water quality, which can be monitored directly by satellites or modeled using satellite data on land use, climate, and other environmental determinants (Haag et al., 2009; Hendricks et al., 2014). Similarly, satellite data have been used to track the expansion of palm oil, intended for biodiesel and other market uses, across the tropics (Koh et al., 2011; Carlson et al., 2012). These data have helped identify solutions to stymie the widespread environmental consequences of palm oil on rainforests, biodiversity, and local communities (Rose et al., 2015; Leidner and Buchanan, 2018; Meijaard et al., 2020).

Satellite data can also support interventions to minimize the environmental impacts of energy infrastructure on natural habitats and existing land conditions. For example, a pilot study from the Electric Power Research Institute (EPRI) tested the use of satellite data in identifying the effects of energy infrastructure on monarch butterfly habitats and wetlands (Madsen, 2021)[5].

[5]EPRI Program on Technology Innovation: New Frontiers in Milkweed Detection — Evaluating the Potential of Satellite Data and Machine Learning. https://www.epri.com/research/products/000000003002016599.

Another analysis used satellite data to monitor impacts along Azerbaijan gas and oil pipeline right-of-ways (ROW) spanning 10 million square miles (Bayramov, 2013).

Satellite data have also been used to assess the impacts of energy systems on water quality, particularly those arising from thermal power plants (i.e., bioenergy, fossil fuels, and nuclear). For example, studies have used satellite data to estimate water demand (Luo et al., 2018), monitor thermal discharge from power plants (Wu et al., 2007), detect turbidity (Alkan, 2009), and estimate water quality impacts (Sridhar and Vincent, 2009). Remotely sensed data have also been used as inputs to advanced modeling and prediction of water quality outcomes from bioenergy production. Off the coast of the U.S., the size of the Gulf of Mexico hypoxic zone, the world's second largest oxygen-depleted "dead zone" (Dybas, 2005), can be tracked and modeled using satellite data (Haag et al., 2009), and contributions from changes in bioenergy-related land use can be estimated using satellite-based inputs (Donner and Kucharik, 2008; Hendricks et al., 2014). Similarly, satellite data have been used to track the size and occurrence of harmful algal blooms (Klemas, 2012; Shen et al., 2012) and estimate the contribution of bioenergy to water quality impairments (Hamada et al., 2015; Lin et al., 2015; Chen et al., 2017). Satellite data have also been critical in real-time monitoring of oil spills (see **Figure 6**).

Satellite-based assessments can inform water resources conservation and planning for energy and other uses (Bastiaanssen et al., 2012). For example, analyses of evapotranspiration can inform estimates of potential water use associated with bioenergy feedstock production (Bhattarai et al., 2017; Wagle et al., 2017). These estimates can also be compared to the evapotranspiration of alternative (e.g., food) crops or native ecosystems, and inform assessments of the overall water use intensity of bioenergy feedstocks relative to other energy systems and land uses (Sanders and Masri, 2016).

The ability of satellites to capture frequent observations of changes in land and water use creates exceptional opportunities to evaluate the causal outcomes of energy policies, many of which began after routine satellite data collection (Blackman, 2013; Donaldson and Storeygard, 2016). Publicly available, space-based data can provide transparency and credibility for certification schemes that go beyond industry-reported results. For example, the Roundtable for Sustainable Palm Oil (RSPO) certification schemes rely on satellite technology to strengthen fire prevention efforts and protect forests (RSPO, 2021). Bonsucro's certification scheme for sugarcane production, which is used as a feedstock for ethanol, also relies on satellite data to map changes in land use (Bonsucro, 2021).

Health and Air Quality Impacts

A wide range of gas and particle species are emitted from fossil fuel combustion in the energy system, especially nitrogen oxides (NO_x), carbon monoxide (CO), and sulfur dioxide (SO_2), as well as suspended liquid and solid particles, referred to as particulate matter (PM). These traditional air pollutants represent the most direct linkage between energy policy and health outcomes. The World Health Organization (WHO) estimates that 92% of the global population lives in areas where air quality levels exceed WHO limits (World Health Organization, 2016), and 4.2 million people die each year due to outdoor air pollution[6].

In many ways, the experience of the air quality and health communities serves as a success story for satellite data integration into existing energy-related decision frameworks (Holloway et al., 2021). As satellite technology advanced to detect gases and particles in the atmosphere, early research highlighted the potential for these datasets to inform model evaluation, support improved emission inventories, and assess surface abundance of health-relevant pollutants. In 2011, NASA launched the first Applied Sciences team around the theme of air quality (Jacob et al., 2014), which was expanded to address health and air quality in 2016 (Holloway et al., 2018) and renewed in 2021. The three generations of these teams represent a systematic research and outreach enterprise, wherein applied research projects have advanced rapidly over the past 10 years, in collaboration with stakeholder partners.

These experiences highlight key areas where satellite data can inform energy-related air quality and health issues (World Health Organization, 2016). Nitrogen dioxide (NO_2) has emerged as perhaps the most useful air quality indicator from satellites, which has been used as an indicator of NO_x emissions, including trends in NO_x emissions associated with emission controls on power plants as well as transportation patterns, fuel shifts, and economic changes. As an example, satellite NO_2 from the TROPOMI instrument was used as an indicator of energy use changes during the early stage of the COVID-19 lockdowns in early 2020 (see **Figure 7**) (NASA Earth Observatory, 2020).

Because most NO_2 in the troposphere is emitted near the surface, the column abundance detected by satellites is well-correlated with concentrations detected by ground monitors (Goldberg et al., 2021). Furthermore, the short atmospheric lifetime of NO_2 limits mixing of the pollutant in the atmosphere, such that satellite images capture the sources of emissions and track closely with spatial patterns in combustion activities at the ground level. Satellite NO_2 has been used to evaluate health outcomes from NO_2 (Anenberg et al., 2022) and environmental justice dimensions of air pollution exposure (Kerr et al., 2021). NO_2 is also a key ingredient in ozone production near the surface, and thus an important factor in ozone control strategies (Duncan et al., 2010; Witman et al., 2014).

Many other chemical species observed from space bear relevance to energy emissions, air quality, and health. For example, satellite-derived SO_2 can be an important indicator of power plant emissions (Lu et al., 2013), satellite observations of CO show the impact of global pollution transport (NASA, 2015), and satellite observations of "aerosol optical depth" have been used quantify global exposure to fine PM (van Donkelaar et al., 2010). Beyond tracking fuel combustion, satellite data have been used to assess upstream emissions from energy processes, such as dust impacts of cropland expansion from bioenergy (Lambert et al., 2020) and air emissions associated with the pre-harvest sugarcane field

[6]World Health Organization. Air pollution. https://www.who.int/westernpacific/health-topics/air-pollution.

FIGURE 6 | Satellites were able to spot an oil slick from a major oil spill in Southern California in 2021. **(L)** is an image from October 2, 2021, from OLI on Landsat 8, and **(R)** is a SAR image from the ESA Sentinel-1B satellite. [Credit: NASA Earth Observatory image by Joshua Stevens, using Landsat data from the U.S. Geological Survey and modified Copernicus Sentinel data processed by the ESA (NASA Earth Observatory, 2021b)].

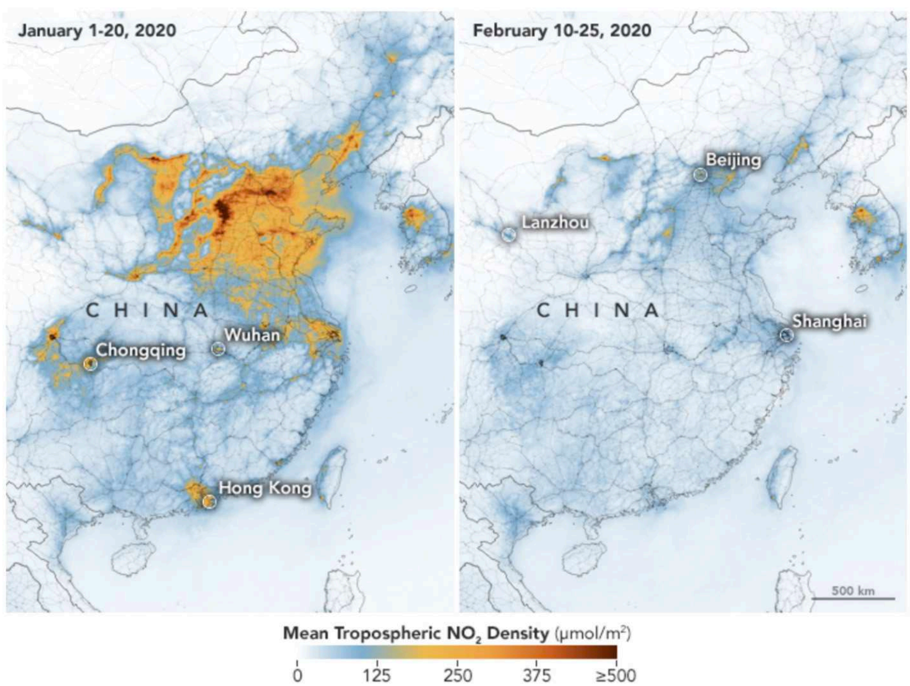


FIGURE 7 | Changes in NO_2 concentration due to COVID-19 lockdown in China using data collected from TROPOMI from ESA's Sentinel-5P satellite (Credit: NASA Earth Observatory images by Joshua Stevens, using modified Copernicus Sentinel 5P data processed by the ESA).

burning phase of ethanol production in Brazil (Tsao et al., 2012).

Oil and Gas Emissions

Oil and gas operations release a wide range of chemical emissions, including volatile organic compounds (VOCs, associated with ozone formation and also posing direct health risks) and the powerful greenhouse gas methane. Both the federal government and some states in the U.S. are beginning to consider satellite data to assess oil and gas VOC emissions for regulatory purposes. A proposed rule in New Mexico incentivizes the use and development of new technologies for leak detection and repair (LDAR) such as remote monitoring via satellites or aircraft, aiming to increase the accuracy and speed of reporting as part of their ozone control effort. Colorado provides operators with the opportunity to submit an Alternative Approved Instrument Monitoring Method (AIMM) for identifying VOC ozone precursors. These changes provide the opportunity to include other monitoring methods as an alternative to current ground-based measurement approaches.

Methane emissions from oil and gas systems have been the subject of significant recent policy action. At the 2021 United Nations Climate Change Conference (COP26), more than 100 countries joined the U.S. and EU in launching the Global Methane Pledge, an initiative to reduce methane emissions by at least 30% from 2020 levels by 2030 (European Commission, 2021). Satellite-based inventory methodologies can play a crucial role in achieving these goals by providing timely data for monitoring and verifying country commitments. The United Nations Environment Program (UNEP) is supporting this effort through the International Methane Emissions Observatory (IMEO), which will use multiple data sources from satellites, ground-based sensors, and national and company inventories (UN Environment Programme, 2021). These data can be combined to identify and reconcile gaps and inconsistencies and enable global stakeholders to track whether emissions reductions are being achieved and take targeted action.

Satellites can monitor oil and gas infrastructure on a frequent basis with an emphasis on high-risk areas, quickly detecting very large emissions sources. The natural gas supply chain is characterized by super-emitter behavior, where a small percentage of sources are responsible for the majority of emissions. A meta-analysis of approximately 15,000 measurements from 18 individual studies in the U.S. showed that the largest 5% of methane leaks typically contribute over 50% of the total emissions by volume (Brandt et al., 2016), and similar phenomena have been observed for individual production sites (Zavala-Araiza et al., 2017) and across sources and sectors (Duren et al., 2019). A recent study also used satellite data to identify large methane releases from "ultra-emitters" worldwide (Lauvaux et al., 2022).

Policies targeting super-emitters could be a cost-effective strategy for reducing overall emissions (Ravikumar et al., 2020; Edwards et al., 2021). Multiple types of measurements can work together to assess methane emissions in a tiered system-of-systems approach, integrating space-based platforms with airborne instruments and ground sensors (Esparza and Mattson, 2021). This tiered approach can enable more complete monitoring, detection, and repair of emissions sources without the need to deploy an impracticably large number of ground-based sensors, consistent with other examples of using satellite data to complement other measurement approaches.

In recent years, greenhouse gas monitoring satellites from the private sector have complemented technology from government space agencies[7]. For example, the company GHGSat currently has three methane sensing satellites in orbit with spatial resolutions as low as 30 m, allowing for detection of point sources such as individual oil and gas wells. GHGSat has an ongoing collaboration with the Netherlands Institute for Space Research (SRON) whereby elevated methane levels detected by TROPOMI, which makes measurements in 2,600 km swaths at 7 km resolution, are followed up with high-resolution GHGSat imagery that can attribute these methane hot spots to specific facilities (European Space Agency, 2020).

In early 2019, a GHGSat satellite was imaging a natural source of methane emissions known as a mud volcano in the western part of Turkmenistan when it serendipitously discovered an enormous methane leak – assessed to have been 10,000 to 43,000 kg/h – from a compressor station at the nearby Korpezhe oil and gas field. Other nearby leaks of similar magnitudes were also identified. These were some of the largest methane leaks ever detected by satellite at the time. Archived TROPOMI data confirmed the magnitude of these emissions sources going back at least 14 months (Varon et al., 2019). GHGSat worked with the diplomatic community to identify the industrial operator and contact the relevant authorities, and for a period of time the leaks were stopped. However, in February 2021, another GHGSat satellite detected new leaks from eight natural gas pipelines and unlit flares in the Galkynysh gas field in Turkmenistan (see **Figure 8**) (Malik Naureen, 2021).

There are barriers to increasing the use of satellite data to inform policy on oil and gas emissions. For example, the EPA has had an alternative means of emission limitation (AMEL) program since 1977 (42 U. S., and Code § 7401, 1977), but the current AMEL application process is complex and requires EPA review prior to public notice and public hearing events. This complexity may limit the use of satellite data in satisfying regulatory requirements, such those targeting methane and VOC emissions from new and existing sources in the oil and gas sector.

Energy-Related CO_2 Emissions

Accurate estimates of the distribution and magnitude of CO_2 emissions from energy systems are important for improving predictions of climate change, designing policies to reduce emissions, and monitoring and verifying their effects. Historically, anthropogenic CO_2 emissions have been inferred through bottom-up approaches using reported or estimated data on fuel consumption, emission factors, and

[7]Group on Earth Observations, Climate TRACE and World Geospatial Industry Council. Greenhouse Gas Monitoring from Space: A Mapping of Capabilities Across Public, Private, and Hybrid Satellite Missions. GEO Observations Blog http://www.earthobservations.org/geo_blog_obs.php?id=533.

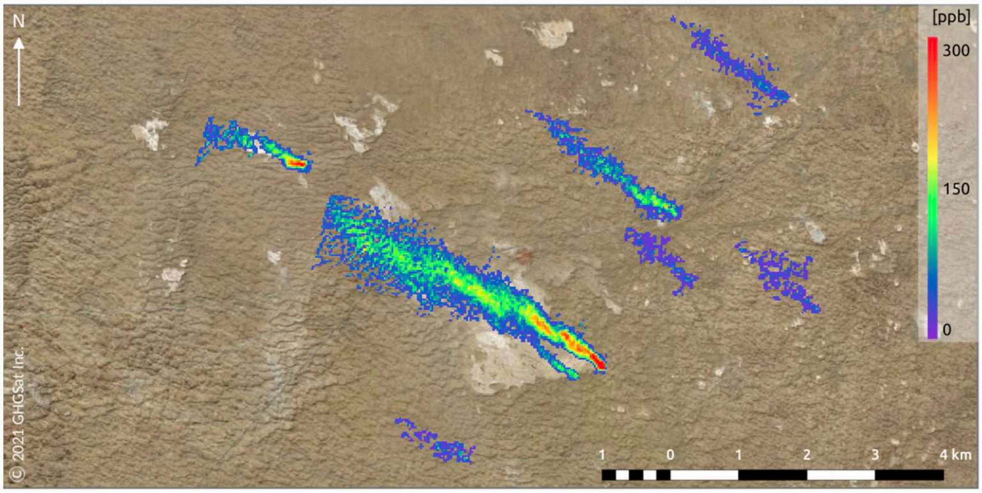


FIGURE 8 | Eight separate methane plumes captured by GHGSat in a single image, representing a total emission rate of about 10,000 kg/hour. Four of the larger plumes on the left are emissions from pipelines (likely problems with valves), with the remaining emissions from unlit flares.

fuel properties for thermal power plants, transportation, and industry. However, there are uncertainties in these data, even in high-income economies (Wheeler and Ummel, 2008; Gurney et al., 2019). For example, there is generally about a 20% difference between U.S. thermal power plant emissions estimated from fuel usage and those reported from a continuous emissions monitoring system (CEMS) program (Ackerman and Sundquist, 2008). Data uncertainties and gaps have prompted policymakers to look to satellite data to enhance tracking of greenhouse gas emissions and to monitor and verify reduction efforts.

The first space-based measurements of the global distribution of near-surface greenhouse gases were performed by an instrument called SCIAMACHY (European Space Agency, 2005), which operated aboard the ESA's Envisat satellite between 2002 and 2012. The first satellites dedicated to greenhouse gas measurements were GOSAT, launched by the Japan Aerospace Exploration Agency in 2009, and the Orbiting Carbon Observatory-2 (OCO-2), launched by NASA in 2014 (see example in **Figure 9**) (Yokota et al., 2009; Crisp, 2015). These were followed by TROPOMI aboard the ESA Sentinel-5P satellite, which has been in operation since 2017, as well as the GOSAT-2 satellite launched in 2018 and the OCO-3 instrument that was installed on the ISS in 2019. Tracking of greenhouse gas emissions with satellites is set to expand in the upcoming years: the Environmental Defense Fund (MethaneSat), the State of California (Carbon Mapper), and NASA (Geostationary Carbon Cycle Observatory, or Geocarb) are all planning launches of satellites to track emissions between 2022 and 2025 (Dennis, 2021). The growth in new dedicated satellite instrumentation, combined with existing measurements, may allow for easier independent monitoring, verification, and enforcement of the national emissions reduction commitments under Paris Agreement (Ganesan et al., 2019; Kaminski et al., 2022).

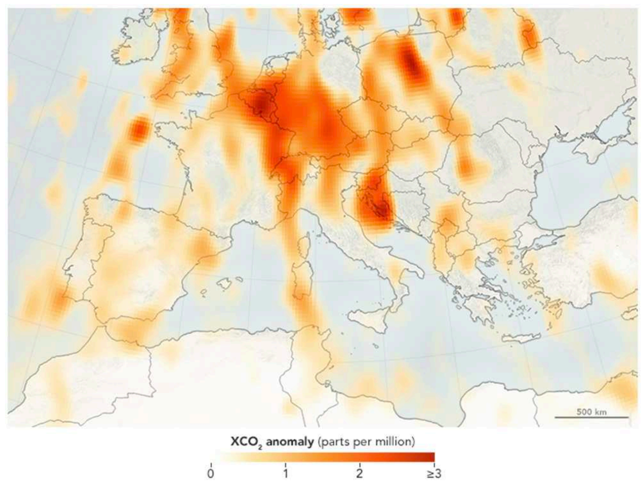


FIGURE 9 | Tracking human contribution to atmospheric CO_2 using data from NASA's OCO-2 satellite. Here, anomalies are shown from between 2014-2016. [NASA Earth Observatory maps by Joshua Stevens, using OCO-2 anomaly data courtesy of Hakkarainen et al. (2016) (NASA, 2016)].

Satellite observations of column-averaged CO_2 concentrations have demonstrated that, in some circumstances, satellites can provide top-down constraints on source emissions, but the capabilities of current satellite instruments are limited (Nassar et al., 2017, 2021; Hill and Nassar, 2019; Zheng et al., 2019; Wu et al., 2020). Data from GOSAT and OCO-2 do show statistically significant CO_2 enhancements over metropolitan regions (Kort et al., 2012; Schneising et al., 2013; Janardanan et al., 2016; Buchwitz et al., 2018; Wang et al., 2018; Reuter et al., 2019), and top-down methods have been applied to a few large thermal power plants (Bovensmann et al., 2010; Velazco et al., 2011), which are some of the largest point sources of anthropogenic CO_2 (Janssens-Maenhout et al., 2019). A recent analysis presented the first quantification of CO_2 emissions from individual power plants using OCO-2 observations (Nassar et al., 2017). However, because of the narrow swath (∼10 km at nadir) and 16-day repeat cycle of the sensor, the number of clear-day overpasses is too few for the development of a global CO_2 emissions inventory (Kiel et al., 2021; Nassar et al., 2021). The sparse sampling of the OCO-2 sensor will partly be overcome by the planned CO_2 imaging satellites that have denser spatial coverage, such as the CO_2 Monitoring mission (CO2M) and the GeoCarb instrument (Moore et al., 2018; Sierk et al., 2019).

An alternate method uses auxiliary satellite data, such as co-emitted NO_x, as a proxy for CO_2 emissions. Recent studies have shown that using NO_2 data for plume detection improves quantification of annual CO_2 emissions from point and urban sources as compared to CO_2 data alone (Kuhlmann et al., 2019, 2021; Reuter et al., 2019). This method takes advantage of the higher spatial resolution and spatiotemporal coverage of satellite NO_2, from which NO_x emissions are inferred, and have been shown to compare well to independent observations (Beirle et al., 2011; Duncan et al., 2013, 2016; de Foy et al., 2015; Lu et al., 2015; Krotkov et al., 2016; Liu et al., 2016, 2017; Goldberg et al., 2019). This approach is particularly useful for identifying new combustion sources as they come online and changes in existing point sources and urban areas (Duncan et al., 2016; Krotkov et al., 2016). The highest resolution satellite instrument for NO_2 is TROPOMI (2017-present) (Levelt et al., 2006, 2018; Veefkind et al., 2012; Munro et al., 2016; Krotkov et al., 2017). NO_2 and satellite-based CO and CO_2 data can also provide constraints on emissions inventories and be useful in monitoring trends and understanding regional-scale combustion (Silva and Arellano, 2017; Goldberg et al., 2019; Liu et al., 2020; Park et al., 2021).

CO_2 emissions from individual power plants and large cities may also be inferred from satellite NO_2. For power plants, these calculations have been performed using linear relationships between reported NO_x and CO_2 emissions by coal type, firing method, and emission control device (Liu et al., 2020). Ratios of NO_x to CO_2 emissions derived from U.S. power plants, where power plants have CEMS stack-height emissions monitors, offer a reasonable approximation for power plants in other countries, especially where coal type is known (Zoogman et al., 2017; Kim et al., 2019; Timmermans et al., 2019). City-scale emissions may be inferred through related statistical approaches to fit a collection of satellite-observed NO_2 plumes and inferred CO_2 emissions (Goldberg et al., 2019). While conducted and validated in the U.S., these approaches show potential for estimating CO_2 emissions in other countries as well.

Moving forward, a synergistic combination of bottom-up and top-down approaches would likely provide the greatest constraint on global anthropogenic CO_2 emissions. CO2M will carry instruments to observe both NO_2 and CO_2, which will allow for the estimation of NO_x/CO_2 ratios, although these ratios may have large regional and technological uncertainties (Kuhlmann et al., 2021). It has been shown that satellite NO_2 and CO_2 data could be used to infer a ratio to allow the estimation of CO_2 emissions using TROMPOMI and OCO-2 data for an individual power plant (Hakkarainen et al., 2021). These methods ideally would be complemented by a database with region-specific NO_x/CO_2 ratios from CEMS measurements or other bottom-up sources.

ENERGY RESILIENCE

Extreme weather events have long been a major risk factor for energy infrastructure, with climate change worsening these risks. Satellite data can provide a cost effective means for tracking vulnerable energy infrastructure, planning for new climate normals, and providing real-time support for operations and maintenance.

Energy Resilience and Global Change

Power outages, infrastructure damage, and challenges with adequately managing energy demand are well-known consequences of extreme weather and weather-related disruptions, including storms, heat waves, wildfires, and flooding (IEA, 2021a). In the U.S., for example, blackouts from extreme weather events cost an estimated \$20 to \$55 billion annually (Nik et al., 2021), and hurricanes are a major cause of power outages that have contributed to substantial loss of life and infrastructure (Alemazkoor et al., 2020). Extreme heat stresses the electric grid, resulting in increased demand for air conditioning and a loss in transmission and distribution efficiency (Añel et al., 2017). In February 2021, historic snowfall and ice across Texas led to blackouts that left millions of people without power (Nazir, 2021). Transmission line failure caused by extreme wind or heat can also result in wildfires, as in the 2009 Australian "Black Saturday" fires, where line failures ignited one of the most disastrous bushfires in Australian history, resulting in 173 deaths and \$4 billion (Australian) in property damage (Mitchell, 2013).

Within the energy management sector, there is a strong push to design climate-resilient infrastructure that can continue operating or recover quickly after immediate shocks and adapt to long-term changes in climate and environmental conditions (IEA, 2021b). In the U.S., increased emphasis on embedding climate adaptation and resilience into federal programs could support investments in the energy sector. Efforts currently underway include the Biden Administration's Executive Order 14008 on Tackling the Climate Crisis at Home and Abroad, Build Back Better Agenda, Infrastructure Investment and Jobs Act, and Justice40 Initiative, which focus on shifting energy supply to reduce environmental and health risks and support economic development for communities impacted by energy transitions (The White House, 2021a,b). Complementing these efforts, the Department of Energy (DOE) is deploying climate-resilient energy technologies nationwide, including in underserved communities (U.S. Department of Energy, 2021).

Satellite data can be used to better understand the impacts of a changing climate on energy infrastructure, advance the development of forecast models, and reduce the effort needed to assess environmental risks, which in turn can improve site-specific resilience planning (Leibrand et al., 2019). To support the climate adaptation and resilience efforts underway in the Biden

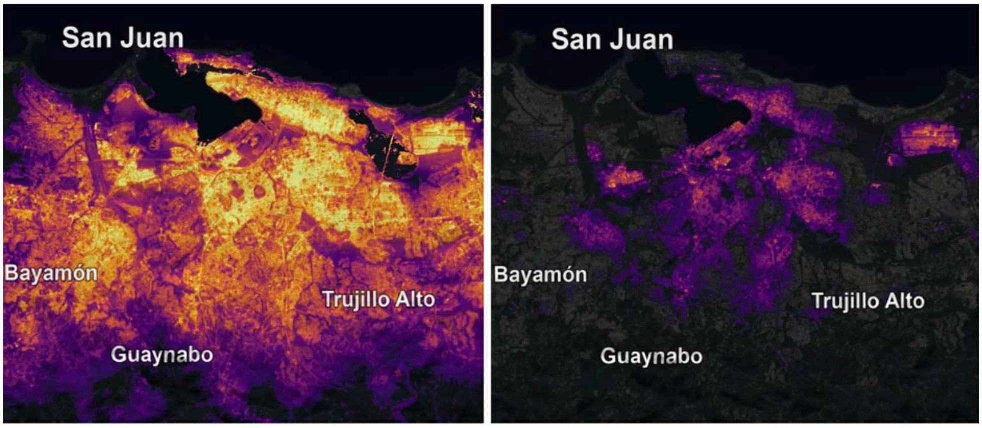


FIGURE 10 | Baseline (pre-storm) view of San Juan, Puerto Rico, nighttime lights **(L)** and average nighttime lights two months (Sep. 20 - Nov. 20, 2017) after Hurricane Maria **(R)**. (Credit: Kel Elkins, NASA Scientific Visualization Studio).

Administration, NASA and NOAA are planning to provide data and services to stakeholders to increase understanding of threats and vulnerabilities due to climate change (Margetta, 2021; US Department of Commerce, 2021). Satellite data can be used to inform planning to mitigate various energy infrastructure risks (Hauer and Miller, 2021). For example, several utility companies' wildfire mitigation plans use satellite data to monitor wildfire risks (Horizon West Transmission, 2021; Idaho Power, 2021; Pacific Gas and Electric Company, 2021; San Diego Gas and Electric Company, 2021; Southern California Edison Company, 2021), and satellite data has been used to identify vegetation encroachment and stressed or dead trees (Matikainen et al., 2016; Mahdi Elsiddig Haroun et al., 2020)[4,8].

Satellite data are also already being used to support disaster response in the energy sector. For example, in 2004, Eskom, South Africa's largest energy company developed a mobile fire alert system to mitigate line faults and provide near-real-time fire information. This system relies in part on NASA MODIS data and Meteosat Second Generation (MSG) data from the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) and is still in use today (Davies et al., 2008)[9]. MODIS data are also an integral part of NASA's Fire Information for Resource Management System (FIRMS), which integrates data from both MODIS and VIIRS instruments to deliver global active fire and hotspot data in near real time, within three hours of satellite observation (NASA, 2021).

Following Hurricane Maria's devastating impacts on Puerto Rico in 2017, NTL data from NASA's Black Marble product were used to understand the extent of power outages and characteristics of areas that withstood the greatest impact. These data were later used to monitor the effects of electricity restoration policies (see example in **Figure 10**), including how these policies can exacerbate inequality and unintentionally burden vulnerable populations (Romn et al., 2019). The high spatial resolution of Black Marble enables researchers to overcome four primary limitations of power outage data: timeliness, continuous data collection, consistent data across large geographic areas, and availability of data at a very fine spatial resolution (Romn et al., 2019).

[8]EPRI Identification of At-Risk Trees Using Satellite Imagery. https://www.epri.com/research/products/000000003002019050.

[9]Advanced Fire Information System (AFIS). AFIS. http://www.afis.co.za/.

DATA DISTRIBUTION

Distribution of satellite data for use by decision makers and researchers is a continuing challenge, especially as the number and complexity of data products grows. While researchers and high-end data users may choose to navigate data distribution platforms, many stakeholders prefer GIS-enabled web interfaces developed for their application area. The most developed energy-specific platform for satellite data distribution is the NASA Prediction Of Worldwide Energy Resources (POWER) project. POWER has provided Earth observation data for energy applications since 2002, with the goal of improving the accessibility and use of NASA data to support community research in three focus areas: (1) renewable energy development, (2) building energy efficiency and sustainability, and (3) agroclimatology applications (Zell et al., 2008; Eckman and Stackhouse, 2012). POWER allows users to select community-specific parameters, units, time periods, and output formats to efficiently retrieve data. The output data can then be directly applied in decision support tools, modeling and forecasting packages, and as inputs to scientific research.

The solar energy parameters in POWER are compiled using multiple satellite data sources. Hourly to long-term averaged parameters are provided for each parameter and can be used to support applications such as solar cooking, sizing solar panels, and designing battery backup systems. The daily and hourly time series include the basic solar and meteorological parameters as well as additional calculated parameters such as diffuse and direct normal radiation. For example, a community solar installation in a rural village in West Africa appeared to be working poorly. POWER data revealed that the solar array was in fact performing up to specifications, but cloudiness affected the available solar energy. Local management made the necessary adjustments to power usage and billing, drawing from data only available via satellite (NASA POWER and NASA ARSET, 2021). Similarly, a consultant responsible for solar installations at various locations in North Carolina uses the low latency data products to help assess changes in output (NASA POWER and NASA ARSET, 2021).

POWER data products have also been used to support a variety of energy system operation and maintenance decisions. For example, the Ottawa Renewable Energy Cooperative used the RETScreen Expert Clean Energy Technology software suite, which directly links the POWER web suite, to assess the potential benefits of paying for snow removal for a rooftop PV system by comparing lost generation to the building's actual load (NASA POWER and NASA ARSET, 2021). NASA's Office of Strategic Infrastructure also uses RETScreen for building energy management (Rosenzweig et al., 2014). Additionally, satellite estimates suggest that increases in the solar irradiance in the maize growing regions of the U.S. from the late 1980's through about 2012 were responsible for 27% of the productivity increase observed during this time period, relevant to corn-based biofuels (Tollenaar et al., 2017). Finally, the value of packaging correlated solar, wind, and other meteorological parameters has been demonstrated for a smart energy management system for hybrid solar-wind-biomass systems (Bhattacharjee and Nandi, 2021).

DISCUSSION

Applications of satellite data are growing across a wide range of energy policy and planning problems. Recent developments are increasing the potential for satellite data to support energy decision-making, with new public and private satellites being launched, advances in data processing techniques, and efforts by government and private organizations to increase uptake in new user communities. With more complete spatial and temporal coverage, satellite data can fill gaps in traditional data sources. Often the value of space-based data is greatest through integration with existing data and decision tools. Many types of satellite data have been collected for years, enabling analysts to track changes in energy supply, demand, and impacts over time and evaluate the effectiveness of policy interventions. While collecting satellite observations entails high fixed costs, the marginal costs are generally low, especially for datasets that are freely available to the public. Since data collection is remote, it also does not directly disturb local communities or the environment.

Our review points to many applications and opportunities for further use of space-based measurements. For energy supply, this includes resource potential and risk assessment to inform siting, development, and maintenance of energy infrastructure as well as real-time resource availability to support grid management and ensure reliability of supply. For energy demand, it includes energy use patterns to predict energy needs and identify locations with unserved demand. For energy impacts, it includes the effects of energy use on climate, air quality, and water and land systems, as well as monitoring efforts to reduce these impacts. Satellite data are also playing an increasing role in supporting investments in energy resilience, both in advance and in the aftermath of disruptions to energy access. The expansive coverage of many measurements allows for global indexing of critical metrics, and the increase in temporal resolution of new products means that satellite data can be used to track progress toward policy commitments to reduce energy-related emissions, increase energy access, and support sustainable energy transitions around the world.

The technical limitations in the use of satellite data for energy applications are primarily driven by insufficient spatial and temporal resolution. For example, polar orbiting satellites such as the Landsat 8 Operational Land Imager (OLI) and the recently launched Landsat 9 Operational Land Imager 2 (OLI-2) provide radiance measurements that are high spatial resolution (30 m) and multispectral (visible, near-infrared, and shortwave infrared bands). These measurements are suitable for land use characterization at urban and individual agricultural field scales. However, the 16-day repeat cycle, relatively narrow swath width (165 km), and likelihood of cloudy scenes limits temporal sampling to typically a single observation at one location each month, which does not allow for rapid responses to changing conditions. The European Sentinel-2 MultiSpectral Instrument (MSI) partially addresses these limitations with a 5-day repeat cycle and higher spatial resolution (10 m for the visible channels and 20 m for near-infrared and shortwave infrared). Other polar orbiting instruments, such as VIIRS, have more channels and larger swath widths (3000 km) and can observe the entire planet each day. However, VIIRS has significantly lower spatial resolution than either Landsat or Sentinel-2.

Unlike polar orbiting satellites, geostationary satellites continuously observe the same area and consequently have very high temporal sampling, and a constellation of geostationary satellites can allow for near global, continuous sampling of the Earth throughout the day. The U.S. Advanced Baseline Imager (ABI) on GOES East and West, Japanese Advanced Himawari Imager (AHI), and European Meteosat Third Generation (MTG) are all third-generation geostationary instruments with similar retrieval capabilities as VIIRS but even coarser spatial resolution. A new generation of instruments such as the recently launched Geostationary Environmental Monitoring Spectrometer (GEMS), Tropospheric Emissions: Monitoring of Pollution (TEMPO), and the European UltraViolet/Visible/Near-Infrared (UVN) instrument will provide the first geostationary hourly ultraviolet radiance measurements suitable for a wide variety of energy applications, from tracking photosynthetic activity for biofuel production to monitoring NO_2 emissions from fossil fuel combustion. The GeoCarb instrument will provide similar measures of photosynthetic activity as well as geostationary CO_2, CO, and CH_4 retrievals over North America.

Combining high spatial resolution polar orbiting measurements with high temporal resolution measurements from geostationary satellites will create unprecedented opportunities for energy policy and planning. For example, accurate short-term cloud forecasts are critical for optimizing electric power generation and load balancing. Improved use of geostationary cloud measurements for solar PV forecasting could include data assimilation for short-term surface irradiance forecasts and advanced pattern recognition estimates of cloud motion. Using satellite data for improved monitoring and prediction of droughts (in particular, flash droughts) and resulting changes in photosynthetic activity could have significant impacts on improving biofuel production efficiency. Satellite data with high spatial and temporal resolution also can uncover patterns of energy injustice. For example, new satellite-based research can help us understand who has access to energy infrastructure and the reliability, quality, and impacts of energy services for different groups. With the ability to monitor changes over time, we can also assess equity in energy transitions.

Beyond spatial and temporal resolution, analysts must also carefully consider the appropriateness of satellite data for energy policy and planning applications. It is particularly important to understand the distribution of potential errors in satellite measurements when they are used as a proxy for energy-relevant variables, and how these errors affect causal inference (Jain, 2020). For example, NTL datasets measure nighttime luminosity but are often used to estimate energy use and access, economic activity, and other variables. This approach can lead to biased inferences on the effects of policies if NTL data undercounts energy access in areas with intermittent service or underestimates economic activity in high luminosity areas due to saturation effects. Other types of errors can also be systematic, such as when agroforestry or plantations are classified as tree cover or when small-scale logging is undetected, which may lead to an underestimate of the effects of policies on forest loss. These challenges underscore the importance of data validation and the value of combining other types of data with satellite measurements to create a more complete picture of the energy system.

While researchers are actively working to address the technical limitations of satellite data for energy applications, addressing social and structural barriers will be equally important. While social science studies specifically on the use of satellite data for decision-making are more limited, in the case of air quality – especially for policy organizations implementing the Clean Air Act in the U.S. – a range of social and structural as well as technical barriers impede data use relative to traditional monitoring and assessment methods (Milford and Knight, 2017). Satellite data do not fit with decision and policy frameworks in a clear manner, and users have expressed uncertainty about whether data will be accepted for regulatory purposes. Research also indicates a number of social barriers to satellite data use, including difficulty finding data, data formats that are unfamiliar or difficult to use, and lack of staff time, training, and expertise to acquire and process data. Two-way dialogue between end-users and satellite experts has helped identify specific areas where space-based data can contribute effectively to information needs.

Collaboration between researchers with expertise in satellite data analysis, energy systems and policy, and a broader set of social science disciplines will be essential to realizing the potential for satellite data to support energy decision-making. Research has pointed to the importance of active communication between experts and decision-makers and investing in translational work to bridge the gap between scientific data and decision processes (Cash et al., 2003; Klemun et al., 2020). Researchers themselves can engage in this boundary work – for example, NASA encourages engagement with the full satellite data application process through its Application Readiness Levels (ARLs), which range from initial discovery to full integration of satellite data into a partner's decision-making systems and processes. However, research also points to the vital role of boundary organizations and boundary objects (including data portals, interactive maps, and training workshops) in facilitating this work. Several organizations are actively working to enhance the usability of satellite data for energy applications.

Satellite applications for energy planning and policy are growing rapidly, with novel information needs to support sustainable energy transitions, a suite of new satellites recently launched or planned to be launched soon, and advances in methods for analyzing satellite data products and translating the results into useful information for energy decision-making. Our review suggests that, while there are many energy application areas where satellite data are already playing an important role, there is significant untapped potential to apply satellite data to support decision-making around energy supply, demand, impacts, and resilience. As advances in satellite data analysis open up new opportunities to support decision-making, active dialogue between experts in satellite data and energy planning and policy, as well as decision-makers across energy sectors, will be essential to maximize the usefulness of satellite data for sustainable energy transitions.

AUTHOR CONTRIBUTIONS

MRE led the conceptualization, organization, and editing of the review paper with TH and RBP. All authors drafted sections of the paper, edited the final manuscript, and approved the submitted version.

FUNDING

TH acknowledges support from NASA Health and Air Quality Applied Sciences Team Grant Number 80NSSC21K0427. HG and TL acknowledge support from the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409. KS acknowledges support from NASA Grant Number 80NSSC19K0638. Battelle expresses its gratitude for the support from Richard Eckman, Paul Stackhouse, and Brock Blevins of NASA, Craig Zamuda of U.S. DOE and DOE Partnership for Energy Sector Resilience members, Technical Advisory Group members, Chris Johnson of Battelle, Bradley Macpherson of Booz Allen Hamilton, the GEO community, and NASA Grant Number 80NSSC18K0325. This research was partially financially supported by the University of Wisconsin-Madison Office of Sustainability.

REFERENCES

42 U. S., and Code § 7401. (1977). https://www.govinfo.gov/content/pkg/STATUTE-91/pdf/STATUTE-91-Pg685.pdf.

Ackerman, K. V., and Sundquist, E. T. (2008). Comparison of two U.S. power-plant carbon dioxide emissions data sets. *Environ. Sci. Technol.* 42, 5688–5693. doi: 10.1021/es800221q

Ahsbahs, T., Badger, M., Volker, P., Hansen, K. S., and Hasager, C. B. (2018). Applications of satellite winds for the offshore wind farm site Anholt. *Wind Energy Sci.* 3, 573–588. doi: 10.5194/wes-3-573-2018

Aklin, M., Harish, S. P., and Urpelainen, J. (2018). A global analysis of progress in household electrification. *Energy Policy* 122, 421–428. doi: 10.1016/j.enpol.2018.07.018

Alemazkoor, N., Rachunok, B., Chavas, D. R., Staid, A., Louhghalam, A., Nateghi, R., et al. (2020). Hurricane-induced power outage risk under climate change is primarily driven by the uncertainty in projections of future hurricane frequency. *Sci. Rep.* 10, 15270. doi: 10.1038/s41598-020-72207-z

Alkan, M. (2009). Detection of Coal Ash Turbidity in Marine Environment Using Remote Sensing. *Fresenius Environ. Bull.* 18, 2072–2078.

Añel, J. A., Fernández-González, M., Labandeira, X., López-Otero, X., and De la Torre, L. (2017). Impact of cold waves and heat waves on the energy production sector. *Atmosphere* 8, 209. doi: 10.3390/atmos8110209

Anenberg, S. C., Mohegh, A., Goldberg, D. L., Kerr, G. H., Brauer, M., Burkart, K., et al. (2022). Long-term trends in urban NO_2 concentrations and associated paediatric asthma incidence: estimates from global datasets. *Lancet Planet. Health* 6, e49–e58. doi: 10.1016/S2542-5196(21)00255-2

Arderne, C., Zorn, C., Nicolas, C., and Koks, E. E. (2020). Predictive mapping of the global power system using open data. *Sci. Data* 7, 19. doi: 10.1038/s41597-019-0347-4

Avtar, R., Tripathi, S., and Aggarwal, A. K. (2019). Assessment of energy–population–urbanization nexus with changing energy industry scenario in India. *Land* 8, 124. doi: 10.3390/land8080124

Badger, M., Peña, A., Hahmann, A. N., Mouche, A. A., and Hasager, C. B. (2016). Extrapolating satellite winds to turbine operating heights. *J. Appl. Meteorol. Climatol.* 55, 975–991. doi: 10.1175/JAMC-D-15-0197.1

Bastiaanssen, W. G. M., Simons, G., and Sheridan, P. F. (2012). *The Potential of Satellite-based Information for Monitoring Water and Production Related Aspects of the Sustainability of Biofuels*.

Baxter, R. E., and Calvert, K. E. (2017). Estimating available abandoned cropland in the United States: possibilities for energy crop production. *Ann. Am. Assoc. Geogr.* 107, 1162–1178. doi: 10.1080/24694452.2017.1298985

Bayramov, E. (2013). "Optimization of environmental monitoring principles along oil and gas pipelines using GIS and remote sensing," in *2013 7th International Conference on Application of Information and Communication Technologies* p. 1–3. doi: 10.1109/ICAICT.2013.6722681

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T. (2011). Megacity emissions and lifetimes of nitrogen oxides probed from space. *Science* 333, 1737–1739. doi: 10.1126/science.1207824

Berndes, G., Ahlgren, S., Börjesson, P., and Cowie, A. L. (2013). Bioenergy and land use change—state of the art. *WIREs Energy Environ.* 2, 282–303. doi: 10.1002/wene.41

Bhattacharjee, S., and Nandi, C. (2021). Design of a voting based smart energy management system of the renewable energy based hybrid energy system for a small community. *Energy* 214, 118977. doi: 10.1016/j.energy.2020.118977

Bhattarai, N., Wagle, P., Gowda, P. H., and Kakani, V. G. (2017). Utility of remote sensing-based surface energy balance models to track water stress in rain-fed switchgrass under dry and wet conditions. *ISPRS J. Photogramm. Remote Sens.* 133, 128–141. doi: 10.1016/j.isprsjprs.2017.10.010

Blackman, A. (2013). Evaluating forest conservation policies in developing countries using remote sensing data: an introduction and practical guide. *For. Policy Econ.* 34, 1–16. doi: 10.1016/j.forpol.2013.04.006

Bonsucro (2021). Bonsucro Draft Production Standard Version 5, 7. Available online at: https://www.bonsucro.com/wp-content/uploads/2021/06/Bonsucro-Production-Standard-draft-v.5.7-1.pdf (accessed July 11, 2022).

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., et al. (2010). Remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications. *Atmospheric Meas. Tech.* 3, 781–811. doi: 10.5194/amt-3-781-2010

Brandt, A. R., Heath, G. A., and Cooley, D. (2016). Methane leaks from natural gas systems follow extreme distributions. *Environ. Sci. Technol.* 50, 12512–12520. doi: 10.1021/acs.est.6b04303

Brown, J. C., Hanley, E., Bergtold, J., Caldas, M., Barve, V., Peterson, D., et al. (2014). Ethanol plant location and intensification vs. extensification of corn cropping in Kansas. *Appl. Geogr.* 53, 141–148. doi: 10.1016/j.apgeog.2014.05.021

Buchwitz, M., Reuter, M., Schneising, O., Nol, S., Gier, B., Bovensmann, H., et al. (2018). Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016. *Atmospheric Chem. Phys.* 18, 17355–17370. doi: 10.5194/acp-18-17355-2018

Burlig, F., and Preonas, L. (2016). *Out of the Darkness and Into the Light?* Development Effects of Rural Electrification. EPIC https://epic.uchicago.edu/research/out-of-the-darkness-and-into-the-light-development-effects-of-rural-electrification/ (accessed July 11, 2022).

Cai, X., Zhang, X., and Wang, D. (2011). Land availability for biofuel production. *Environ. Sci. Technol.* 45, 334–339. doi: 10.1021/es103338e

Carley, S., and Konisky, D. M. (2020). The justice and equity implications of the clean energy transition. *Nat. Energy* 5, 569–577. doi: 10.1038/s41560-020-0641-6

Carlson, K. M., Curran, L. M., Ratnasari, D., Pittman, A. M., Soares-Filho, B. S., Asner, G. P., et al. (2012). Committed carbon emissions, deforestation, and community land conversion from oil palm plantation expansion in West Kalimantan, Indonesia. *Proc. Natl. Acad. Sci.* 109, 7559–7564. doi: 10.1073/pnas.1200452109

Cash, D., Clark, W., Alcock, F., Dickson, N., Selin, N., Guston, D., et al. (2003). Knowledge systems for sustainable development. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8086–8091. doi: 10.1073/pnas.1231332100

Chen, Y., Ale, S., Rajan, N., and Munster, C. (2017). Assessing the hydrologic and water quality impacts of biofuel-induced changes in land use and management. *GCB Bioenergy* 9, 1461–1475. doi: 10.1111/gcbb.12434

Cheng, L., Zang, H., Wei, Z., Ding, T., Xu, R., Sun, G., et al. (2022). Short-term Solar Power Prediction Learning Directly from Satellite Images With Regions of Interest. *IEEE Trans. Sustain. Energy* 13, 629–639. doi: 10.1109/TSTE.2021.3123476

Creutzig, F., Agoston, P., Minx, J. C., Canadell, J. G., Andrew, R. M., Qur, C. L., et al. (2016). Urban infrastructure choices structure climate solutions. *Nat. Clim. Change* 6, 1054–1056. doi: 10.1038/nclimate3169

Creutzig, F., Erb, K. H., Haberl, H., Hof, C., Hunsberger, C., et al. (2021). Considering sustainability thresholds for BECCS in IPCC and biodiversity assessments. *GCB Bioenergy* 13, 510–515. doi: 10.1111/gcbb.12798

Crisp, D. (2015). "Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2)," in *Earth Observing Systems XX* vol. 9607 p. 960702 (SPIE). doi: 10.1117/12.2187291

Cui, R. Y., Hultman, N., Edwards, M. R., He, L., Sen, A., Surana, K., et al. (2019). Quantifying operational lifetimes for coal power plants under the Paris goals. *Nat. Commun.* 10, 4759. doi: 10.1038/s41467-019-12618-3

Davies, D., Vosloo, H., Vannan, S., and Frost, P. (2008). "Near real-time fire alert system in South Africa: From desktop to mobile service," in *Proceedings of the 7th ACM Conference on Designing Interactive Systems* (New York, NY), 315–322. doi: 10.1145/1394445.1394479

Davis, G. (2007). History of the NOAA satellite program. *J. Appl. Remote Sens.* 1, 012504. doi: 10.1117/1.2642347

de Foy, B., Lu, Z., Streets, D. G., Lamsal, L. N., and Duncan, B. N. (2015). Estimates of power plant NO_x emissions and lifetimes from OMI NO_2 satellite retrievals. *Atmos. Environ.* 116, 1–11. doi: 10.1016/j.atmosenv.2015.05.056

de Miguel, A. S., Castao, J. G., Zamorano, J., Pascual, S., Ngeles, M., Cayuela, L., et al. (2014). Atlas of astronaut photos of Earth at night. *Astron. Geophys.* 55, 436 doi: 10.1093/astrogeo/atu165

Dennis, B. (2021). *How Satellites Could Help Hold Countries to Emissions Promises Made at COP26 Summit*. Washington Post. Available online at: https://www.washingtonpost.com/climate-environment/2021/11/09/cop26-satellites-emissions

Doll, C. N. H., and Pachauri, S. (2010). Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery. *Energy Policy* 38, 5661–5670. doi: 10.1016/j.enpol.2010.05.014

Donaldson, D., and Storeygard, A. (2016). The View from Above: Applications of Satellite Data in Economics. *J. Econ. Perspect.* 30, 171–198. doi: 10.1257/jep.30.4.171

Donner, S. D., and Kucharik, C. J. (2008). Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. *Proc. Natl. Acad. Sci.* 105, 4513–4518. doi: 10.1073/pnas.0708300105

Dugoua, E., Kennedy, R., and Urpelainen, J. (2017). *Satellite Data for the Social Sciences: Measuring Rural Electrification with Nighttime Lights*. Available online at: https://papers.ssrn.com/abstract=3083505 (accessed July 11, 2022).

Duncan, B. N., Lamsal, L. N., Thompson, A. M., Yoshida, Y., Lu, Z., Streets, D. G., et al. (2016). A space-based, high-resolution view of notable changes in urban NO_x pollution around the world (2005–2014). *J. Geophys. Res. Atmospheres* 121, 976–996. doi: 10.1002/2015JD024121

Duncan, B. N., Yoshida, Y., de Foy, B., Lamsal, L. N., Streets, D. G., Lu, Z., et al. (2013). The observed response of Ozone Monitoring Instrument (OMI) NO_2 columns to NO_x emission controls on power plants in the United States: 2005–2011. *Atmos. Environ.* 81, 102–111. doi: 10.1016/j.atmosenv.2013.08.068

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., et al. (2010). Application of OMI observations to a space-based indicator of NO_x and VOC controls on surface ozone formation. *Atmos. Environ.* 44, 2213–2223. doi: 10.1016/j.atmosenv.2010.03.010

Duren, R. M., Thorpe, A. K., Foster, K. T., Rafiq, T., Hopkins, F. M., Yadav, V., et al. (2019). California's methane super-emitters. *Nature* 575, 180–184. doi: 10.1038/s41586-019-1720-3

Dybas, C. L. (2005). Ocean dead zones spreading in world oceans. *BioScience* 55. doi: 10.1641/0006-3568(2005)055[0552:DZSIWO]2.0.CO;2

Eckman, R. S., and Stackhouse, P. W. (2012). CEOS contributions to informing energy management and policy decision making using space-based Earth observations. *Appl. Energy* 90, 206–210. doi: 10.1016/j.apenergy.2011.03.001

Edwards, M. R., Giang, A., Macey, G. P., Magavi, Z., Nicholas, D., Ackley, R., et al. (2021). Repair failures call for new policies to tackle leaky natural gas distribution systems. *Environ. Sci. Technol.* 55, 6561–6570. doi: 10.1021/acs.est.0c07531

Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F. C., and Taneja, J. (2021). Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sens.* 13, 922. doi: 10.3390/rs13050922

Esparza, Á., and Mattson, R. (2021). *Monitoring Leaks at Oil and Gas Facilities Using Satellites in Support of ESG Initiatives*. ROGTEC. Available online at: https://www.rogtecmagazine.com/monitoring-leaks-at-oil-and-gas-facilities-using-satellites-in-support-of-esg-initiatives/ (accessed July 11, 2022).

Esparza, Á. E., and Gauthier, J. F. (2021). *Monitoring Leaks at Oil and Gas Facilities Using the Same Sensor on Aircraft and Satellite Platforms*. OnePetro. doi: 10.15530/urtec-2021-5666

European Commission (2021). Launch by US, EU and Partners of the Global Methane Pledge. *European Commission* Available online at: https://ec.europa.eu/commission/presscorner/detail/en/statement_21_5766 (accessed July 11, 2022).

European Space Agency (2005). *Envisat enables first global check of regional methane emissions*. Available online at: https://www.esa.int/Applications/Observing_the_Earth/Envisat/Envisat_enables_first_global_check_of_regional_methane_emissions (accessed July 11, 2022).

European Space Agency (2020). *Detecting methane emissions during COVID-19*. Available online at: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Detecting_methane_emissions_during_COVID-19 (accessed July 11, 2022).

Falchetta, G., Kasamba, C., and Parkinson, S. C. (2020b). Monitoring hydropower reliability in Malawi with satellite data and machine learning. *Environ. Res. Lett.* 15, 014011. doi: 10.1088/1748-9326/ab6562

Falchetta, G., and Noussan, M. (2019). Interannual variation in night-time light radiance predicts changes in national electricity consumption conditional on income-level and region. *Energies* 12, 456. doi: 10.3390/en12030456

Falchetta, G., Pachauri, S., Byers, E., Danylo, O., and Parkinson, S. C. (2020a). Satellite Observations Reveal Inequalities in the Progress and Effectiveness of Recent Electrification in Sub-Saharan Africa. *One Earth* 2, 364–379. doi: 10.1016/j.oneear.2020.03.007

Falchetta, G., Pachauri, S., Parkinson, S., and Byers, E. (2019). A high-resolution gridded dataset to assess electrification in sub-Saharan Africa. *Sci. Data* 6, 110. doi: 10.1038/s41597-019-0122-6

Fletcher, S., Lickley, M., and Strzepek, K. (2019). Learning about climate change uncertainty enables flexible water infrastructure planning. *Nat. Commun.* 10, 1782. doi: 10.1038/s41467-019-09677-x

Fobi, S., Deshpande, V., Ondiek, S., Modi, V., and Taneja, J. A. (2018). Longitudinal study of electricity consumption growth in Kenya. *Energy Policy* 123, 569–578. doi: 10.1016/j.enpol.2018.08.065

Fragkias, M., Lobo, J., and Seto, K. C. (2017). A comparison of nighttime lights data for urban energy research: Insights from scaling analysis in the US system of cities. *Environ. Plan. B Urban Anal. City Sci.* 44, 1077–1096. doi: 10.1177/0265813516658477

Frolking, S., Milliman, T., Mahtta, R., Paget, A., Long, D. G., Seto, K. C. et al. (2022). A global urban microwave backscatter time series data set for 1993–2020 using ERS, QuikSCAT, and ASCAT data. *Sci. Data* 9, 88. doi: 10.1038/s41597-022-01193-w

Frolking, S., Milliman, T., Seto, K. C., and Friedl, M. A. (2013). A global fingerprint of macro-scale changes in urban structure from 1999 to 2009. *Environ. Res. Lett.* 8, 024004. doi: 10.1088/1748-9326/8/2/024004

Ganesan, A. L., Schwietzke, S., Poulter, B., Arnold, T., Lan, X., Rigby, M., et al. (2019). Advancing scientific understanding of the global methane budget in support of the Paris Agreement. *Glob. Biogeochem. Cycles* 33, 1475–1512. doi: 10.1029/2018GB006065

Gelfand, I., Sahajpal, R., Zhang, X., Izaurralde, R. C., Gross, K. L., Robertson, G. P., et al. (2013). Sustainable bioenergy production from marginal lands in the US Midwest. *Nature* 493, 514–517. doi: 10.1038/nature11811

Goga, T., Feranec, J., Bucha, T., Rusnk, M., Sakov, I., Barka, I., et al. (2019). A review of the application of remote sensing data for abandoned agricultural land identification with focus on Central and Eastern Europe. *Remote Sens.* 11, 2759. doi: 10.3390/rs11232759

Goldberg, D. L., Anenberg, S. C., Kerr, G. H., Mohegh, A., Lu, Z., and Streets, D. G. (2021). TROPOMI NO_2 in the United States: A Detailed Look at the Annual Averages, Weekly Cycles, Effects of Temperature, and Correlation With Surface NO_2 Concentrations. *Earths Future* 9, e2020EF001665. doi: 10.1029/2020EF001665

Goldberg, D. L., Lu, Z., Oda, T., Lamsal, L. N., Liu, F., Griffin, D., et al. (2019). Exploiting OMI NO_2 satellite observations to infer fossil-fuel CO_2 emissions from U.S. *megacities. Sci. Total Environ.* 695, 133805. doi: 10.1016/j.scitotenv.2019.133805

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., et al. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. doi: 10.1016/j.rse.2017.06.031

Gu, Y., and Wylie, B. (2016). Mapping marginal croplands suitable for cellulosic feedstock crops in the Great Plains, United States. *GCB Bioenergy* 9, 836–844. doi: 10.1111./gcbb.12388

Gurney, K. R., Liang, J., O'Keeffe, D., Patarasuk, R., Hutchins, M., Huang, J., et al. (2019). Comparison of Global Downscaled Versus Bottom-Up Fossil Fuel CO_2 Emissions at the Urban Scale in Four U.S. Urban Areas. *J. Geophys. Res. Atmospheres* 124, 2823–2840. doi: 10.1029/2018JD028859

Haag, A. S., Walker, N. D., Pilley, C. T., Comeaux, J. C., Calvasina, J. J., Firth, R., et al. (2009). "Satellite surveillance of the Gulf and beyond: An overview of the LSU Earth Scan Laboratory's ocean observing capabilities," in *OCEANS 2009.* p. 1–8. doi: 10.23919/OCEANS.2009.5422173

Haberl, H., Erb, K-. H., Krausmann, F., Running, S., Searchinger, T. D., et al. (2013). Bioenergy: how much can we expect for 2050? *Environ. Res. Lett.* 8, 031004. doi: 10.1088/1748-9326/8/3/031004

Hakkarainen, J., Ialongo, I., and Tamminen, J. (2016). Direct space-based observations of anthropogenic CO_2 emission areas from OCO-2. *Geophys. Res. Lett.* 43, 11400–6 doi: 10.1002/2016GL070885

Hakkarainen, J., Szelg, M. E., Ialongo, I., Retscher, C., Oda, T., Crisp, D., et al. (2021). Analyzing nitrogen oxides to carbon dioxide emission ratios from space: A case study of Matimba Power Station in South Africa. *Atmospheric Environ. X* 10, 100110. doi: 10.1016/j.aeaoa.2021.100110

Hamada, Y., Ssegane, H., and Negri, M. C. (2015). Mapping intra-field yield variation using high resolution satellite imagery to integrate bioenergy and environmental stewardship in an agricultural watershed. *Remote Sens.* 7, 9753–9768. doi: 10.3390/rs70809753

Hauer, R., and Miller, R. (2021). *Utility Vegetation Management in North America: Results from a 2019 Utility Forestry Census of Tree Activities and Operations*.

He, S., Gu, L., Tian, J., Deng, L., Yin, J., Liao, Z., et al. (2021). Machine learning improvement of streamflow simulation by utilizing remote sensing data and potential application in guiding reservoir operation. *Sustainability* 13, 3645. doi: 10.3390/su13073645

Hendricks, N. P., Sinnathamby, S., Douglas-Mankin, K., Smith, A., Sumner, D. A., Earnhart, D. H., et al. (2014). The environmental effects of crop price increases: Nitrogen losses in the U.S. Corn Belt. *J. Environ. Econ. Manag.* 68, 507–26. doi: 10.1016/j.jeem.2014.09.002

Hernández, D. (2015). Sacrifice Along the Energy Continuum: A Call for Energy Justice. *Environ. Justice* 8, 151–156. doi: 10.1089/env.2015.0015

Hewson, R., Mshiu, E., Hecker, C., van der Werff, H., van Ruitenbeek, F., Alkema, D., et al. (2020). The application of day and night time ASTER satellite imagery for geothermal and mineral mapping in East Africa. *Int. J. Appl. Earth Obs. Geoinformation* 85, 101991. doi: 10.1016/j.jag.2019.101991

Hill, T., and Nassar, R. (2019). Pixel size and revisit rate requirements for monitoring power plant CO_2 emissions from space. *Remote Sens.* 11, 1608. doi: 10.3390/rs11131608

Holloway, T., Jacob, D. J., and Miller, D. (2018). Short history of NASA applied science teams for air quality and health. *J. Appl. Remote Sens.* 12, 1. doi: 10.1117/1.JRS.12.042611

Holloway, T., Miller, D., Anenberg, S., Diao, M., Duncan, B., Fiore, A. M., et al. (2021). Satellite monitoring for air quality and health. *Annu. Rev. Biomed. Data Sci.* 4, 417–447. doi: 10.1146/annurev-biodatasci-110920-093120

Horizon West Transmission, L. (2021). Wildfire Mitigation Plan 2021. Available online at: https://energysafety.ca.gov/wp-content/uploads/docs/wmp-2021/hwt_2021_wmp.pdf

Howari, F. (2015). Prospecting for geothermal energy through satellite based thermal data: Review and the way forward. *J Environ Sci Manage*. 1, 265–74. doi: 10.7508/GJESM.2015.04.001

Hultman, N. E., Clarke, L., Frisch, C., Kennedy, K., McJeon, H., Cyrs, T., et al. (2020). Fusing subnational with national climate action is central to decarbonization: the case of the United States. *Nat. Commun.* 11, 5255. doi: 10.1038/s41467-020-18903-w

Idaho Power (2021). Wildfire Mitigation Plan 2022. Available online at: https://docs.idahopower.com/pdfs/Safety/2022Wildfire%20MitigationPlan.pdf (accessed July 11, 2022).

IEA (2019). *Africa Energy Outlook – Analysis*. IEA. Available online at: https://www.iea.org/reports/africa-energy-outlook-2019 (accessed July 11, 2022).

IEA (2021a). *Global Energy Review – Analysis*. IEA. Available online at: https://www.iea.org/reports/global-energy-review-2021 (accessed July 11, 2022).

IEA (2021b). *Climate Resilience Policy Indicator*. Available online at: https://www.iea.org/reports/climate-resilience-policy-indicator/climate-resilience-policy-indicator (accessed July 11, 2022).

IEA (2021c). *Access to Electricity*. Available online at: https://www.iea.org/reports/sdg7-data-and-projections/access-to-electricity (accessed July 11, 2022).

Inman, R. H., Pedro, H. T. C., and Coimbra, C. F. M. (2013). Solar forecasting methods for renewable energy integration. *Prog. Energy Combust. Sci.* 39, 535–576. doi: 10.1016/j.pecs.2013.06.002

International Hydropower Association (2020). Hydropower Status Report. Available online at: https://www.hydropower.org/publications/2020-hydropower-status-report (accessed July 11, 2022).

International Hydropower Association (2021). Hydropower Status Report. Available online at: https://www.hydropower.org/publications/2021-hydropower-status-report (accessed July 11, 2022).

IPCC (2018). Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways. *Special Report Global Warming of 1.5°C* Available online at: https://www.ipcc.ch/sr15/

IPCC (2022). "Climate Change 2022: Mitigation of Climate Change," in *Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds P. R. Shukla, J. Skea, R. Slade, A. vAl Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley (Cambridge, UK; New York, NY: Cambridge University Press), doi: 10.1017/9781009157926

Ishii, T., Simo-Serra, E., Iizuka, S., Mochizuki, Y., Sugimoto, A., Ishikawa, H., et al. (2016). "Detection by classification of buildings in multispectral satellite imagery," in *2016 23rd International Conference on Pattern Recognition (ICPR)* p. 3344–3349 (IEEE). doi: 10.1109/ICPR.2016.7900150

Jacob, D. J., Holloway, T., and Haynes, J. D. (2014). The NASA air quality applied sciences team (AQAST). *Environ. Manag.* 2. Available online at: https://pubs.awma.org/flip/EM-Feb-2014/coverstory.pdf

Jain, M. (2020). The Benefits and Pitfalls of Using Satellite Data for Causal Inference. *Rev. Environ. Econ. Policy* 14, 157–169. doi: 10.1093/reep/rez023

Janardanan, R., Maksyutov, S., Oda, T., Saito, M., Kaiser, J. W., Ganshin, A., et al. (2016). Comparing GOSAT observations of localized CO_2 enhancements by large emitters with inventory-based estimates. *Geophys. Res. Lett.* 43, 3486–3493. doi: 10.1002/2016GL067843

Janjai, S., Pankaew, P., Laksanaboonsong, J., and Kitichantaropas, P. (2011). Estimation of solar radiation over Cambodia from long-term satellite data. *Renew. Energy* 36, 1214–1220. doi: 10.1016/j.renene.2010.09.023

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., et al. (2019). EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth Syst. Sci. Data* 11, 959–1002. doi: 10.5194/essd-11-959-2019

Jiang, C., Guan, K., Khanna, M., Chen, L., and Peng, J. (2021). Assessing marginal land availability based on land use change information in the contiguous United States. *Environ. Sci. Technol.* 55, 10794–10804. doi: 10.1021/acs.est.1c02236

Kaminski, T., Scholze, M., Rayner, P., Vobeck, M., Buchwitz, M., Reuter, M., et al. (2022). Assimilation of atmospheric CO_2 observations from space can support national CO_2 emission inventories. *Environ. Res. Lett.* 17, 014015. doi: 10.1088/1748-9326/ac3cea

Kang, S., Post, W., Wang, D., Nichols, J., Bandaru, V., West, T., et al. (2013). Hierarchical marginal land assessment for land use planning. *Land Use Policy* 30, 106–113. doi: 10.1016/j.landusepol.2012.03.002

Karlsson, K-. G., Anttila, K., Trentmann, J., Stengel, M., Fokke Meirink, J., et al. (2017). CLARA-A2: the second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data. *Atmospheric Chem. Phys.* 17, 5809–5828. doi: 10.5194/acp-17-5809-2017

Kerr, G. H., Goldberg, D. L., and Anenberg, S. C. (2021). COVID-19 pandemic reveals persistent disparities in nitrogen dioxide pollution. *Proc. Natl. Acad. Sci.* 118. doi: 10.1073/pnas.2022409118

Khaki, M., Hendricks Franssen, H.-., J., and Han, S. C. (2020). Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation. *Sci. Rep.* 10, 18791. doi: 10.1038/s41598-020-75710-5

Khanna, M., Chen, L., Basso, B., Cai, X., Field, J. L., Guan, K., et al. (2021). Redefining marginal land for bioenergy crop production. *GCB Bioenergy* 13, 1590–1609. doi: 10.1111/gcbb.12877

Kiel, M., Eldering, A., Roten, D. D., Lin, J. C., Feng, S., Lei, R., et al. (2021). Urban-focused satellite CO_2 observations from the Orbiting Carbon Observatory-3: A first look at the Los Angeles megacity. *Remote Sens. Environ.* 258, 112314. doi: 10.1016/j.rse.2021.112314

Kim, J., Jeong, U., Ahn, M. H., Kim, J., Park, R., et al. (2019). New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bull. Am. Meteorol. Soc.* 101, 00.

Kimball, J., and McDonald, K. (2020). An Earth System Data Record for Land Surface Freeze-Thaw State: Quantifying Terrestrial Water Mobility Constraints to Global Ecosystem Processes. *NASA*. Available online at: https://www.earthdata.nasa.gov/esds/competitive-programs/measures/esdr-for-land-surface-freeze-thaw-state

Klemas, V. (2012). Remote sensing of algal blooms: an overview with case studies. *J. Coast. Res.* 28, 34–43. doi: 10.2112/JCOASTRES-D-11-00051.1

Klemun, M. M., Edwards, M. R., and Trancik, J. E. (2020). Research priorities for supporting subnational climate policies. *WIREs Clim. Change* 11, e646. doi: 10.1002/wcc.646

Koh, L. P., Miettinen, J., Liew, S. C., and Ghazoul, J. (2011). Remotely sensed evidence of tropical peatland conversion to oil palm. *Proc. Natl. Acad. Sci.* 108, 5127–5132. doi: 10.1073/pnas.1018776108

Korkovelos, A., Khavari, B., Sahlberg, A., Howells, M., and Arderne, C. (2019). The role of open access data in geospatial electrification planning and the achievement of SDG7. An OnSSET-based case study for Malawi. *Energies* 12, 1395. doi: 10.3390/en12071395

Kort, E. A., Frankenberg, C., Miller, C. E., and Oda, T. (2012). Space-based observations of megacity carbon dioxide. *Geophys. Res. Lett.* 39. doi: 10.1029/2012GL052738

Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., et al. (2017). The version 3 OMI NO_2 standard product. *Atmospheric Meas. Tech.* 10, 3133–3149. doi: 10.5194/amt-10-3133-2017

Krotkov, N. A., McLinden, C. A., Li, C., Lamsal, L. N., Celarier, E. A., Marchenko, S. V., et al. (2016). Aura OMI observations of regional SO_2 and NO_2 pollution changes from 2005 to 2015. *Atmospheric Chem. Phys.* 16, 4605–4629. doi: 10.5194/acp-16-4605-2016

Kruitwagen, L., Story, K. T., Friedrich, J., Byers, L., Skillman, S., Hepburn, C. A., et al. (2021). Global inventory of photovoltaic solar energy generating units. *Nature* 598, 604–610. doi: 10.1038/s41586-021-03957-7

Kuhlmann, G., Broquet, G., Marshall, J., Clment, V., Lscher, A., Meijer, Y., et al. (2019). Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission. *Atmospheric Meas. Tech.* 12, 6695–6719. doi: 10.5194/amt-12-6695-2019

Kuhlmann, G., Henne, S., Meijer, Y., and Brunner, D. (2021). Quantifying CO_2 Emissions of Power Plants With CO_2 and NO_2 Imaging Satellites. *Front. Remote Sens.* 2, 14. doi: 10.3389/frsen.2021.689838

Kumar, A., Ramsankaran, R., Brocca, L., and Muñoz-Arriola, F. A. (2021). Simple machine learning approach to model real-time streamflow using satellite inputs: Demonstration in a data scarce catchment. *J. Hydrol.* 595, 126046. doi: 10.1016/j.jhydrol.2021.126046

Lambert, A., Hallar, A. G., Garcia, M., Strong, C., Andrews, E., Hand, J. L., et al. (2020). Dust Impacts of Rapid Agricultural Expansion on the Great Plains. *Geophys. Res. Lett.* 47, e2020GL. doi: 10.1029/2020GL090347

Lark, T. J., Hendricks, N. P., Smith, A., Pates, N., Spawn-Lee, S. A., Bougie, M., et al. (2022). Environmental outcomes of the US Renewable Fuel Standard. *Proc. Natl. Acad. Sci.* 119, e2101084119. doi: 10.1073/pnas.2101084119

Lauvaux, T., Giron, C., Mazzolini, M., D'Aspremont, A., Duren, R., Cusworth, D., et al. (2022). Global assessment of oil and gas methane ultra-emitters. *Science*. 375:557–61. doi: 10.1126/science.abj4351

Leibrand, A., Sadoff, N., Maslak, T., and Thomas, A. (2019). Using Earth observations to help developing countries improve access to reliable, sustainable, and modern energy. *Front. Environ. Sci.* 7, 123. doi: 10.3389/fenvs.2019.00123

Leidner, A., and Buchanan, G. (2018). *Satellite Remote Sensing for Conservation Action: Case Studies from Aquatic and Terrestrial Ecosystems.* Cambridge: Cambridge University Press. doi: 10.1017/9781108631129

Leland, A., Hoekman, S. K., and Liu, X. (2018). Review of modifications to indirect land use change modeling and resulting carbon intensity values within the California Low Carbon Fuel Standard regulations. *J. Clean. Prod.* 180, 698–707. doi: 10.1016/j.jclepro.2018.01.077

Levelt, P., van den Oord, G., Dobber, M., Malkki, A., Visser, H., Vries, J., et al. (2006). The ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.* 44, 1093–1101. doi: 10.1109/TGRS.2006.872333

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., et al. (2018). The ozone monitoring instrument: overview of 14 years in space. *Atmospheric Chem. Phys.* 18, 5699–5745. doi: 10.5194/acp-18-5699-2018

Li, X., Li, X., Li, D., He, X., and Jendryke, M. (2019). A preliminary investigation of Luojia-1 night-time light imagery. *Remote Sens. Lett.* 10, 526–535. doi: 10.1080/2150704X.2019.1577573

Li, X., Wagner, F., Peng, W., Yang, J., and Mauzerall, D. L. (2017). Reduction of solar photovoltaic resources due to air pollution in China. *Proc. Natl. Acad. Sci.* 114, 11867–11872. doi: 10.1073/pnas.1711462114

Li, X., Zhou, Y., Gong, P., Seto, K., and Clinton, N. (2020b). Developing a method to estimate building height from Sentinel-1 data. *Remote Sens. Environ.* 240, 111705. doi: 10.1016/j.rse.2020.111705

Li, X., Zhou, Y., Zhao, M., and Zhao, X. A. (2020a). Harmonized global nighttime light dataset 1992–2018. *Sci. Data* 7, 168. doi: 10.1038/s41597-020-0510-y

Lin, Z., Anar, M. J., and Zheng, H. (2015). Hydrologic and water-quality impacts of agricultural land use changes incurred from bioenergy policies. *J. Hydrol.* 525, 429–440. doi: 10.1016/j.jhydrol.2015.04.001

Liu, F., Beirle, S., Zhang, Q., Drner, S., He, K., Wagner, T., et al. (2016). NO_x lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations. *Atmospheric Chem. Phys.* 16, 5283–5298. doi: 10.5194/acp-16-5283-2016

Liu, F., Beirle, S., Zhang, Q., van der, A., R. J., Zheng, B., et al. (2017). NO_x mission trends over Chinese cities estimated from OMI observations during 2005 to 2015. *Atmospheric Chem. Phys.* 17, 9261–9275. doi: 10.5194/acp-17-9261-2017

Liu, F., Duncan, B. N., Krotkov, N. A., Lamsal, L. N., Beirle, S., Griffin, D., et al. (2020). Methodology to constrain carbon dioxide emissions from coal-fired power plants using satellite observations of co-emitted nitrogen dioxide. *Atmospheric Chem. Phys.* 20, 99–116. doi: 10.5194/acp-20-99-2020

Longato, D., Gaglio, M., Boschetti, M., and Gissi, E. (2019). Bioenergy and ecosystem services trade-offs and synergies in marginal agricultural lands: A remote-sensing-based assessment method. *J. Clean. Prod.* 237, 117672. doi: 10.1016/j.jclepro.2019.117672

Lu, Z., Streets, D. G., de Foy, B., Lamsal, L. N., Duncan, B. N., and Xing, J. (2015). Emissions of nitrogen oxides from US urban areas: estimation from Ozone Monitoring Instrument retrievals for 2005–2014. *Atmospheric Chem. Phys.* 15, 10367–10383. doi: 10.5194/acp-15-10367-2015

Lu, Z., Streets, D. G., Foy, d. e., and Krotkov, B. (2013). Ozone Monitoring Instrument Observations of Interannual Increases in SO_2 Emissions from Indian Coal-Fired Power Plants during 2005–2012. *Environ. Sci. Technol.* 47, 13993–14000. doi: 10.1021/es4039648

Luo, T., Krishnaswami, A., and Li, X. (2018). A Methodology to Estimate Water Demand for Thermal Power Plants in Data-Scarce Regions Using Satellite Images. *WRI.* Available online at: https://www.wri.org/research/methodology-estimate-water-demand-thermal-power-plants-data-scarce-regions-using-satellite

Madsen, B. (2021). Identifying Wetlands with Deep Learning. *EPRI* Available online at: https://storymaps.arcgis.com/stories/4f98297b48a94efbbbe0199681539980

Mahdi Elsiddig Haroun, F., Mohamad Deros, S., and Din, N. A. (2020). Review of vegetation encroachment detection in power transmission lines using optical sensing satellite imagery. arXiv preprint arXiv:2010.01757. doi: 10.30534/ijatcse/2020/8691.42020

Mahtta, R., Mahendra, A., and Seto, K. C. (2019). Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+. *Environ. Res. Lett.* 14, 124077. doi: 10.1088/1748-9326/ab59bf

Majidi Nezhad, M., Heydari, A., Pirshayan, E., Groppi, D., and Astiaso Garcia, D. A. (2021a). Novel forecasting model for wind speed assessment using sentinel family satellites images and machine learning method. *Renew. Energy* 179, 2198–2211. doi: 10.1016/j.renene.2021.08.013

Majidi Nezhad, M., Nastasi, B., Groppi, D., Lamagna, M., Piras, G., and Astiaso Garcia, D. (2021b). Green energy sources assessment using sentinel-1 satellite remote sensing. *Front. Energy Res.* 9, 649305. doi: 10.3389/fenrg.2021.649305

Malik Naureen, S. (2021). *New Climate Satellite Spotted Giant Methane Leak as It Happened.* Bloomberg News. Available online at: https://www.bloomberg.com/news/articles/2021-02-12/new-climate-satellite-spotted-giant-methane-leak-as-it-happened

Mann, M. L., Melaas, E. K., and Malik, A. (2016). Using VIIRS day/night band to measure electricity supply reliability: preliminary results from Maharashtra, India. *Remote Sens.* 8, 711. doi: 10.3390/rs8090711

Margetta, R. (2021). NASA releases climate action plan. *NASA.* Available online at: http://www.nasa.gov/press-release/nasa-releases-climate-action-plan

Matikainen, L., Lehtomki, M., Ahokas, E., Hyypp, J., Karjalainen, M., Jaakkola, A., et al. (2016). Remote sensing methods for power line corridor surveys. *ISPRS J. Photogramm. Remote Sens.* 119, 10–31. doi: 10.1016/j.isprsjprs.2016.04.011

Medina-Lopez, E., McMillan, D., Lazic, J., Hart, E., Zen, S., Angeloudis, A., et al. (2021). Satellite data for the offshore renewable energy sector:

Synergies and innovation opportunities. *Remote Sens. Environ.* 264, 112588. doi: 10.1016/j.rse.2021.112588

Meijaard, E., Brooks, T. M., Carlson, K. M., Slade, E. M., Garcia-Ulloa, J., Gaveau, D. L., et al. (2020). The environmental impacts of palm oil in context. *Nat. Plants* 6, 1418–1426. doi: 10.1038/s41477-020-00813-w

Mellors, R. J., Xu, X., Matzel, E., and Fu, P. (2018). "New potential of InSAR for geothermal Systems," in *Proceedings of the 43rd Workshop on Geothermal Reservoir Engineering* (Stanford, CA).

Milford, J. B., and Knight, D. (2017). Increasing the use of Earth science data and models in air quality management. *J. Air Waste Manag. Assoc.* 67, 431–444. doi: 10.1080/10962247.2016.1248303

Min, B., Baugh, K., Monroe, T., Goldblatt, R., Steward, B., Kosmidou-Bradley, W., et al. (2021). *Light Every Night – New nighttime light data set and tools for development*. World Bank Blogs. Available online at: https://blogs.worldbank.org/opendata/light-every-night-new-nighttime-light-data-set-and-tools-development (accessed July 11, 2022).

Min, B., Gaba, K. M., Sarr, O. F., and Agalassou, A. (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *Int. J. Remote Sens.* 34, 8118–8141. doi: 10.1080/01431161.2013.833358

Mitchell, J. W. (2013). Power line failures and catastrophic wildfires under extreme weather conditions. *Eng. Fail. Anal.* 35, 726–735. doi: 10.1016/j.engfailanal.2013.07.006

Moner-Girona, M., Bdis, K., Morrissey, J., Kougias, I., Hankins, M., Huld, T., et al. (2019). Decentralized rural electrification in Kenya: speeding up universal energy access. *Energy Sustain. Dev.* 52, 128–146. doi: 10.1016/j.esd.2019.07.009

Moner-Girona, M., Kakoulaki, G., Falchetta, G., Weiss, D. J., and Taylor, N. (2021). Achieving universal electrification of rural healthcare facilities in sub-Saharan Africa with decentralized renewable energy technologies. *Joule* 5, 2687–2714. doi: 10.1016/j.joule.2021.09.010

Moore, I. I. I., Crowell, B., Rayner, S. M., Kumer, P. J., O'Dell, J., O'Brien, C. W., et al. (2018). The potential of the geostationary carbon cycle observatory (GeoCarb) to provide multi-scale constraints on the carbon cycle in the Americas. *Front. Environ. Sci.* 6, 109. doi: 10.3389/fenvs.2018.00109

Motamed, M., McPhail, L., and Williams, R. (2016). Corn area response to local ethanol markets in the United States: a grid cell level analysis. *Am. J. Agric. Econ.* 98, 726–743. doi: 10.1093/ajae/aav095

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., et al. (2016). The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview. *Atmospheric Meas. Tech.* 9, 1279–1301. doi: 10.5194/amt-9-1279-2016

Næss, J. S., Cavalett, O., and Cherubini, F. (2021). The land–energy–water nexus of global bioenergy potentials from abandoned cropland. *Nat. Sustain.* 4, 525–536. doi: 10.1038/s41893-020-00680-5

Nalepa, R. A., and Bauer, D. M. (2012). Marginal lands: the role of remote sensing in constructing landscapes for agrofuel development. *J. Peasant Stud.* 39, 403–422. doi: 10.1080/03066150.2012.665890

NASA (2015). Jet Propulsion Laboratory AIRS Animation Shows Carbon Monoxide Plume Transport from U.S. Fires. *AIRS* Available online at: https://airs.jpl.nasa.gov/news/136/airs-animation-shows-carbon-monoxide-plume-transport-from-us-fires (accessed July 11, 2022).

NASA (2016). Earth Observatory Satellite Detects Human Contribution to Atmospheric CO_2. Available online at: https://earthobservatory.nasa.gov/images/89117/satellite-detects-human-contribution-to-atmospheric-co2

NASA (2021). Fire Information for Resource Management System (FIRMS) | Earthdata. Available online at: https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/ (accessed July 11, 2022).

NASA Earth Observatory (2020). Airborne Nitrogen Dioxide Plummets Over China. Available online at: https://earthobservatory.nasa.gov/images/146362/airborne-nitrogen-dioxide-plummets-over-china (accessed July 11, 2022).

NASA Earth Observatory (2021a). Plugging-in Sub-Saharan Africa. Available online at: https://earthobservatory.nasa.gov/images/148069/plugging-in-sub-saharan-africa

NASA Earth Observatory (2021b). Satellites View California Oil Spill. Available online at: https://earthobservatory.nasa.gov/images/148929/satellites-view-california-oil-spill

NASA Earth Observatory (2022). Iberian Peninsula Drought. Available online at: https://earthobservatory.nasa.gov/images/149469/iberian-peninsula-drought (accessed July 11, 2022).

NASA Earthdata (2021). Nighttime Lights. Available online at: https://earthdata.nasa.gov/learn/backgrounders/nighttime-lights/

NASA Goddard Space Flight Center (2021). *NASA Black Marble Product*. Available online at: https://viirsland.gsfc.nasa.gov/Products/NASA/BlackMarble.html

NASA POWER and NASA ARSET (2021). *How to Utilize NASA Resources for Energy Efficiency and Renewable Energy Applications*. Available online at: https://appliedsciences.nasa.gov/sites/default/files/2021-06/Energy_Management_Part4.pdf

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., Crisp, D., et al. (2017). Quantifying CO_2 emissions from individual power plants from space. *Geophys. Res. Lett.* 44, 10,045–10,053 doi: 10.1002/2017GL074702

Nassar, R., Mastrogiacomo, J. P., Bateman-Hemphill, W., McCracken, C., MacDonald, C. G., et al. (2021). Advances in quantifying power plant CO_2 emissions with OCO-2. *Remote Sens. Environ.* 264, 112579. doi: 10.1016/j.rse.2021.112579

Nazir, H. (2021). Lessons Learned from the February 2021 Texas Power Outage. *Canadian Energy Research Institute* Available online at: https://ceri.ca/market-updates/lessons-learned-from-the-february-2021-texas-power-outage

Nik, V. M., Perera, A. T. D., and Chen, D. (2021). Towards climate resilient urban energy systems: a review. *Natl. Sci. Rev.* 8, nwaa134. doi: 10.1093/nsr/nwaa134

Noktosat (2021). *Noktosat Satellite Images - The earth at night in colour*. Available online at: https://www.noktosat.com

Pacific Gas and Electric Company (2021). 2021 Wildfire Mitigation Plan Report. Available online at: https://energysafety.ca.gov/wp-content/uploads/docs/wmp-2021/pge-2021-wmp-update.pdf

Pandey, B., Brelsford, C., and Seto, K. C. (2022). Infrastructure inequality is a characteristic of urbanization. *Proc. Natl. Acad. Sci.* 119, e2119890119. doi: 10.1073/pnas.2119890119

Park, H., Jeong, S., Park, H., Labzovskii, L. D., and Bowman, K. W. (2021). An assessment of emission characteristics of Northern Hemisphere cities using spaceborne observations of CO_2, CO, and NO_2. *Remote Sens. Environ.* 254, 112246. doi: 10.1016/j.rse.2020.112246

Pelz, S., Pachauri, S., and Groh, S. (2018). A critical review of modern approaches for multidimensional energy poverty measurement. *WIREs Energy Environ.* 7, e304. doi: 10.1002/wene.304

Pelz, S., Pachauri, S., and Rao, N. (2021). Application of an alternative framework for measuring progress towards SDG 7.1. *Environ. Res. Lett.* 16, 084048. doi: 10.1088/1748-9326/ac16a1

Perez, R., Cebecauer, T., and Šúri, M. (2013). "Chapter 2 - Semi-Empirical Satellite Models," in *Solar Energy Forecasting and Resource Assessment* (ed. Kleissl, J.) p. 21–48 (Academic Press). doi: 10.1016/B978-0-12-397177-7.00002-4

Peters, I. M., Liu, H., Reindl, T., and Buonassisi, T. (2018). Global prediction of photovoltaic field performance differences using open-source satellite data. *Joule.* 2, 307–322. doi: 10.1016/j.joule.2017.11.012

Pfenninger, S., and Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* 114, 1251–1265. doi: 10.1016/j.energy.2016.08.060

Popp, J., Lakner, Z., Harangi-Rákos, M., and Fári, M. (2014). The effect of bioenergy expansion: Food, energy, and environment. *Renew. Sustain. Energy Rev.* 32, 559–578. doi: 10.1016/j.rser.2014.01.056

Ravikumar, A. P., Roda-Stuart, D., Liu, R., Bradley, A., Bergerson, J., Nie, Y., et al. (2020). Repeated leak detection and repair surveys reduce methane emissions over scale of years. *Environ. Res. Lett.* 15, 034029. doi: 10.1088/1748-9326/ab6ae1

Reames, T. G. (2016). Targeting energy justice: Exploring spatial, racial/ethnic and socioeconomic disparities in urban residential heating energy efficiency. *Energy Policy* 97, 549–558. doi: 10.1016/j.enpol.2016.07.048

Reba, M., and Seto, K. (2020). A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Remote Sens. Environ.* 242, 111739. doi: 10.1016/j.rse.2020.111739

Receveur, M., Sigmundsson, F., Drouin, V., and Parks, M. (2019). Ground deformation due to steam cap processes at Reykjanes, SW-Iceland: effects of geothermal exploitation inferred from interferometric analysis of Sentinel-1 images 2015–2017. *Geophys. J. Int.* 216, 2183–2212. doi: 10.1093/gji/ggy540

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., et al. (2019). Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. *Atmospheric Chem. Phys.* 19, 9371–9383. doi: 10.5194/acp-19-9371-2019

Román, M. O., and Stokes, E. C. (2015). Holidays in lights: Tracking cultural patterns in demand for energy services. *Earths Future.* 3, 182–205. doi: 10.1002/2014EF000285

Romn, M. O., Stokes, E. C., Shrestha, R., Wang, Z., Schultz, L., Carlo, E. A. S., et al. (2019). Satellite-based assessment of electricity restoration efforts in Puerto Rico after Hurricane Maria. *PLoS ONE* 14, e0218883. doi: 10.1371/journal.pone.0218883

Romn, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., et al. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sens. Environ.* 210, 113–143. doi: 10.1016/j.rse.2018.03.017

Rose, R. A., Byler, D., Eastman, J. R., Fleishman, E., Geller, G., Goetz, S., et al. (2015). Ten ways remote sensing can contribute to conservation. *Conserv. Biol.* 29, 350–359. doi: 10.1111/cobi.12397

Rosenzweig, C., Horton, R. M., Bader, D. A., Brown, M. E., DeYoung, R., Dominguez, O., et al. (2014). Enhancing climate resilience at NASA centers: a collaboration between science and stewardship. *Bull. Am. Meteorol. Soc.* 95, 1351–1363. doi: 10.1175/BAMS-D-12-00169.1

RSPO (2021). - *Roundtable on Sustainable Palm Oil*. Hotspot Hub — RSPO strengthens fire prevention with satellite technology. Available online at: https://rspo.org/news-and-events/news/hotspot-hub--rspo-strengthens-fire-prevention-with-satellite-technology (accessed July 11, 2022).

Rutan, D. A., Kato, S., Doelling, D. R., Rose, F. G., Nguyen, L. T., Caldwell, T. E., et al. (2015). Synoptic product: methodology and validation of surface radiant flux. *J. Atmospheric Ocean. Technol.* 32, 1121–1143. doi: 10.1175/JTECH-D-14-00165.1

San Diego Gas and Electric Company (2021). 2020-2022 Wildfire Mitigation Plan Update. Available online at: https://energysafety.ca.gov/wp-content/uploads/docs/misc/wmp/2021/utility/sdge/sdge-2021-wmp-update.pdf (accessed July 11, 2022).

Sanders, K. T., and Masri, S. F. (2016). The energy-water agriculture nexus: the past, present and future of holistic resource management via remote sensing technologies. *J. Clean. Prod.* 117, 73–88. doi: 10.1016/j.jclepro.2016.01.034

Schneising, O., Heymann, J., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., et al. (2013). Anthropogenic carbon dioxide source areas observed from space: assessment of regional enhancements and trends. *Atmospheric Chem. Phys.* 13, 2445–2454. doi: 10.5194/acp-13-2445-2013

Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., Shelby, J., et al. (2018). The National Solar Radiation Data Base (NSRDB). *Renew. Sustain. Energy Rev.* 89, 51–60. doi: 10.1016/j.rser.2018.03.003

Seto, K., Dhakal, S., Bigio, A., Blanco, H., Delgado, G., Dewar, D., et al. (2014). *Chapter 12 - Human Settlements, Infrastructure and Spatial Planning*. Cambridge: Cambridge University Press.

Shen, L., Xu, H., and Guo, X. (2012). Satellite remote sensing of harmful algal blooms (HABs) and a potential synthesized framework. *Sensors* 12, 7778–7803. doi: 10.3390/s120607778

Sierk, B., Bezy, J-. L., Löscher, A., and Meijer, Y. (2019). "The European CO_2 Monitoring Mission: observing anthropogenic greenhouse gas emissions from space," in *International Conference on Space Optics—ICSO 2018* (Chania: SPIE), 11180, 237–250. doi: 10.1117/12.2535941

Silva, S., and Arellano, A. (2017). Characterizing Regional-Scale Combustion Using Satellite Retrievals of CO, NO_2 and CO_2. *Remote Sens.* 9, 744. doi: 10.3390/rs9070744

Smith, W. K., Cleveland, C. C., Reed, S. C., Miller, N. L., and Running, S. W. (2012). Bioenergy potential of the united states constrained by satellite observations of existing productivity. *Environ. Sci. Technol.* 46, 3536–3544. doi: 10.1021/es203935d

Soukissian, T. H., and Papadopoulos, A. (2015). Effects of different wind data sources in offshore wind power assessment. *Renew. Energy* 77, 101–114. doi: 10.1016/j.renene.2014.12.009

Southern California Edison Company (2021). 2021 Wildfire Mitigation Plan Update. Available online at: https://energysafety.ca.gov/wp-content/uploads/docs/wmp-2021/sce-2021-wmp-update.pdf (accessed July 11, 2022).

Sridhar, J. B. M., and Vincent, R. K. (2009). Mapping and Estimation of Phosphorus and Copper Concentrations in Fly Ash Spill Area using LANDSAT TM Images. *Sci. Total Environ.* 407, 1031. doi: 10.1016/j.scitotenv.2009.01.021

Stackhouse, P. W., Cox, S. J., Mikovitz, J. C., and Zhang, T. (2021). *GEWEX (Global Energy and Water Exchanges Project): Surface Radiation Budget (SRB) Release 4 Integrated Product (IP4) - Algorithm Theoretical Basis Document and Evaluation*.

Steele, J. E., Sundsy, P. R., Pezzulo, C., Alegana, V. A., Bird, T. J., Blumenstock, J., et al. (2017). Mapping poverty using mobile phone and satellite data. *J. R. Soc. Interface* 14, 20160690. doi: 10.1098/rsif.2016.0690

Stiles, G., and Murove, C. (2019). 2018. SADC Renewable Energy and Energy Efficiency Status Report. *REN21*. (2019). Available online at: https://www.ren21.net/2018-sadc-renewable-energy-and-energy-efficiency-status-report/ (accessed July 11, 2022).

Stokes, E., and Seto, K. (2019). Characterizing urban infrastructural transitions for the Sustainable Development Goals using multi-temporal land, population, and nighttime light data. *Remote Sens. Environ.* 234, 111430. doi: 10.1016/j.rse.2019.111430

Stokes, E. C., and Román, M. O. (2022). Tracking COVID-19 urban activity changes in the Middle East from nighttime lights. *Sci. Rep.* 12, 8096. doi: 10.1038/s41598-022-12211-7

The White House (2021a). Executive Order on Tackling the Climate Crisis at Home and Abroad. Available online at: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/executive-order-on-tackling-the-climate-crisis-at-home-and-abroad/

The White House (2021b). FACT SHEET: President Biden Takes Executive Actions to Tackle the Climate Crisis at Home and Abroad, Create Jobs, and Restore Scientific Integrity Across Federal Government. Available online at: https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/27/fact-sheet-president-biden-takes-executive-actions-to-tackle-the-climate-crisis-at-home-and-abroad-create-jobs-and-restore-scientific-integrity-across-federal-government/

Timmermans, R., Segers, A., Curier, L., Abida, R., Atti, J. L., El Amraoui, L., et al. (2019). Impact of synthetic space-borne NO_2 observations from the Sentinel-4 and Sentinel-5P missions on tropospheric NO_2 analyses. *Atmospheric Chem. Phys.* 19, 12811–12833. doi: 10.5194/acp-19-12811-2019

Tollenaar, M., Fridgen, J., Tyagi, P., Stackhouse, P. W., and Kumudini, S. (2017). The contribution of solar brightening to the US maize yield trend. *Nat. Clim. Change* 7, 275–278. doi: 10.1038/nclimate3234

Tsao, C-. C., Campbell, J. E., Mena-Carrasco, M., Spak, S. N., Carmichael, G. R., et al. (2012). Increased estimates of air-pollution emissions from Brazilian sugar-cane ethanol. *Nat. Clim. Change* 2, 53–57. doi: 10.1038/nclimate1325

U.S. Department of Energy (2021). 2021 Climate Adaptation and Resilience Plan. Available online at: https://www.energy.gov/sites/default/files/2021-09/DOEClimateAdaptationandResiliencePlanSept21.pdf (accessed July 11, 2022).

UN Environment Programme (2021). *New global methane pledge aims to tackle climate change*. UNEP Available online at: http://www.unep.org/news-and-stories/story/new-global-methane-pledge-aims-tackle-climate-change(accessed July 11, 2022).

US Department of Commerce (2021). Department of Commerce Releases Climate Action Plan for Adaptation and Resilience. Available online at: https://www.commerce.gov/news/press-releases/2021/10/department-commerce-releases-climate-action-plan-adaptation-and (accessed July 11, 2022).

US EPA (2010). Office of Transportation and Air Quality, Assessment and Standards Division. Renewable Fuel Standard Program (RFS2) Regulatory Impact Analysis. Available online at: https://www.eesi.org/files/420r10006.pdf (accessed July 11, 2022).

van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., et al. (2010). Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. *Environ. Health Perspect.* 118, 847–855. doi: 10.1289/ehp.0901623

Vargas, C. A., Caracciolo, L., and Ball, P. J. (2022). Geothermal energy as a means to decarbonize the energy mix of megacities. *Commun. Earth Environ.* 3, 1–11. doi: 10.1038/s43247-022-00386-w

Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S., et al. (2019). Satellite Discovery of Anomalously Large Methane Point Sources From Oil/Gas Production. *Geophys. Res. Lett.* 46, 13507–13516. doi: 10.1029/2019GL083798

Veefkind, J. P., Aben, I., McMullan, K., Frster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.* 120, 70–83. doi: 10.1016/j.rse.2011.09.027

Velazco, V. A., Buchwitz, M., Bovensmann, H., Reuter, M., Schneising, O., Heymann, J., et al. (2011). Towards space based verification of CO_2 emissions from strong localized sources: fossil fuel power plant emissions as seen by a CarbonSat constellation. *Atmospheric Meas. Tech.* 4, 2809–2822. doi: 10.5194/amt-4-2809-2011

Wagle, P., Bhattarai, N., Gowda, P. H., and Kakani, V. G. (2017). Performance of five surface energy balance models for estimating daily evapotranspiration in high biomass sorghum. *ISPRS J. Photogramm. Remote Sens.* 128, 192–203. doi: 10.1016/j.isprsjprs.2017.03.022

Wang, E., Cruse, R. M., Sharma-Acharya, B., Herzmann, D. E., Gelder, B. K., James, D. E., et al. (2020). Strategic switchgrass (Panicum virgatum) production within row cropping systems: Regional-scale assessment of soil erosion loss and water runoff impacts. *GCB Bioenergy* 12, 955–967. doi: 10.1111/gcbb.12749

Wang, S., Zhang, Y., Hakkarainen, J., Ju, W., Liu, Y., Jiang, F., et al. (2018). Distinguishing Anthropogenic CO_2 emissions from different energy intensive industrial sources using OCO-2 observations: a case study in northern China. *J. Geophys. Res. Atmospheres* 123, 9462–9473. doi: 10.1029/2018JD029005

Wheeler, D., and Ummel, K. (2008). *Calculating Carma: Global Estimation of CO_2 Emissions from the Power Sector*. Available online at: https://papers.ssrn.com/abstract=1138690 doi: 10.2139/ssrn.1138690 (accessed July 11, 2022).

Witman, S., Halloway, T., and Reddy, P. J. (2014). Integrating Satellite Data into Air Quality Management. Experience from Colorado. *NASA* Available online at: https://pubs.awma.org/gsearch/em/2014/2/witman.pdf (accessed July 11, 2022).

World Bank (2019). Tracking SDG7: The Energy Progress Report 2019. Available online at: https://www.worldbank.org/en/topic/energy/publication/tracking-sdg7-the-energy-progress-report-2019 (accessed July 11, 2022).

World Health Organization (2016). WHO releases country estimates on air pollution exposure and health impact. Available online at: https://www.who.int/news/item/27-09-2016-who-releases-country-estimates-on-air-pollution-exposure-and-health-impact (accessed July 11, 2022).

Wright, C. K., Larson, B., Lark, T. J., and Gibbs, H. K. (2017). Recent grassland losses are concentrated around U.S. ethanol refineries. *Environ. Res. Lett.* 12, 044001. doi: 10.1088/1748-9326/aa6446

Wu, C., Wang, Q., Yang, Z., and Wang, W. (2007). Monitoring heated water pollution of the DaYaWan nuclear power plant using TM images. *Int. J. Remote Sens.* 28, 885–890. doi: 10.1080/01431160600580616

Wu, D., Lin, J. C., Oda, T., and Kort, E. A. (2020). Space-based quantification of per capita CO_2 emissions from cities. *Environ. Res. Lett.* 15, 035004. doi: 10.1088/1748-9326/ab68eb

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., et al. (2009). Global Concentrations of CO_2 and CH_4 Retrieved from GOSAT: first preliminary results. *Sola.* 5, 160–163. doi: 10.2151/sola.2009-041

Yu, J., Wang, Z., Majumdar, A., and Rajagopal, R. (2018). DeepSolar: a machine learning framework to efficiently construct a solar deployment database in the United States. *Joule* 2, 2605–2617. doi: 10.1016/j.joule.2018.11.021

Zavala-Araiza, D., Alvarez, R. A., Lyon, D. R., Allen, D. T., Marchese, A. J., Zimmerle, D. J., et al. (2017). Super-emitters in natural gas infrastructure are caused by abnormal process conditions. *Nat. Commun.* 8, 14012. doi: 10.1038/ncomms14012

Zell, E., Engel-Cox, J., Eckman, R., and Stackhouse, P. (2008). Application of satellite sensor data and models for energy management. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 1, 5–17. doi: 10.1109/JSTARS.2008.2001142

Zen, S., Hart, E., and Mendia-Lopez, E. (2021). The use of satellite products to assess spatial uncertainty and reduce life-time costs of offshore wind farms. *Clean. Environ. Syst.* 2, 100008. doi: 10.1016/j.cesys.2020.100008

Zhang, Y., Rossow, W., Lacis, A., Oinas, V., and Mishchenko, M. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J Geophys. Res. Atmospheres*. 109, 4457. doi: 10.1029/2003JD004457

Zhao, M., Zhou, Y., Li, X., Cao, W., He, C., Yu, B., et al. (2019). Applications of Satellite Remote Sensing of Nighttime Light Observations: Advances, Challenges, and Perspectives. *Remote Sens.* 11, 1971. doi: 10.3390/rs11171971

Zheng, T., Nassar, R., and Baxter, M. (2019). Estimating power plant CO_2 emission using OCO-2 XCO_2 and high resolution WRF-Chem simulations. *Environ. Res. Lett.* 14, 085001. doi: 10.1088/1748-9326/ab25ae

Zhu, Y., Xu, D., Ali, S. H., Ma, R., and Cheng, J. (2019a). Can nighttime light data be used to estimate electric power consumption? New evidence from causal-effect inference. *Energies* 12, 3154. doi: 10.3390/en12163154

Zhu, Z., Zhou, Y., Seto, K. C., Stokes, E. C., Deng, C., Pickett, S. T. A., et al. (2019b). Understanding an urbanizing planet: strategic directions for remote sensing. *Remote Sens. Environ.* 228, 164–182. doi: 10.1016/j.rse.2019.04.020

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., et al. (2017). Tropospheric emissions: monitoring of pollution (TEMPO). *J. Quant. Spectrosc. Radiat. Transf.* 186, 17–39. doi: 10.1016/j.jqsrt.2016.05.008

Zumkehr, A., and Campbell, J. E. (2013). Historical U.S. cropland areas and the potential for bioenergy production on abandoned croplands. *Environ. Sci. Technol.* 47, 3840–3847. doi: 10.1021/es3033132

Conflict of Interest: Authors EC and ÁE were employed by GHGSat Inc. Author MB was employed by Booz Allen Hamilton, Inc. Author AL was employed by Science and Technology Corporation.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Sustainability

TYPE Perspective
PUBLISHED 04 October 2022
DOI 10.3389/frsus.2022.943202

OPEN ACCESS

EDITED BY
Christopher Mutel,
Paul Scherrer Institut (PSI), Switzerland

REVIEWED BY
Leena Kakkori,
University of Eastern Finland, Finland

*CORRESPONDENCE
Go Okui
okui.go.5x@kyoto-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 13 May 2022
ACCEPTED 09 September 2022
PUBLISHED 04 October 2022

CITATION
Okui G and Takeda S (2022) Revisiting the measure of development: A critique of sustainametrics. *Front. Sustain.* 3:943202. doi: 10.3389/frsus.2022.943202

Revisiting the measure of development: A critique of sustainametrics

Go Okui[1]* and Shutaro Takeda[2]

[1]Bioethics and Philosophy Group, Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto, Japan, [2]Urban Institute, Kyushu University, Fukuoka, Japan

For sustainametrics to gain a firm ground as an effective concept, the meaning of development shall be revisited first without depending on any statistical measurement. The word "development" originally meant the act of disclosure or opening a cover to disclose what is inside. Martin Heidegger (1889–1976) analyzed the significance of *alētheuein*, or "to bring the world out of its hidden and covered state and into ours," and explicates that the *alētheia* under the condition of modern technology is dominated by a mode of revealing that is destructive to the earthly beings. Here, the danger inherent in the essence of technology, i.e., enframing [*Ge-stell*], renders human beings incapable of encountering the essence of beings as they are challenged and demanded to frame everything they encounter, including themselves, as mere variables. In contrast to Heidegger's thinking as releasement [*Gelassenheit*], Hannah Arendt's (1906–1975) conception of disclosure is closely tied to action. Following Heidegger's and Arendt's threads of thought, the authors conclude that any measures of development must be fundamentally grounded in disclosure through speech and action in the public realm. In this respect, the experts on the sustainametrics shall inspire fellow citizens to join the discourse by taking the risk of acting and speaking in public, disclosing who they are and what it is really meant for us. The course of development must ultimately be grounded in such an act of disclosure, only through which we may find something worth sustaining in our future development, and sustainametrics is no exception.

KEYWORDS

development, measure, disclosure, Heidegger, Arendt

Introduction

Sustainametrics addresses the question of the measurability of the objects of sustainable development. Sen's Human Development Index and Arrow's Inclusive Wealth Index have laid the ground toward the conception of sustainable development goals and thereby marked the beginning of this question, which is by no means closed at this point. We may recall that these indexes were the products of welfare economics, which have taken pains to articulate, quantify and measure wellbeing that does not appear in the market when left alone. Precisely because they deal with such hidden values that await articulation, measures of this kind will inevitably call for the question of legitimacy:

Who should quantify wellbeing under what right, or even about what and how wellbeing ought to be quantified? There is always an element of arbitrariness in these decisions, no matter how they may be formulated.

The arbitrariness inherent in such decisions reminds those of us who have come to uphold the motto "leave no one behind" that experts are no longer the only ones with the prerogative to adjudicate them. The arbitration on measures of development concerns all of us, including future generations to come. In this respect, the Great Acceleration, the exacerbation of sustainability indicators in the second half of the 20th century that coincided with the technocratic adoption of GDP as the sole indicator of development, reminds us of the grave weight of our task ahead. With this in mind, perhaps we can begin by looking back and thinking about what development has meant to us beyond the surface.

Development and metrics

The immanent problems we face today concerning the global environment are problems that we have caused ourselves through our own activities of development with the advances of modern technology. Today, it has become almost self-evident that development in this context is used synonymously with growth and progress. The tripartite association of productivity growth, social progress and development was already evident in Marx, but, it was not until after World War II that its amalgamation acquired seamless façade of calculability (Coyle, 2014).

The concept of GDP, i.e., gross domestic product, is rooted in the national income calculation first presented by Kuznets in a report submitted to the US Congress in 1934. The national income calculation used herein later led to the concept of GDP. However, Kuznets was concerned about such measures to be used as a deterministic instrument for the country's management as they create "illusions" by oversimplifying the object of what is being measured and "invite abuse" in conflicts between antagonistic social groups (Kuznets, 1934).

Whether it is GDP or susteinametrics, they share the fact that they cannot operate without the use of statistics. The word statistics is rooted in the medieval Latin *statisticum*, meaning "the affairs of states," and is a loanword from the German word *Statistik*, which was originally introduced as "the study of the matters pertaining to the prosperity of empires and states" by Gottfried Achenwall in 1749 (Meitzen and Falkner, 1891; Onions, 1994). Whether it is preparation for war, growth, progress or sustainability, statistics always implicitly point to the direction toward which a community ought to move forward. In other words, what is implicit in the use of statistics as a measure of development is a particular political standpoint from which certain objects are perceived as good under certain teleological end. No matter how effective a certain measure seems to statisticians, should they promote its adoption without endorsing critical examination of its underlining assumption, it will end up failing us as surreptitious propaganda cloaked in the guise of numerical rationality. This point brings us back to the problem at hand which calls for us to firmly grasp the meaning of development without depending on any mathematical formula, any statistical measurement for its definition or the illusion of growth or progress that previous measures projected on us for half a century.

The meaning of development

The word "development" has not always meant growth and progress, nor does economics have a prerogative over its definition. The word "develop" is a variant of "disvelop," which had been in use until around the 17th century. While development in modern English is a form later influenced by modern French "*developer*," "disvelop" is a loan word from Old French "*desveloper*," the earliest variant of which can be found in *chanson de geste* of *Aiol* from 12th century (Normand and Raynaud, 1877; Greimas, 1969; Onions, 1994; Hartman and Malicote, 2014). The negative prefix, "des-," creates an antonym of "velop," which means "to envelop." In this sense, "development" means the act of disclosure or opening a cover of something to disclose what is inside. The meaning of "growth" and "progress" is later derived from the meaning of disclosure and was not originally linked to the meaning of development.

What does it mean to grasp development in the sense of disclosure? Though there are no preceding studies of this kind, we are not left without clues in addressing this question. It may seem surprising to some readers of this volume, but in this paper, we would like to refer to the contribution of Martin Heidegger (1889–1976) and Hannah Arendt (1906–1975) particularly with regard their conception of disclosure as a guiding thread to the issue at hand.

An important clue to examining the meaning of disclosure in Heidegger can be found in the lecture notes on Plato's Sophists given at the University of Marburg in the winter semester of 1924-5. Heidegger devoted more than a hundred pages of his lengthy introduction to an analysis of the Greek word *alētheia*, the etymology of truth, through a detailed interpretation of Aristotle's *Nichomachean Ethics* and *Metaphysics*. In the introduction, Heidegger articulates the composition of the word *alētheia* by breaking down its prefix "*a*-," indicating absence, and its stem "*lētheia*," indicating a hidden state or a state of oblivion. In contrast to the common understanding of truth as *adequation intellectus et rei*, he explicates the fundamental meaning of *alētheia* as the state of being no longer hidden, that is, "disclosure [*Das Erschließen*] [1]" (Heidegger, 1967, 1992, 1997).

Heidegger further analyzes the significance of the verb form of *alētheia*, *alētheuein*, as "to bring the world out of its hidden and covered state and into ours," and stresses that this act of disclosing "appears first in speaking, that is, in speaking with one another, in *legein*" (Heidegger, 1992, 1997). This suggests that the activity denoted by the verb *alētheuein* has a fundamental significance for the existence of human beings, understood by Aristotle as a living being that has language, i.e., *zoon logon echon*, corresponding to the human experience of speaking. Heidegger discerns that "*legein* or to speak constitutes a human being in the most fundamental sense," since "[i]n speaking, Being expresses itself—by speaking about something, about the world" (Heidegger, 1992, 1997). In other words, the disclosing act of speaking about the world constitutes being human so fundamentally, that is in relation to Being, that it takes precedence over all the other activities.

Here, one may question how the act of disclosing fares in the modern world, to which Ancient Greek city-states might seem nothing but a distant past. Tracing Heidegger's interpretation of *alētheia* to his later analysis of technology provides an insight into how a diminutive understanding of development can have catastrophic consequences for human beings whose existence and activities are grounded in the experience of the act of disclosure. Heidegger's later interpretation has important implications in considering what "development" means to human beings in the modern time in which "development," understood as growth and progress, is overshadowing everything from global policies to the minutiae of everyday life while forcing its yardstick onto everything it encounters, debasing its meaning as a mere means to an end.

In his later work concerning the question of technology, Heidegger explicates that the *alētheia* under the condition of modern technology is dominated by a mode of revealing [*Entbergen*] that is destructive to the earthly beings. This mode of *alētheia* endlessly engages in the activities of setting up artifacts while challenging nature to give up what is in store. Hence,

> *The energy concealed in nature is unlocked, what is unlocked is transformed, what is transformed is stored up, what is stored up is, in turn, distributed, and what is distributed is switched about ever anew. Unlocking, transforming, storing, distributing, and switching about are ways of revealing. But the revealing never simply comes to an end* (Heidegger, 1977, 2000).

A resonance between modern development and this mode of *alētheia* in the sense *Entbergen* can hardly be denied. What is described here is precisely what we have witnessed under the name of development leading into the era of Anthropocene where every being the development encounters get stripped of their essence as they are commodified and thrown into the endless vortex of economic activities, wherein human beings are no exception. Here, "the revealing never simply comes to an end" nor does it simply runoff uncontrollably.

> *The revealing reveals to itself its own manifoldly interlocking paths, through regulating their course. This regulating[Steuerung] itself is, for its part, everywhere secured. Regulating and securing even become the chief characteristics of the challenging revealing.* (Heidegger, 1977, 2000).

We may recall that measures, whether it is GDP or susteinametrics, are deemed necessary precisely because they are needed to regulate and steer the course of development. Heidegger calls what he sees in the essence of technology "*Ge-Stell*" or "enframing." There is an inherent danger to *Ge-Stell*, where "human beings are caught [*gestellt]*, demanded, and challenged by a force that is revealed in the essence of technology." Being caught as such, human beings are rendered incapacitated to encounter their own essence, i.e., who they are, as they are challenged and demanded to frame everything they encounter, including human beings and even themselves as mere variables. Under the reign of technological rationality, it becomes impossible to govern otherwise as everything becomes framed on its accord. Thus, according to Heidegger, "the technological state would be the most obsequious and blind servant in the face of the reign of technology."

In an interview with Spiegel, who asked what one can do about these potential dangers of modern technology, Heidegger remarked:

> *Philosophy will be unable to effect any immediate change in the current state of the world. This is true not only of philosophy but of all purely human reflection and endeavor. Only a god can save us. The only possibility available to us is that by thinking and poetizing we prepare a readiness for the appearance of a god, or for the absence of a god in [our] decline, insofar as in view of the absent god we are in a state of decline* (Sheehan, 1981).

---

1 Through the course of his work, he produced many variants of the terminology inspired by *alētheia*, to describe briefly, such as disclosedness [*Erschlossenheit*] of Dasein and the world, unconcealedness [*Unverborgenheit*], uncoveredness [*Enthüllheit*] of being, discoveredness[*Entdecktheit*] of present-at-hand, and revealing [*Entbergung*] of standing-reserves (Inwood, 1999). They have different connotations according to the context in which they are placed. Nonetheless, what remains constant about *alētheia* or "truth understood as Un-hiddenness or Unconcealment, is always on the side of Being," with the sole exception found in the exegesis of Anaximander fragment (Arendt, 1978).

Disclosure as action

In such Heidegger's thoughts on technology, Arendt discerns the avoidance of action in the comportment of *Gelassenheit*, e.i., the "Will-not-to-will," where "the actions of men are inexplicable by themselves and can be understood only as of the work of some hidden purpose or some hidden actor" (Heidegger, 1969; Arendt, 1978). In contrast, Arendt's conception of disclosure (disclosure/*Enthüllung*) is closely tied to human action. For the sake of an action to be "fully revealed," it must be seen, in the void of propaganda that dazzles everybody, and thus requires "the shining brightness we once called glory and which is possible only in the public realm" (Arendt, 1998). In other words, it is only in the uncovered *openness* of the public sphere [*Der öffentliche Raum*], brightly illuminated by being seen by plural beings, that it is possible for the actor to appear through his action (Arendt, 1998, 2015). Needless to say, development, hence, must be fundamentally be grounded in such disclosive action.

Here, it is important to note Arendt's emphasis on the etymological meaning of action, i.e., *archein*, which originally meant to begin something new but came to be understood predominantly to rule in the western tradition since Plato (Arendt, 1998). Politically speaking, rulership finds its expression in the notion of sovereignty, which is rooted in *majestas* in Latin. According to Arendt, one of the decisive differences between American Revolution and French Revolution has to do with their relation to sovereignty. "National sovereignty" is "the majesty of public realm itself as it had come to be understood in the long centuries of absolute kingship" (Arendt, 1990). Since it demands "undivided centralized power," it contradicts "the establishment of a republic" in principle as was seen in the failures of the French revolution and subsequent rise of European nation-states (Arendt, 1990). What was revealed in American Revolution was "an entirely new concept of power and authority," where those who are elected to constitute body politic received the power and authority from the below as "they held fast to the Roman principle that the seat of power lay in the people" (Arendt, 1990). What were defeated or in the European revolutions were council systems which held the same principle of organization as American "townships" of people from which the power to constitute sprang (Arendt, 1990).

Disclosure and measure

The above discussion suggests that any measure in development must be fundamentally grounded in disclosure through speech and action in the public realm. Through the action of beginning something new, then, how can we determine the measure of development? With the issue of statistical indicators in mind, I would like to consider Arendt's interpretation of Solon's "*aphanes metron*" i.e., "non-appearing" or "invisible measure" toward understanding measures of sustainable development.

In her post-humously published book titled *The Life of the Mind*, Arendt quotes Solon's reference to measure in passing. In the passage to which Arendt refers, he says, "it is difficult to see the invisible measure that alone determines the limits of all things" (Tyrtaeus and Theognis, 1999). Since this measure is *aphanes*, i.e., non-appearing or invisible, it concerns things that are "indicated to my senses by what I have seen, though they themselves are not present in sense perception," such as happiness or courage (Arendt, 1978). Hence, Solon answered to Croesus, the king of Lydia renowned for his wealth, that wealth is not what determines one's happiness, but the "invisible measure" of happiness.

According to Arendt, Solon's "invisible measure" corresponds to what was later called "Idea" by Plato and has come to be understood as "concept" in modern times. It is what is conceived in such words as "courage," "justice," "knowledge," and "beauty," nouns derived from words describing the scene of particular events that occurred and appeared as such.

The "invisible measure" includes not only concepts that have been the object of philosophical inquiry, such as "justice" and "beauty," but also more mundane concepts such as "house." Words such as "development" and "measure" can also be counted among these concepts. These words, Arendt emphasizes, are "like a frozen thought that thinking must unfreeze whenever it wants to find out the original meaning" (Arendt, 1978). Thinking, in this sense, "inevitably leads to the destruction and overthrow of all established standards, values, and measures of right and wrong," that is, "the habits and rules of behavior dealt with in morals and ethics." On the side of common sense, thought is indeed fraught with these dangers, but what is even more dangerous for us is the desire for results in thought and the desire to escape from thinking. Arendt says the following about this.

> *[N]onthinking, which seems so recommendable a state for political and moral affairs, also has its dangers. By shielding people against the dangers of examination, it teaches them to hold fast to whatever the prescribed rules of conduct may be at a given time in a given society. What people then get used to is not so much the content of the rules, a close examination of which would always lead them into perplexity, as the possession of rules under which to subsume particulars* (Arendt, 2003).

This quoted passage is also instructive for those of us who are pondering on the question of measures that should instruct the course of sustainable development. We may indeed have become accustomed to the long-lasting "possession of rules" under the dominance of gross domestic product. If we try to offer a new measure for sustainable development, only to teach people to behave according to the rationality built into such measures, we will end up reproducing another "non-thinking" calculations and behaviors. If a society is once again unthinkingly oriented toward a certain measure, it will be forced to continue to harbor potential crises. Herein lies one of the limitations of development measures.

Conclusion

What Arendt sought to clarify in her late studies of political philosophy, which revolved around Kant's Critique of Judgment, was the relationship between thought and action, which is linked through our ability to judge. Thinking, a silent dialogue between "two in one" (Arendt, 1978) in solitude, makes possible the understanding of experience through logos. What is fostered by judgment, on the other hand, is critical thinking in which communicability of thought is at stake. Without a public realm where thoughts can be expressed in the form of speech of an actor, the thinking mind eventually suffocates in destitute of common sense, let alone being critical in any sense. Arendt suggests that the way in which we can restore and preserve such a public realm is through our action and speech as one among equals. It is then up to us to rise to our occasion to exchange opinions and cultivate critical thinking on an equal footing, irrespective of social status or attributes of the participants so that the public realm can be felt to be our reality. This translates to constituting and sustaining the public realm through institutionalization where citizens can make substantive political decisions including the one on susteinametrics.

If the discipline of sustainametrics is to devise and propose measures for the sustainability of the world, then it must be conscious of the limit and prevent the reproduction of its thoughtless adoption by constantly exposing its outcomes to the scrutiny of political debate in the public realm. In this respect, the experts on the sustainametrics can inspire fellow citizens to join the discourse by taking the risk of acting and speaking in public as one of the fellow citizens, disclosing who they are and what it is really meant for us. The course of development must ultimately be grounded in such an act of disclosure, only through which we may find something worth sustaining in our future development, and sustainametrics is no exception.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

Conceptualization, formal analysis, and writing—original draft preparation: GO. Writing—review and editing and project administration: ST. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by JSPS KAKENHI Grant Number 21K00042.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arendt, H. (1978). *The Life of the Mind.* 2 vols. New York, NY: Harcourt.

Arendt, H. (1990). *On Revolution.* New York, NY: Penguin Books.

Arendt, H. (1998). *The Human Condition.* 2nd ed. Chicago, IL: The University of Chicago Press.

Arendt, H. (2003). *Responsibility and Judgment.* Ed J. Kohn. New York, NY: Schocken Books.

Arendt, H. (2015). *Vita Activa: Oder Von Tätigen Leben.* München: Piper Verlag GmbH.

Coyle, D. (2014). *GDP: a Brief but Affectionate History.* Princeton, NJ: Princeton University Press.

Greimas, A. J. (1969). *Dictionnaire de l'ancien français jusqu'au milieu du XIVe siècle*, Paris, Larousse.

Hartman, A. R., and Malicote, S. C. (2014). *Aiol: A chanson de Geste*. Italica Press.

Heidegger, M. (1967). *Sein und Zeit.* Tübingen: Max Niemeyer Verlag.

Heidegger, M. (1969). *Discourse on Thinking.* New York, NY: Harper Torchbooks.

Heidegger, M. (1977). *The Question Concerning Technology, Other Essays*. trans. with an intro. William Lovitt. New York, NY: Garland.

Heidegger, M. (1992). *Platon, Sophistes. GA 19*. Ed I. Schüssler. Frankfurt am Main: Klostermann.

Heidegger, M. (1997). *Plato's Sophist*. Trans. Richard Rojcewicz and Andre Schuwer. Bloomington: Indiana University Press.

Heidegger, M. (2000). *Vorträge und aufsätze, GA 7*. Ed F. W. Von Hermann. Frankfurt am Main: Klostermann.

Inwood, M. J. (1999). *A Heidegger Dictionary*. Malden, MA: Blackwell Publishers. doi: 10.1111/b.9780631190950.1999.00002.x

Kuznets, S. (1934). *National Income, 1929-1932. Senate Document No. 124, 73rd Congress, 2nd Session*. 5f. Available online at: https://www.nber.org/system/files/chapters/c2258/c2258.pdf (accessed October 10 2019).

Meitzen, A., and Falkner, R. P. (1891). History, theory, and technique of statistics. *Part first: History of statistics. Ann. Am. Acad. Polit. Soc. Sci.* 1, 1–100. Available online at: http://www.jstor.org/stable/1008943 (accessed September 16, 2022).

Normand, J., and Raynaud, G. (1877). *Aiol: Chanson de Geste*. Paris: Firmin Didot. p. 270.

Onions, C. T. (1994). *The Oxford Dictionary of English Etymology*. Hong Kong: Oxford University Press.

Sheehan, T. (1981). *Heidegger: The Man and the Thinker*. Chicago, IL: Precedent.

Tyrtaeus, S., and Theognis, M. (1999). *Greek Elegiac Poetry: From the Seventh to the Fifth Centuries BC.* Edited and translated by D. E. Gerber. Loeb Classical Library 258. Cambridge, MA: Harvard University Press.

frontiers | Frontiers in Sustainability

TYPE Original Research
PUBLISHED 23 November 2022
DOI 10.3389/frsus.2022.1072141

Check for updates

OPEN ACCESS

EDITED BY
Alexander Ryota Keeley,
Kyushu University, Japan

REVIEWED BY
Yuki Ishimoto,
Institute of Applied Energy, Japan
Yoshinari Suzuki,
United Nations Industrial Development Organization, Austria

*CORRESPONDENCE
Zhe Jin
jin.zhe.47c@st.kyoto-u.ac.jp

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 17 October 2022
ACCEPTED 01 November 2022
PUBLISHED 23 November 2022

CITATION
Jin Z and Ialnazov DS (2022) Sustainability assessment of individual-level solar energy poverty alleviation program-A case on Jinzhai County, China.
Front. Sustain. 3:1072141.
doi: 10.3389/frsus.2022.1072141

Sustainability assessment of individual-level solar energy poverty alleviation program-A case on Jinzhai County, China

Zhe Jin* and Dimiter S. Ialnazov

Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto, Japan

In this paper we study the Solar Energy for Poverty Alleviation Program (SEPAP) in China, which aims to increase the 3,000 Yuan annually for poor people by installing solar panels. SEPAP was initially launched in 2014 and officially ended in 2020 when President Xi announced that absolute poverty was eliminated in China. During the 6 years, China built and put into operation 26.49 million kilowatts of solar PV systems, benefiting 1,472 counties, 138,091 villages, and 4.18 million poor households. We propose a sustainable assessment framework and apply the Analytic hierarchy process (AHP) and Fuzzy comprehensive evaluation method (FCEM) to evaluate individual-level SEPAP in Jinzhai County, China, based on the findings of 80 semi-structured interviews with professionals and poor households. When examining SEPAP sustainability, we discover that the economic dimension is the most crucial one, with income, employment, training for the poor, and solar panel quality being the most weighted sub-indicators. In 2021, SEPAP could increase by roughly 2,700 Yuan for poor households, which is 90% achieved the governmental goals. We obtain a "Medium-high" outcome for the individual-level SEPAP. We provide two policy recommendations for maintenance work that will help the poor maintain a steady income.

KEYWORDS

Solar Energy for Poverty Alleviation Program (SEPAP), solar photovoltaic (PV) energy, extreme poverty, sustainability assessment, Jinzhai County, China

Introduction

There were 17 Sustainable Development Goals (SDGs) endorsed by all UN members in 2015 for global peace and prosperity. All nations, both developed and developing, are urged to act in response to achieving the 17 SDGs. For example, "Goal 1: No poverty" and "Goal 7: Clean and affordable energy" are the two goals tackling poverty and energy issues, respectively. The global poverty rate (at the US $1.90 poverty line) in 2018 was 8.6%, which was <9.1% in 2017 and is equivalent to a decline of 28 million poor people in 2 years (World Bank, 2022). The electrification rate in 2020 was 90.5%, which means there were still 770 million people living without access to electricity (World Bank, 2022). Unfortunately, the COVID-19 pandemic decelerated years of development and worsened poverty and energy access in undeveloped countries. According to the International Energy Agency (2022), after an average annual decrease of 9% in electrification in 2015–2019, there was a little change in the number of people without access to electricity in 2019–2021. Furthermore, the adverse effects of COVID-19 have downgraded poverty levels in some regions to approximately those recorded 30 years ago. The number of people living below poverty might have increased by 420–580 million in the worst-case scenario, with a 20% loss in income or consumption (Sumner et al., 2020).

China is one of the largest developing countries and had 82 million people living in extreme poverty in the year 2013 (National Bureau of Statistics of China, 2020). Despite China's recent declaration of 100% electrification access, 18.9% of Chinese could still be defined as "energy poor," who are in short of modern energy consumption, with most of them are concentrated in central and western China (Lin and Wang, 2020). Therefore, eradicating poverty and providing stable energy to the people in rural areas have been the priority tasks for the Chinese government. In 2014, China launched an ambitious poverty alleviation program (Solar-energy Poverty Alleviation Program, SEPAP) by implementing solar photovoltaic systems in remote rural areas. It aimed to increase energy capacity by more than 10 GW and generate annual income of ∼3,000 yuan for each poor household (National Development Reform Commissions, 2016). The support has been provided to over 2 million households in ∼35,000 villages across the nation. There were four primary choices: individual-level SEPAP, village-level SEPAP, joint-village-level SEPAP, and utility-scale SEPAP. In individual-level SEPAP, governments and photovoltaic companies assisted the poor in installing solar panels on their rooftops or lands. In village-level and joint-village-level SEPAPs, solar power plants were built in the vicinity of the counties or villages. In utility-scale SEPAP, centralized solar power plants were built in the neighborhoods of these counties or villages. Here, poor households and villages could either utilize or sell their generated electricity to the grid companies. According to the National Energy Administration (2020b), as of July 2020, China had built and put into operation 26.49 GW of photovoltaic power stations for poverty alleviation, benefiting 1,472 counties, 138,091 villages, and 4.18 million poor households, averaging more than 6 kW per poor household.

Because China has a large geographical area and large number of poor people, SEPAP has attracted great attention from academia. Zhang et al. (2020) use a panel dataset of 211 pilot counties between 2013 and 2016 to conduct a difference-in-difference (DID) regression and show that the SEPAP raises per-capita disposable income in a county by roughly 7–8%. Further DID regression analyzes by Liu et al. (2021) revealed that SEPAP has significantly improved the economic conditions and social capital[1] of low-income poor families, but that the expected gains in human and natural capital have not materialized. By adopting the Life Cycle Assessment (LCA) and Net Energy Analysis (NEA) methodologies, Wang et al. (2020) discovered that SEPAP had good energy efficiency and environmental advantages. Lo (2021) found that SEPAP has succeeded in achieving a just energy transition because of the just governmental procedures and the just outcome to the poor. There also have been several attempts to study sustainability perspectives and evaluate their impacts on the poor. Tao et al. (2022) evaluated the comprehensive benefits of SEPAP from the sustainability perspective. They identified an index system consisting of 13 sub-criteria derived from four aspects: economics, technology, society, and environment. The triangular intuitionistic fuzzy number and decision-making trial and evaluation laboratory were applied to calculate the weight of indicators, and the improved matter element extension method with the cloud model was applied to obtain final results. The results of the case study of four projects in Yunnan province demonstrated that the SEPAP has the potential for further development, and its overall benefits are generally at a good level. Huang et al. (2021) conducted a social impact evaluation of SEPAP, which was a multi-criteria analysis of four categories (human life, safety guarantee, and social resources), forming 13 indicators. They found that SEPAP improved poor families' economic conditions and social capital; however, the expected increases in human and natural capital were not observed. Wei (2021) constructed a 16-criteria system based on BOCR[2] theory from the perspectives of benefits, opportunities, costs, and risks and proposed a combined BOCR-AHP-IT2FTOPSIS method to analyze ambiguous data and perform a comprehensive sustainability assessment. In their empirical study in Guangxi Province, China, comparative and sensitivity analyses were used to rank the alternative SEPAP. According to the results of the criteria analysis, investment costs, poor collaboration, cleaner production, creation of jobs, and a reduction in the need for fuel are the major influencing factors for SEPAP sustainability.

On conducting a literature review, two research gaps were found. First, the institutional perspective is missing from the earlier literature on the SEPAP sustainability study. The implementation mechanism requires collaboration between the government and solar enterprises. However, no study has attempted quantification of the efforts made by the government and solar industry, incorporating both technological and institutional perspectives into the evaluation framework. Second, the literature often focuses on village-level SEPAP,

1 "Social capital" measures social attributes (e.g., social relations, social trust, formal and informal social networks), collective appeals, and opportunities to participate in decision-making.

2 BCOR stands for benefit, cost, opportunity and risk.

utility-scale SEPAP, or assessing the SEPAP in one region as a whole. As the four types of SEPAP are different in many ways, there is a lack of sustainability assessment on individual-level SEPAP. The main reason indeed is the lack of data. The majority of the individual-level solar modules have been installed in front of the poor people's home or on its rooftop. It is challenging to collect data because the majority of the poor live in remote areas that are difficult to access. Another reason could be that the Chinese government had designated village-level SEPAP as the main option between 2016 and 2020. However, given the technological advancements and falling cost of distributed solar panels, individual-level solar power plants are now considered as the new growth point for the solar expansion and the most suitable option for small investors in rural area as compared to large-scale solar power plants. In fact, the capacity of individual-level solar power plant has been increasing. The capacity of newly installed individual-level solar power plants in 2021 reached 21.6 GW (National Energy Administration, 2022), while the number in 2020 was 10.12 GW (National Energy Administration, 2020a). Further, 21.5 GW accounts for 41% of the total newly installed solar capacity and 75% of the newly installed distributed solar capacity (National Energy Administration, 2022). The reason is that most recently the Chinese government has been encouraging the installation of individual-level solar power plants, which is suitable for smaller investors. Although utility-scale solar power plants generally have a higher investment return than individual-level solar power plants, smaller investors may not be able to afford such a big investment. China being a vast country makes the evaluation of SEPAP implementation covering all the geographical areas difficult. This study chose Jinzhai County as the study area, because it is one of the earliest counties to implement SEPAP, and 7,803 individual-level SEPAP have been constructed since 2013. With its rich experience with SEPAP, it has also been promoted as the most successful model of SEPAP implementation in China. Before SEPAP initiation, Jinzhai County's poverty rate was 21.72%, but gradually, the poor people all escaped from poverty until 2020. This research intends to conduct a sustainability study on the factors affecting project sustainability and evaluate individual-level SEPAP in order to determine the most important factors/sustainability indicators for the sustainability of individual-level SEPAP and provide recommendations to policymakers regarding future poverty reduction policies using solar energy in rural areas. Figure 1 shows the research framework of sustainability assessment of individual-level SEPAP.

The remaining paper is divided into six sections. Section Sustainable framework and identification of sustainable indicators explains the sustainability framework and the selected sustainability indicators. Section Research methodology discusses the model and steps for analysis. Section Results shows the results of the indicator weight and SEPAP scores. Section Discussion discusses the results and provides policy recommendations for future poverty reduction policies using solar energy in rural areas. Section Conclusion presents the study's conclusions.

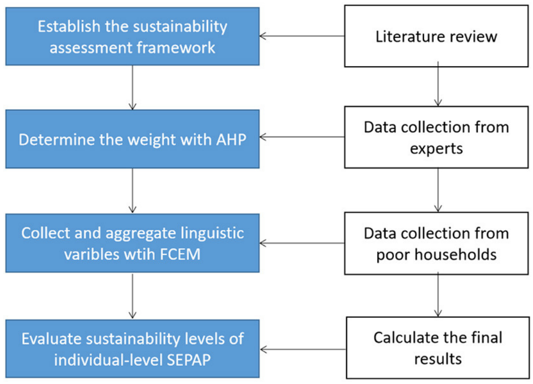


FIGURE 1
Research framework of sustainability assessment of individual-level SEPAP.

Sustainable framework and identification of sustainable indicators

As we entered the 21st century, the "sustainability" of human beings has become the paramount issue globally. With the launch of the SDGs and the effort to achieve the goals, researchers have been increasingly applying sustainability assessment frameworks to evaluate the sustainability in a certain area. The sustainability assessment framework creates a set of indicators that are connected to the SDGs, especially the economic, social, and environmental aspects. Typically, this approach has been used to evaluate electrification and poverty reduction programs in developing countries. In fact, the need to improve energy access continues to be a major motivator for reducing poverty in their rural areas (Thiam, 2011; Cheng et al., 2021). Modern energy's accessibility makes it easier to promote industrial development and raise human living standards.

The sustainability assessment framework suggested by Ilskog (2008) was based on 39 indicators. The proposed indicators cover the five dimensions of sustainability: technical, economic, social/ethical, environmental, and institutional sustainability. In the study by Ilskog and Kjellström (2008), fieldwork data from seven rural electrification projects in Kenya, Tanzania, and Zambia were presented, together with an explanation of how the chosen indicators and proposed framework could be applied to evaluate and compare different electrification programs. The authors suggested concentrating on providing rural electrification through small private and local community-based organizations, as they are the most effective in fostering sustainable development in their rural communities.

TABLE 1 Sustainability assessment framework (five dimensions and 17 indicators).

Economic	Social	Environmental	Technical	Institutional
A1. Affordability	B1. Smaller number of young people leaving the rural areas for the cities	C1. Replaced traditional energy (kerosene/wood)	D1. Quality of solar panel	E1. Information disclosure
A2. Income	B2. Education	C2. Reduced carbon emission	D2. Service availability	E2. Training provided by government to the poor households
A3. Employment	B3. Health		D3. Grid access improvement	E3. Accessibility of the local government and for issue reporting
	B4. Social Activities			E4. Trust between the poor people and the local government
				E5. Trust between the poor people and the maintenance company

Yadoo and Cruickshank (2012) applied the sustainability assessment framework to three case areas in Nepal, Peru, and Kenya to explore sustainable welfare benefits generated by renewable energy mini-grids. Due to the data availability and the unique local context, they adapted a framework with 44 set indicators in the five dimensions. To stimulate private sector investment, the authors proposed that policy efforts should concentrate on increasing public knowledge of renewable energy mini-grids, enhancing institutional, technical, and regulatory frameworks and establishing creative financing methods.

Boliko and Ialnazov (2019) compared four electricity projects in rural Kenya to ensure future sustainable development, following the methodologies employed by Ilskog (2008) and Yadoo and Cruickshank (2012). Their findings show that private sector-led off-grid solar electrification programs performed better than the others under evaluation, and hence, policymakers should continue to support these activities.

In our paper we use the sustainability assessment framework constructed by Ilskog (2008), Yadoo and Cruickshank (2012), and Boliko and Ialnazov (2019) with improvements to fit the local context in Jinzhai County, China. Our sustainability assessment framework (five dimensions and 17 indicators) is shown in Table 1. Below are the descriptions of the undertaken sustainability assessment framework.

The economic dimension assesses the economic impact of the project on poor households in Jinzhai County. Affordability is the capability of poor households to pay the cost of solar system. Income is the measure of whether SEPAP has raised the income of poor households. The employment indicator points at whether SEPAP has resulted in new employment opportunities.

The social dimension of the framework focuses on the impact that SEPAP has had on various aspects of the daily lives of the poor people living in Jinzhai County. Education and health evaluate its impacts on children's education and physical health, respectively, and social activities point at additional social activities that have been made possible as a result of SEPAP (such as increased social entertainment). The percentage of young individuals staying in rural areas reveals whether young poor people stayed or have returned to Jinzhai County in search of employment. The reason for adding this indicator is to check whether the introduction of solar energy has had an impact on the number of young people leaving the rural areas for the cities. If that number was smaller, then this result would help to prevent the further decline of the rural areas.

The environmental dimension focuses on how SEPAP has been able to lessen reliance on conventional energy sources, particularly the use of traditional biomass for energy production in rural areas, which often has negative effects on overall health, climate, and nearby natural environment. Kerosene for lighting and charcoal or firewood for cooking are examples of traditional biomass sources. Thus, this study examines whether SEPAP has reduced carbon emissions and substituted traditional biomass in terms of energy consumption (replacing kerosene, wood, and charcoal).

The technical dimension focuses on the technical performance of the SEPAP-related equipment and services. Therefore, the indicators that are included in the dimension are grid access improvement, service availability, and solar panel durability. The term "quality of solar panel" was measured by the time of professional maintenance work needed. Service availability measures the speed and quality of maintenance operations. The grid access improvement shows whether the grid service has been improved so that the poor can use stable electricity.

The institutional dimension examines how well the government and solar companies can support SEPAP's overall operation. Poor households' information disclosure reflects whether they could assess various SEPAP related information. Training for the poor households shows whether the government provided the necessary training to the poor households. Accessibility of the local government for issue reporting refers to the ease with which poor households (consumers) can report a particular issue on SEPAP. Trust between the poor and local government indicates the extent to which the poor could trust the local government during the entire program period. Similarly, trust between the poor and the maintenance company checks if they trust the maintenance company during the program period. The trust indicators are important because they demonstrate the degree of satisfaction of the poor people with the overall operation of the SEPAP.

Research methodology

After preparing the sustainability assessment framework for SEPAP, an appropriate method was required to weight the indicators and calculate the SEPAP scores. In this study, we used an analytic hierarchy process (AHP) to calculate the weight of each indicator. AHP not only enables multi-criteria comparison but also combines both quantitative and qualitative indicators together to rank both the indicators and alternatives. Then, the general fuzzy evaluation model was used to create a fuzzy mapping between each of the evaluation factors and transform a set of categorical appraisal grades into a numeral for final evaluation. Below are the descriptions of each method.

Designing the comparison matrix and obtaining the weight of each indicator

Upon setting the index system, the weights of indices denominating their relative importance are defined in Table 2.

According to the first-level evaluation indicator and the second-level evaluation indicator after classification, the following comparison matrix is constructed. The weight of the indicator was determined by calculating the specific feature vector through the comparison matrix:

$$P = \left[p_{ij}\right] = \begin{bmatrix} 1 & p_{12} & \cdots & p_{1n} \\ 1/p_{12} & 1 & \cdots & p_{2n} \\ \vdots & \ldots & \ddots & \vdots \\ 1/p_{1n} & 1/p_{2n} & \cdots & 1 \end{bmatrix}$$

TABLE 2 Analytic hierarchy process (AHP) matrix scale and definition.

Standard value	Definition
1	Equally important
3	Slightly important
5	More important
7	Obviously important
9	Absolutely important

Calculate consistency ratio

The priorities were derived as the normalized values of the right-hand eigenvector, which were associated with the largest eigenvalue (lmax) of the reciprocal matrix formed from the pairwise comparisons. The closer the lmax is to n (the number of elements being compared), the more consistent the judgments are. If judgments are perfectly consistent, lmax will be equal to n. Thus, the difference between lmax and n can be used as a measure of inconsistency. Instead of using this difference directly, a consistency index (CI) is shown as follows:

$$CI = \frac{(l\ \max - n)}{(n-1)}$$

Then, the consistency ratio (CR) can be calculated after obtaining the value of random index (RI). RI is the average random consistency index, which is listed in Table 3. If CR < 0.1, it can be considered that the judgment matrix has satisfactory consistency.

$$CR = \frac{CI}{RI}$$

Evaluating each indicator

The poor households and government officials (interviewees) were also given a qualitative indicator based on the 5-scale evaluation as {High, Medium high, Medium, Medium low, low}. Quantitative data of indicators was transformed to its actual number on the 5-scale evaluation, as shown in Table 4. Then, the probability of each choice for each indicator was mapped using a fuzzy matrix, R, such that if there are n factors and m levels of evaluation grades.

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}$$

Obtaining the final score of different types of SEPAP

The linear weighted sum function method was used to calculate the results of each stratified evaluation. The final indicator can be calculated as follows:

$$G = \sum_{1}^{N} w_{ij} r_{ij}$$

TABLE 3 Random consistency indicator RI coefficient table.

N	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.52	0.89	1.12	1.24	1.32	1.41	1.45	1.49	1.51

TABLE 4 Indicators using quantitative data and the transformation criteria to fuzzy evaluation.

Indicator	High	Medium high	Medium	Medium low	Low
Income (yuan/year)	≥3,000	(2,400, 3,000)	(1,800, 2400)	(1,000, 1,800)	(0, 1,000)
Reduced Carbon emission (kg/year)	≥2,430	(1,944, 2,430)	(1,458, 1,944)	(802, 1,458)	(0, 802)
Quality of solar panel (Times of maintenance needed)	0	1	2	3	≥4

where G represents the final score of the SEPAP, and w_{ij} and r_{ij} represent the weighting factor and evaluation result of each indicator, respectively.

Results

Field study

We conducted 80 semi-structured interviews using a questionnaire with 20 professionals, including professors, government officials, and project managers, who have rich experience in SEPAP, and 60 poor households, who have installed individual-level SEPAP in August–September 2022. To reduce in part the sample bias, we have interviewed poor households in three different villages in Jinzhai County. The data collected from the professionals were used for indicator weighting, and the household data was used for indicator evaluation.

The individual-level SEPAP was initially an experiment conducted by the Jinzhai government in 2013. The Jinzhai government helped eight poor households install solar photovoltaic (PV) in different areas of the county. All projects were completely funded by the local government, and each of them had a capacity of 3 kW. After 1 year, the government found that each poor household could earn an average of 3,000 yuan by selling the generated electricity to the grid company for 20 years. Therefore, in 2014–2015, large-scale installations were completed for 7,803 poor households. Similar to the experiment, each solar PV project had a capacity of 3 kW, and each individual-level SEPAP had a total investment of 24,000 Chinese yuan, of which the local government provided 8,000 yuan; another 8,000 yuan came from the donation by the solar PV company, and the remaining 8,000 yuan from the poor households. Because most of the poor households could not afford the initial investment of 8,000 yuan, the local banks provided an interest-free loan to them. The bank loans were agreed to be repaid by half of the annual income gained by selling the generated electricity until 2020. All the individual-level SEPAP were connected to the grid. The electricity generated was sold to the grid company at a selling price of 1 yuan/kWh along with the government subsidy on this price.

Indicator weighting

There were six comparison matrices determined based on the average weightage given by 20 researchers. Specifically, in this study, judgment matrices was constructed with a comparison matrix of primary indicators and five comparison matrices of sub-indicators. The values obtained from the judgment matrix are shown in Table 5, and the weighting results are shown in Table 6.

As seen from the Table 5, all CR values are lower than 0.1; therefore, they all passed the consistency test.

As can be seen from Figure 2 and Table 6, the economic dimension received the largest weight, making up more than half of all the weights. Therefore, it is clear that the SEPAP program has reduced poverty, whose sustainability was mostly determined by economic indicators. The institutional dimension came second after the economic dimension, demonstrating the importance of the roles played by the government and solar companies in the implementation of SEPAP. Additionally, the social and technological dimensions accounted for more than 10%. Lastly, the environmental indicators cannot be ignored.

Figure 3 shows the rank of normalized weights of each sub-indicator. It appears that higher income, which was also SEPAP's original goal, is the most crucial factor for benefiting poor households. Second, promoting employment is important, because it can provide the poor a reliable source of income. The training of the poor, quality of solar panels, education, trust

TABLE 5 Eigenvector, largest eigenvalue, consistency indicator, and consistency check coefficient comparison matrices.

	W	CI	RI	CR
Main criteria layer	(0.5161, 0.1353, 0.0532, 0.1117, 0.1882)	0.044	1.120	0.039
Economic	(0.1062, 0.6334, 0.2605)	0.029	0.520	0.037
Social	(0.1414, 0.4360, 0.3407, 0.0819)	0.064	0.890	0.072
Environmental	(0.7500, 0.2500)	0	0	Null[a]
Technological	(0.6434, 0.2828, 0.0738)	0.033	0.520	0.063
Institutional	(0.0722, 0.4489, 0.0722, 0.2428, 0.1640)	0.052	1.120	0.047

[a]Because the environmental dimension has only two indicators, it does not need to take a consistency test.

TABLE 6 Weighting factor of sustainability framework indicators.

Criteria layer	Weights	Sub-criteria layer	Weights	Normalized weights	Rank
Economic	0.5161	Affordability	0.1062	0.0548	6
		Income	0.6334	0.3269	1
		Employment	0.2605	0.1344	2
Social	0.1353	Smaller number of young people leaving the rural areas for the cities	0.1414	0.0191	12
		Education	0.4360	0.0590	5
		Health	0.3407	0.0461	7
		Social activities	0.0819	0.0111	16
Environmental	0.0532	Replaced traditional energy	0.7500	0.0399	9
		Reduced carbon emission	0.2500	0.0133	15
Technological	0.1117	The quality of solar panel	0.6434	0.0719	4
		Service availability	0.2828	0.0316	10
		Grid access improvement	0.0738	0.0082	17
Institutional	0.1882	Information disclosure (transparency)	0.0722	0.0136	13
		Training to the poor	0.4489	0.0845	3
		Issuing reporting	0.0722	0.0136	13
		Trust between the poor people and the local government	0.2428	0.0457	8
		Trust between the poor people and the maintenance company	0.1640	0.0309	11

between the local government and the poor, and affordability are also essential and account for more than 5% of the total weight.

Indicator evaluation

The data gained from the interviews with poor households was normalized, converted into the probability of each choice for each indicator, and then mapped using the fuzzy matrix R, as shown in Table 7. On the Left side, A1 to D5 are the codes of the indicators. On the first line, the classifications from "High" to "Low" are the evaluation levels. For instance, the value "0.55" in the second line indicates that 55% of the interviewees rated their income as "high," while "0" in the second line indicates that no one reported their income as "low".

$$G = \sum_{1}^{N} w_{ij} r_{ij} = \{0.250, 0.336, 0.194, 0.133,\ 0.086\}$$

Therefore, the final evaluation result of Individual-level SEPAP was "medium high."

Economic sustainability

Generally speaking, the majority of poor households were satisfied with the loan policy for the installation fees. This is due

TABLE 7 Evaluation results based on 60 interviews with individual-level Solar-energy Poverty Alleviation Program (SEPAP) households.

	High	Medium high	Medium	Medium low	Low
A1	0.55	0.27	0.13	0.05	0
A2	0.28	0.52	0.17	0.03	0
A3	0.10	0.07	0.18	0.40	0.25
B1	0.03	0.07	0.10	0.20	0.60
B2	0.15	0.35	0.22	0.12	0.17
B3	0.20	0.32	0.30	0.10	0.08
B4	0.32	0.27	0.25	0.13	0.03
C1	0.10	0.13	0.12	0.30	0.35
C2	0.28	0.52	0.17	0.03	0
D1	0.48	0.32	0.15	0.05	0
D2	0.30	0.33	0.18	0.13	0.05
D3	0.42	0.27	0.17	0.08	0.07
E1	0.25	0.45	0.25	0.03	0.02
E2	0.13	0.22	0.32	0.23	0.10
E3	0.18	0.37	0.30	0.12	0.03
E4	0.32	0.35	0.23	0.08	0.02
E5	0.25	0.32	0.23	0.15	0.05

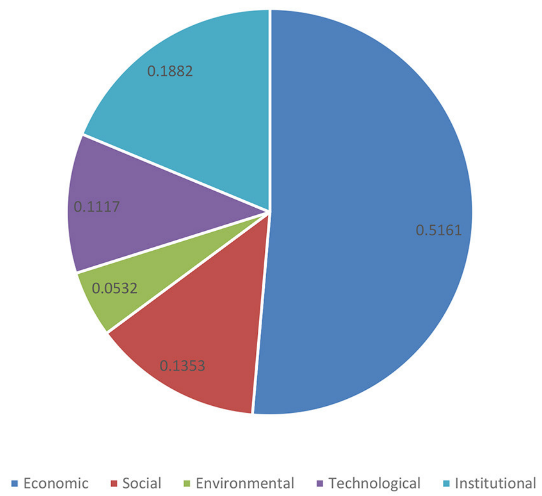


FIGURE 2
Weights of each main dimension.

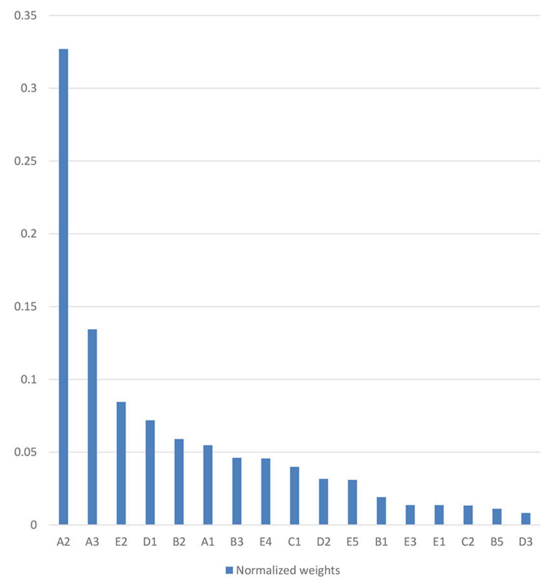


FIGURE 3
Normalized weights of each sub-indicator.

to the fact that, out of the entire 24,000 yuan installation fees, only a third was to be paid by them, and the rest was covered by the government and solar companies. For the poor households, local banks offered no-interest loans, which would be repaid by taking a deduction from SEPAP revenue each year until 2020. As a result, the poor households paid nothing for the installation. However, the poor households who installed SEPAP before mid-2014 were not eligible for the no-interest loan policy, as it was implemented after that. In fact, 8,000 yuan took up a marked percentage of their savings, and some of them borrowed money from their relatives.

In terms of increase in income, individual-level SEPAP has enabled poor households to gain ∼2,700 yuan by 2021, which was 90% of the government's goals (3,000 yuan). Less than 30% of interviewees reported having an income of >3,000 yuan, and over 50% of households received money ranging from 2,400 to

3,000 yuan. Additionally, the interview revealed that 20% of the respondents earn <2,400 yuan annually. Even worse, in 2021, two people earned an annual income from SEPAP ranging from 1,000 to 1,800 yuan. The primary cause was the malfunctioning of their solar panels for a while, which the poor households were unaware of.

The factor of increased employment had both direct and indirect impacts. In the direct impact, the villagers were employed as PV power plant maintenance personnel, and the new employment was produced by the PV company in the county. For example, one solar company built a factory to produce solar module components, and hired some local people. There were two types of indirect impacts: first, the money earned through the PV poverty alleviation program was used for skill development to increase steady employment with the acquired skills; second, there have been some attempts to combine solar energy with agriculture and fishing, which was a new "Solar+" business model that has the potential to expand employment. However, it was clear from the interview that both direct and indirect impacts were limited. Regarding the direct effects, each village only employs one maintenance worker for maintenance tasks, and PV companies prefer to hire employees with extensive experience for jobs involving PV rather than poor households, who would need expensive training. As for the indirect impacts, only 16% of interviewees mentioned that they used the income to learn more skills, and 21% of interviewees gained benefits from the "Solar+" industrial development.

Social sustainability

From the interview, it is hard to say that there was a trend of young people returning to the rural area for work. Most young people decide to stay in big cities. It must be acknowledged that there are more job opportunities in big cities; thus, the income there would be considerably higher than in rural areas. The poor might quickly escape poverty by working in large cities. Each province in China sets its own minimum monthly salary. According to the interview, poor households in Jinchai County were more likely to move to Shanghai for employment than any other large city, owing to the geographical closeness of the county to Shanghai. Shanghai's minimum monthly wage is 2,590 yuan (Shanghai Municipal Bureau of Human Resources Social Security, 2022); therefore, 2 months of work there would allow one to overcome poverty according to the Chinese poverty standard. As a result, the majority of young people do not choose to move back to the country. Despite difficulties for the majority of elderly individuals in commuting outside for work, some have stated that if they were younger, they would like to move out for work.

The effects of SEPAP on education varied widely among poor households. Some poor households believe that the effects are positive. Twenty five percent of poor households indicated that they could purchase stationery and reference books, and 30% of interviewees mentioned the increase in the nighttime study. However, some other poor households did not recognize the importance of SEPAP on education. Over 30% of households choose "medium-low" and "low" impacts for education. Below are some of their explanations:

Interviewee A: "The tuition and book fees were already exempted, and 9-year compulsory education was guaranteed by the government; I do not find any relation between SEPAP and education."
Interviewee B: "I have never paid any educational fees with money I received from SEPAP."
Interviewee C: "It is hard to say. Although we have benefited from SEPAP as our income increased, this is not related to education."

The health effects of SEPAP appeared to be more positive than those of education. In fact, the "351" policy, where the government guaranteed that an annual self-pay medical fee would not be more than 3,000, 5,000, and 10,000 yuan for medical treatment within the county, municipality, and provincial medical institutions in Anhui province, respectively, has actually helped in covering the majority of the medical costs. Approximately 20% of impoverished households indicated using the money they received from SEPAP to pay for the portion that the government was unable to cover. In addition, nearly 70% of poor households said that after the implementation of SEPAP, their quality of life has improved and they feel much healthier.

Interviewee A: "Before 2014, meat could only be eaten at festivals. Due to the increase in income after SEPAP installation, my family can eat more high-protein food."
Interviewee B: "My child was a bit malnourished and thin before, but since the government's poverty alleviation program came up, we can now afford more food and the child's health has improved."

The impact of increased social activities was also discovered during the interview. Poor households had more time to participate in daily social activities after the implementation of SEPAP than before it. The interviewee's village arranged more social activities, including watching movies and live sporting events, after the launch of the SEPAP program in 2014. The poor also have more time to play cards and chess together than they had before.

Environmental sustainability

There was no marked impact of SEPAP on poor households as an alternative to traditional energy sources. SEPAP connects individual-level solar PV to the grid, and the electricity generated is all sold to the grid company; hence, there is no noteworthy change in the way poor households use electricity. According to the interviews, most poor households maintain the habit of using wood for cooking and heating. Because most of the wood burned comes from the woods of the nearby mountains, it costs nothing. On the contrary, using electricity from the grid for cooking and heating would add an extra monthly expense. It should be noted that ∼10% of poor households have relocated with government assistance. They also changed their energy consumption and began utilizing more clean energy. Additionally, some poor households have increased their incomes because of SEPAP and other poverty alleviation programs. Consequently, they moved out of poverty and can now afford more electricity consumption from the grid.

As compared to replacing traditional energy, the impact on reducing CO_2 emissions was much better. The amount of CO_2 emissions was affected by the actual electricity generated. Yanzhe et al. (2021) estimated that coal-fired electricity generation emits 0.839 kg/kWh of CO_2, whereas solar energy generation emits 0.029 kg/kWh of CO_2. The government aimed to increase the poor's income by 3,000 yuan per year by installing 3 kW of solar power plants, which would produce 3,000 kWh of electricity and be sold at a price of 1 yuan/kWh. Thus, SEPAP would ideally cause an annual decrease of 2,430 kg in CO_2 emissions.

Technological sustainability

Most of the existing literature uses the decay rate as an indicator of solar technical performance (Tao et al., 2022). However, this study uses the actual times of maintenance work as an indicator to show the actual quality of solar panels. According to the questionnaire results, ∼50% of the respondents said that their solar panels are in good condition and they have never called a maintenance company for professional maintenance work. However, 30% of households asked about the maintenance work once, but they stated that most of their panels did not have a major problem. Approximately 10% of the solar panels were broken and required replacement. These solar panels were all constructed in 2013 and early 2014. Because there were no strict rules concerning the project entry requirements at that time, all solar companies were able to install solar panels for poor households. In order to lower the cost, some solar companies used low-quality solar panels, which broke afterwards. Even worse, many of these solar companies have gone bankrupt, which has caused big problems for maintenance work.

As for the speed and quality of the maintenance work, in general, poor households were satisfied. The interviewees claimed that the majority of issues are quickly solved. The technicians arrive on the second day after they call the maintenance company. Minor problems are solved on the same day, whereas some major problems that need replacement are solved within a week.

As individual-level SEPAP are connected to the grid, Jinzhai County has completed several rounds of grid upgrade work. Simultaneously, poor households have claimed that the electricity supply is more reliable now than it was before. Before 2014, frequent blackouts occurred during the summer and winter. However, after the grid upgrade, these situations improved. Nevertheless, some poor households asserted that a few extreme weather events had caused power outages.

Institutional sustainability

The respondents gave information disclosure a good rating. First, majority of the SEPAP related policy documents are available online. Poor households can also access the government website and enquire online regarding SEPAP. Second, the majority of the interviewees demonstrated that SEPAP was explained during the village residents' meeting at the beginning of the program. Additionally, it was disclosed who was eligible for SEPAP at that time. In addition, the list of poor households that joined SEPAP can also be found on the government website.

The interviews revealed that the training given to the poor was insufficient. Most of the poor said that there was no official training provided by the government, and therefore, they were unaware of the maintenance work. In addition, some of the poor households complained that they did not even realize the solar panel was broken for months. This caused lower revenue generation from SEPAP for the poor. In fact, we also interviewed government officials in Jinzhai County in 2021 and found that the frequency of SEPAP training for government officials varies among government agencies, with at least one time for the past SEPAP period. The majority of county-level agencies hold annual meetings to summarize the progress of the past year and make specific action plans. The most common pattern for government officials was to be trained when a new policy is initiated.

As for the accessibility to the government for issue reporting, most poor households responded positively. The interviewees generally consult with village-level government officials about SEPAP-related questions. However, it should be emphasized that the reaction takes considerably longer if the village-level government officials are unable to resolve the issue and have to refer it to the higher levels.

The program's sustainability was built on the trust of the poor. It was clear from the interview that people trusted the government and maintenance company both. However, trust in the maintenance company was weaker than that in the government for two primary reasons. First, despite not knowing what the solar panel would be used for, most households agreed to install it. Some admitted that they had a kind of blind trust in the government. They explained that "the government would not lie to us" and "the government serves the people." Second, the goal of an average increase in income by 3,000 yuan annually was achieved by some poor households as the government promised, which further enhanced their trust in the local government. However, if the electricity generation and revenue earned by SEPAP decrease, the trust in the government and the maintenance company will gradually erode.

Discussion

SEPAP was a government-led program with the primary purpose of increasing the income of poor households. The scheme enabled poor households to earn ∼3,000 yuan per year from the program owing to the government's substantial subsidy. Therefore, the sustainability of SEPAP is mostly dependent on the economic aspects, i.e., the actual income increased with the project, increased employment, and affordability. The technical and governmental aspects of the project are undoubtedly important, as they have a considerable impact on SEPAP's actual revenue, and it is also necessary to consider the social and environmental benefits of SEPAP, because the program cannot be sustained without these benefits.

Overall, it is difficult to say that SEPAP has succeeded in Jinzhai County, as 70% of the impoverished households were unable to boost their yearly income by 3,000 yuan. However, the income increase effect is still positive. According to the seventh national population census in Anhui Provincial Bureau of Statistics (2021), the average population per family household was 2.61 people. The poverty threshold was 2,800 yuan in 2014 (National Bureau of Statistics of China, 2018). We used this data to calculate the average income growth rate. If the loan was paid off with half of the SEPAP income and the remaining half was regarded as net revenue, the average income growth rate in 2014 was 18.47% (=2,700/2/2.61/2,800), which was a large increase for poor households. If thanks to SEPAP poor households' income increased each year, this could help them to cross the line and escape poverty. Nevertheless, it is important to remember that PV poverty alleviation has little impact on new employment. Additionally, the evaluation of both income and affordability would have deteriorated without the subsidies to the initial installation costs and electricity selling prices to the grid company. The social and environmental impacts were also somehow limited. For some indicators, such as improved education, improved health, increased social activities, and reduction of CO_2 emissions, the impacts were positive, but for others, such as the smaller number of young people leaving the rural areas for the cities and changing energy consumption structures, there has been no fundamental change so far. Technically speaking, most of the solar panels were of good quality, but low-quality PV exists, and it may have had an adverse effect on the sustainability of SEPAP. The stable grid can meet the electricity demand of poor households, and the quality and speed of maintenance work provided by the maintenance company are acceptable for poor households.

SEPAP have considerable short-term advantages, as it can immediately raise the income of poor households and allow them to escape poverty. Additionally, the 20 years of the program provided the poor with a stable income and simultaneously prevented them from slipping into poverty again. However, SEPAP was more alike to a direct income transfer program given the significant role that governmental funding plays in the program. Meanwhile, SEPAP cannot fundamentally solve the poverty problem from the root level because it does not increase the capability of most poor households to work. Hence, we argue that the combination of SEPAP and other poverty alleviation projects can achieve better poverty alleviation results. It was evident from the interview that the combination of SEPAP and 9-year compulsory education had covered most of the education fees of the poor children and increased their education, the combination of SEPAP and "351" medical insurance policy made medical treatments affordable for poor households, and the combination of SEPAP and relocation programs gradually changed the traditional energy consumption source of poor households.

However, we should also note that SEPAP is not suitable for every poor household. The government has already given some low-income households a living subsidy to ensure their basic needs before the launching of SEPAP, and joining SEPAP may lead these people to become lazier. They would believe that they could survive without work and would not look for a job, which equates to utilizing more government assistance to feed lazy people. Such poverty alleviation is definitely not an effective solution. Finally, based on the interview results, we have suggested the following two policy recommendations for future governmental actions on SEPAP maintenance.

Develop a remote monitoring system and help poor households to install this system on their cell phones

The importance of maintenance is self-evident, and understanding the power generation situation of PV at all times is of utmost importance. From the interviews, we noted that that most poor households have smartphones; therefore, they can install an online monitoring application. At the same time, there are already some applications of remote monitoring systems in other provinces and counties. Therefore, developing such a system is necessary and feasible, which enables the poor to access real-time power generation data and real-time income, where poor households can immediately report any noticeable abnormalities to the maintenance company for quick resolution of the problem.

Training for poor households on maintenance knowledge

From the interview, we noted that most poor households have received some kind of explanation about SEPAP, but they lack the knowledge of how to carry out maintenance work correctly. However, the amount of PV power output is being hampered by the weeds in front of their houses and the dust on the solar panels. Therefore, the government must provide training on maintenance work knowledge to poor households. It is suggested that training can be given to village workers first, and then, village cadres can guide poor households to carry out proper maintenance work, thereby lowering the amount lost because of improper maintenance labor.

Conclusion

By assessing individual-level SEPAP in Jinzhai County through a sustainability framework and an AHP-based fussy comprehensive method, this study contributes to a deeper understanding of the impacts and implementation status of SEPAP in rural China. Furthermore, this research is also important to future sustainable development, especially in the context of the transition from poverty alleviation to rural revitalization, as solar energy has been recognized as an important energy source in the process of rural revitalization (State Council, 2018).

However, our research has some limitations. First, the small sample of interviews of only individual-level SEPAP stakeholders limited our ability to compare different types of SEPAP in Jinzhai County. Second, the limited results of our fieldwork in Jinzhai County restrict our conclusions and policy recommendations to being applicable to other rural areas in China. To overcome these issues, more research should be conducted in the future with a larger sample size, on different types of SEPAP, and with field data from other counties as well.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

ZJ: conceptualization, methodology, data analysis, and writing original draft. DI: conceptualization, reviewing, and editing. All authors contributed to the article and approved the submitted version.

Funding

ZJ reports financial support was provided by Japan Science and Technology Agency. The Grant Number is JPMJSP2110.

Acknowledgments

The authors thank Dr. Elisabeth Ilskog and other co-researchers who contributed to create the sustainability framework. The authors thank all those who helped to conduct the interview in Jinzhai County. The authors also thank Ms. Nancy for the English proofreading.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anhui Provincial Bureau of Statistics (2021). *The Main Data of the Seventh National Population Census in Anhui Province* (in Chinese). Available online at: http://tjj.ah.gov.cn/ssah/qwfbjd/tjgb/sjtjgb/145782311.html (accessed September 21, 2022).

Boliko, C. M., and Ialnazov, D. S. (2019). An assessment of rural electrification projects in Kenya using a sustainability framework. *Energy Policy* 133, 110928. doi: 10.1016/j.enpol.2019.110928

Cheng, X., Chen, J., Jiang, S., Dai, Y., Zeng, J., Shuai, C., et al. (2021). Pursuing sustainable development goals: a review of renewable energy and poverty alleviation nexus. *Environ. Dev.* 40, 100679. doi: 10.1016/j.envdev.2021.100679

Huang, Y., Huang, B., Song, J., Xu, X., Chen, X., Zhang, Z., et al. (2021). Social impact assessment of photovoltaic poverty alleviation program in China. *J. Clean. Prod.* 290, 125208. doi: 10.1016/j.jclepro.2020.125208

Ilskog, E. (2008). Indicators for assessment of rural electrification—An approach for the comparison of apples and pears. *Energy Policy* 36, 2665–2673. doi: 10.1016/j.enpol.2008.03.023

Ilskog, E., and Kjellström, B. (2008). And then they lived sustainably ever after?—Assessment of rural electrification cases by means of indicators. *Energy Policy* 36, 2674–2684. doi: 10.1016/j.enpol.2008.03.022

International Energy Agency (2022). *Access to Electricity*. Available online at: https://www.iea.org/reports/sdg7-data-and-projections/access-to-electricity (accessed August 15, 2022).

Lin, B., and Wang, Y. (2020). Does energy poverty really exist in China? From the perspective of residential electricity consumption. *Energy Policy* 143, 111557. doi: 10.1016/j.enpol.2020.111557

Liu, J., Huang, F., Wang, Z., and Shuai, C. (2021). What is the anti-poverty effect of solar PV poverty alleviation projects? Evidence from rural China. *Energy* 218, 119498. doi: 10.1016/j.energy.2020.119498

Lo, K. (2021). Can authoritarian regimes achieve just energy transition? Evidence from China's solar photovoltaic poverty alleviation initiative. *Energy Res. Soc. Sci.* 82, 102315. doi: 10.1016/j.erss.2021.102315

National Bureau of Statistics of China (2018). *Rural Poverty Monitoring Report 2018* (in Chinese). Available online at: https://www.yearbookchina.com/navibooklist-n3019032313-1.html (accessed September 23, 2022).

National Bureau of Statistics of China (2020). *Poverty Monitoring Report in China 2020* (in Chinese). Available online at: https://www.yearbookchina.com/navibooklist-n3021031904-1.html (accessed September 23, 2022).

National Development and Reform Commissions (2016). *Opinions on the Implementation of SEPAP* (in Chinese). Available online at: https://www.ndrc.gov.cn/xxgk/zcfb/tz/201604/t20160401_963011.html?code=&state=123 (accessed August 10, 2022).

National Energy Administration (2020a). *Household Solar PV Project Information* (in Chinese). Available online at: http://www.nea.gov.cn/2020-12/16/c_139594051.htm (accessed October 28, 2022).

National Energy Administration (2020b). *Solar Energy for Poverty Alleviation: Construction Tasks Are Fully Completed, and Poverty Alleviation Methods Are Widely Popular* (in Chinese). Available online at: http://www.nea.gov.cn/2020-10/30/c_139478911.htm (accessed August 9, 2022).

National Energy Administration (2022). *Household Solar PV Project Information* (in Chinese). Available online at: http://www.nea.gov.cn/2022-01/19/c_1310431338.htm (accessed October 26, 2022).

Shanghai Municipal Bureau of Human Resources and Social Security (2022). *The City's Minimum Wage Will Not Be Adjusted in 2022, and the Monthly Minimum Wage Will Continue to be 2590 Yuan* (in Chinese). Available online at: https://m12333.cn/policy/kbcf.html (accessed September 26, 2022).

State Council (2018). *The Strategy of Rural Revitalization* (in Chinese). Available online at: http://www.gov.cn/zhengce/2018-09/26/content_5325534.htm (accessed September 25, 2022).

Sumner, A., Hoy, C., and Ortiz-Juarez, E. (2020). "Estimates of the impact of COVID-19 on global poverty (No. 2020/43)," in *WIDER Working Paper*. doi: 10.35188/UNU-WIDER/2020/800-9

Tao, Y., Wu, Y., Wu, M., Luo, X., He, F., Gao, R., et al. (2022). Multi-criteria decision making for comprehensive benefits assessment of photovoltaic poverty alleviation project under sustainability perspective: a case study in Yunnan, China. *J. Clean. Prod.* 346, 131175. doi: 10.1016/j.jclepro.2022.131175

Thiam, D. R. (2011). Renewable energy, poverty alleviation and developing nations: evidence from Senegal. *J. Energy South Afr.* 22, 23–34. doi: 10.17159/2413-3051/2011/v22i3a3219

Wang, C., Cheng, X., Shuai, C., Huang, F., Zhang, P., Zhou, M., et al. (2020). Evaluation of energy and environmental performances of solar photovoltaic-based targeted poverty alleviation plants in China. *Energy Sustain. Dev.* 56, 73–87. doi: 10.1016/j.esd.2020.04.003

Wei, Q. (2021). Sustainability evaluation of photovoltaic poverty alleviation projects using an integrated MCDM method: a case study in Guangxi, China. *J. Clean. Prod.* 302, 127040. doi: 10.1016/j.jclepro.2021.127040

World Bank (2022). *Access to Electricity (% of Population)*. Available online at: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS (accessed August 20, 2022).

Yadoo, A., and Cruickshank, H. (2012). The role for low carbon electrification technologies in poverty reduction and climate change strategies: a focus on renewable energy mini-grids with case studies in Nepal, Peru and Kenya. *Energy Policy* 42, 591–602 doi: 10.1016/j.enpol.2011.12.029

Yanzhe, W., Sheng, Z., Yu, W., Xuying, Q., Fubing, C., and Xunmin, O. (2021). Comprehensive assessment of the environmental impact of China's nuclear and other power generation technologies. *Tsinghua Sci. Technol.* 61, 377–384. doi: 10.16511/j.cnki.qhdxxb.2021.25.006

Zhang, H., Wu, K., Qiu, Y., Chan, G., Wang, S., Zhou, D., et al. (2020). Solar photovoltaic interventions have reduced rural poverty in China. *Nat. Commun.* 11, 1–10. doi: 10.1038/s41467-020-15826-4

TYPE Original Research
PUBLISHED 15 December 2022
DOI 10.3389/frsus.2022.910215

OPEN ACCESS

EDITED BY
Shutaro Takeda,
Kyoto University, Japan

REVIEWED BY
Mario Fargnoli,
Mercatorum University, Italy
Kei Saito,
Kyoto University, Japan

*CORRESPONDENCE
Eva Ponce-Cueto
eponce@mit.edu

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 29 April 2022
ACCEPTED 24 November 2022
PUBLISHED 15 December 2022

CITATION
Betts K, Gutierrez-Franco E and Ponce-Cueto E (2022) Key metrics to measure the performance and impact of reusable packaging in circular supply chains.
Front. Sustain. 3:910215.
doi: 10.3389/frsus.2022.910215

Key metrics to measure the performance and impact of reusable packaging in circular supply chains

Kellen Betts, Edgar Gutierrez-Franco and Eva Ponce-Cueto*

Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, MA, United States

Circular supply chains comprise the industrial production and supply chain systems used by companies to eliminate waste and recover value in products and materials. There are a variety of circular strategies including recycling in waste management, returns and repair in consumer-facing industries, and reusable packaging in supply chains. Successful implementation and management of these circular strategies requires the ability to measure and report on progress across different functions and processes. In this paper, we propose a new set of metrics to measure the performance of reusing items in a circular supply chain. We review the literature on metrics in circular supply chain management and reusable packaging in supply chains. We then classify the proposed metrics based on whether they measure the circularity of a circular supply chain initiative or impact of implementing the initiative. They also are segmented based on the level of detail they incorporate from the product level to the system level. We then demonstrate the use of the proposed metrics through a case study with an omnichannel retail company. We find that product-level metrics facilitate the comparison of different types of reusable and single-use packages with the potential to reduce lifecycle greenhouse gas emissions. We also find that measuring system-level Total Logistics Costs helps identify potential challenges with the feasibility of a reusable packaging system including the implications of not recovering packages and amortizing initial costs across multiple use cycles. Our aim in this paper is to address the gap in circularity and impact metrics focused on reuse strategies in supply chains. This new set of metrics provides companies with a tool to measure and report on progress toward a circular economy. It also suggests future avenues for research to assess the economic, environmental, and social dimensions of sustainability.

KEYWORDS

circular supply chain (CSC), reusable packaging, sustainability, performance indicator (PI), retailer

 10.3389/frsus.2022.910215

Introduction

Supply chains are the backbone of the modern economy. They also play a central role in the emission of greenhouse gases (Scott et al., 2018), depletion of critical material resources (Sovacool et al., 2020), accumulation of waste in the environment (Geyer et al., 2017; Tisserant et al., 2017), and other pressures on the natural environment (Hoekstra and Wiedmann, 2014). One solution that has been proposed to address these challenges is the concept of a circular economy (CE), which decouples economic growth from material depletion and waste generation by encouraging the reduction and reuse of products and materials (Stahel, 2016; Geissdoerfer et al., 2017). The concept of a CE was first introduced in Pearce and Turner (1990), and has gained increasing attention over the last two decades from governments, scholars, companies, and consumers (Ghisellini et al., 2016; Schöggl et al., 2020). CE is seen as a new business model that operationalizes sustainable development (Kirchherr et al., 2017). This requires balancing consideration of the economic, environmental, and social (EES) aspects of the economy, sector, or individual industrial process (Ghisellini et al., 2016). There are now national and international policies related to CE in China, the European Union, Norway, the United Kingdom, and other countries (Fan and Fang, 2020; Mhatre et al., 2021). Ellen MacArthur Foundation (2013), WBCSD (2021), WEF (2021), and other organizations have raised awareness of CE concepts at consumer, corporate, and policy levels. There is also a growing body of research that has been published on the topic of CE (Schöggl et al., 2020; De Pascale et al., 2021; Sarja et al., 2021).

By definition, the concept of a CE refers to the industrial production and supply chain systems that are designed to eliminate waste and recover value in products and materials (Batista et al., 2018). Once products are made, reverse logistics operations recover them from consumers and bring them back into the forward supply chain (Govindan and Soleimani, 2017). Closed-loop supply chains integrate both forward and reverse supply chains, but do not account for open-loop flows of by-products and wastes (Govindan and Soleimani, 2017). Circular supply chains (CSC) comprise both closed-loop and open-loop flows of products and materials in a CE (Batista et al., 2018; Geissdoerfer et al., 2018). In this paper, we embrace the definition of circular supply chain management (CSCM) proposed by (Farooque et al., 2019, p. 884): "Circular supply chain management is the integration of circular thinking into the management of the supply chain and its surrounding industrial and natural ecosystems. It systematically restores technical materials and regenerates biological materials toward a zero-waste vision through system-wide innovation in business models and supply chain functions from product/service design to end-of-life and waste management, involving all stakeholders in a product/service life cycle including parts/product manufacturers, service providers, consumers, and users."

CSCM is characterized by a variety of performance objectives including: minimizing waste, inventory, and the use of materials, water, and energy; and maximizing the availability of products, the number of recovery flows, and the efficient use of supply chain assets (Vegter et al., 2020). To achieve these objectives, nine circular strategies (called "R's" or "R-strategies") have been proposed Van Buren et al. (2016) including: Refuse (preventing raw material use), Reduce (increase efficiency in production or use), Reuse (secondhand or sharing), Repair (and maintenance), Refurbish (restoring an old product), Remanufacture (new product from old parts), Repurpose (reuse with different purpose), Recycle (reuse materials), and Recover (energy from incineration). A 10th R is sometimes added to the framework for Rethinking product use (Kirchherr et al., 2017). These strategies form closed loops for product and material flows in a CE with product life extension and reuse strategies referred to as "inner circles" or "inner loops," and end-of-life (EOL) management and recycling forming "outer circles" or "outer loops" (Ellen MacArthur Foundation, 2013).

Given the variety of CE strategies, efforts have been made to prioritize them based on application within the value chain and the impact on critical materials, waste generation, and other CE objectives (Ellen MacArthur Foundation, 2013; Kirchherr et al., 2017; Kalmykova et al., 2018). Shorter loop strategies focused on product use and life extension (e.g., reuse, repair, and refurbish) are considered to hold the greater potential to reduce environmental impact and create economic value than the outer loops focused on EOL management (e.g., remanufacturing and recycling; Ellen MacArthur Foundation, 2013; Stahel, 2013). And though many of the concepts of CE came from the waste management sector, there is a long history of sharing, reusing, and repairing products in the consumer segment as well as reusable packaging in supply chains (Hazen et al., 2021).

In the consumer segment, Kalmykova et al. (2018) found that short loop strategies are one of the most active areas of implementation. Out of the over 100 implementation cases they sampled across the full value chain from materials sourcing and product design to recycling and disposal, ∼10% of the cases were focused on reuse, sharing, and product-as-a-service or pay-per-use strategies. However, they found almost 50% of the cases were implemented in the collection, disposal, recovery, and recycling segments of the value chain. These outer loops are also where there is the most institutional support in China, the EU, and Japan, which can have the greatest impact on successful implementation of CE strategies in supply chains (Govindan and Hasanagic, 2018; Kalmykova et al., 2018).

In this research we focus on key metrics to measure the impact of reusable packaging in circular supply chains. Packaging plays an important role in supply chains, enabling the safe and efficient storage, handling, transportation, and sale of goods (Meherishi et al., 2019). And while reusable packaging is a well-known strategy, most packaging is designed to be disposed of after a single use (Escursell et al., 2021). As a result, packaging also uses large quantities of material resources and is a significant source of municipal and industrial solid wastes. In Europe, for example, packaging uses 40% of plastics and 50% of paper while contributing 36% of solid municipal waste (Coelho et al., 2020). Globally, plastic packaging accounts for roughly half of all plastic waste generated (Geyer et al., 2017). Reusable packaging, such as glass milk containers or fabric grocery bags, can replace single-use packaging throughout the supply chain resulting in lower life cycle greenhouse gas (GHG) emissions and reducing post-consumer waste (Goellner and Sparrow, 2014; Zimmermann and Bliklen, 2020; Fashion for Good, 2021). However, these benefits depend on the characteristics and performance of the reusable packaging system (Goudenege et al., 2013; Accorsi et al., 2014; Zimmermann and Bliklen, 2020).

Successful implementation and management of reusable packaging and other CE strategies in supply chains requires the ability to measure and report on progress across different functions and processes (Vegter et al., 2021). Metrics provide information for decision making, controlling the execution of strategy, and reporting to a variety of stakeholders (Neely et al., 1995; Maestrini et al., 2017). Supply chain metrics typically focus on economic performance (Maestrini et al., 2017). However, Dai and Tang (2021) argue that supply chain operations should be incorporated into environmental, social, and governance (ESG) measures, and ESG measures should play a central role in supply chain management (SCM) practices (Sarja et al., 2021). The literature on metrics in CSCM is sparse, but Vegter et al. (2021) found that performance metrics in CSCM expand on the economic focus in SCM by incorporating elements of circularity, environmental and, currently to a lesser degree, social considerations. Most of these metrics are in early stages of development (Acampora et al., 2017; Walker et al., 2018; Vegter et al., 2021), and despite the variety of R-strategies in a CE, product and enterprise-level metrics currently focus on recycling, remanufacturing, and other EOL strategies rather than reuse and other short-loop strategies (Kristensen and Mosgaard, 2020).

Reuse, and more specifically reusable packaging in supply chains, is a strategy to make progress toward reducing pressure on critical resources, reducing waste generation, and realizing the vision of a CE. In order to do so, key metrics are needed to manage the performance of a CSC with reusable packages or other items. However, the literature on metrics in CSCM is sparse (Vegter et al., 2021) and primarily focuses on recycling and other EOL management strategies rather than reuse and product life extension (Saidani et al., 2019; Kristensen and Mosgaard, 2020; De Pascale et al., 2021). Thus, the objectives of this paper are to:

- Review the literature on metrics in circular supply chain management and reusable packaging in supply chains.
- Propose a new set of metrics to measure the performance of reusing items in a circular supply chain.
- Demonstrate, through a case study, the use of these new metrics with an omnichannel retail company.

In the next section we review the literature on metrics used to measure circularity and EES impact in a CE and for CSCM. We also review the literature on metrics used to manage the performance of reusable packaging in supply chains. Building on this background, we then propose a new set of metrics to measure circularity and impact of reusable packaging in circular supply chains. We also demonstrate the use of these metrics with a case study and conclude with a discussion of what we found in this study and future research directions.

Literature review

There is a growing body of research on developing and testing metrics for supply chains in a CE. This paper will focus on research related to measuring the performance of reusing items in a circular supply chain. Google Scholar and MIT Library databases were used to find relevant publications including early works on topics related to circular economy and a focus on publications from the last 5 years. The primary search terms used were "circular economy," "circular supply chain," and "reusable packaging."

This paper will not provide a comprehensive or systematic review of this literature. Several existing literature reviews aim to do this and provide further guidance for this paper. Corona et al. (2019) reviewed 19 circularity metrics and assessment frameworks. They define criteria that a set of metrics should meet including the degree to which it measures what it intends to (validity), the consistency and robustness of the metric (reliability), and how practical it is (utility). The focus of their review is on the validity of metrics based on eight CE goals. They classify metrics into three groups: metrics measuring the degree of circularity ("circularity indices"), metrics that assess the effects of circularity ("CE assessment indicators"), and assessment frameworks such as Life Cycle Assessment (LCA) and Material Flow Analysis (MFA). For each category, they identified metrics focused on the product/enterprise level and/or at the sector/region/global level. They found there is a trade-off between scope (R-strategies) and practical usability with most indicators focused on a narrow aspect of CE. Assessment frameworks like LCA, in contrast, were more comprehensive in scope, but more complex to apply and interpret in practice. They also found that all 19 metrics accounted for environmental considerations, seven for economic considerations, and one for social considerations.

Saidani et al. (2019) conducted a similar review of 55 indicators across 10 dimensions including levels (micro, meso, and macro), loops (three of the R's), performance (circularity or impacts), perspective (actual or potential), uses, traversability (generic or sector-specific), dimension, units, format, and sources. They found the majority (>90%) of the indicators at all levels consider recycling while less than half consider reuse. Of the 20 indicators focused on the product or enterprise level (micro), a majority (80%) measure circularity while 40% measure EES impacts and 20% consider both. They also found that only three micro-level indicators were sector-specific. Previous work has found micro-level indicators are still in early stages of development with a low degree of adoption in industrial practices (Acampora et al., 2017; Walker et al., 2018), and Saidani et al. (2019) suggest that more work may be needed to advance the development of new indicators adapted to more specific contexts.

More recently, De Pascale et al. (2021) found similar results in their review of 61 CE indicators across levels, loops (six of the R's), sustainability impact (EES), and other dimensions. The majority (26 of 29) of micro-level indicators considered outer loop strategies (recycling and remanufacturing) while half take reuse strategies into account and none focused on reuse. They also found the majority (19 of 29) of the micro-level indicators focus on environmental and economic impacts, five focus on economic, and two on environmental. No micro-level indicators were focused on social considerations alone, although seven considered all three EES impacts.

Kristensen and Mosgaard (2020) found similar results in their review of thirty CE indicators focused at the micro-level. The majority of the indicators (19) include economic considerations, 12 directly and five indirectly include environmental considerations, and only four include social considerations. They also found that the majority of indicators (21) had a narrow focus on a single aspect of a CE, which may present an overly simplified measure of circularity and risks hindering sustainability more broadly by sub-optimizing. Conversely, they conclude that multidimensional indicators covering CE principals more broadly lack practical usability at the enterprise-level and prioritization of CE principals. While more narrow indicators may avoid these limitations, the indicators Kristensen and Mosgaard (2020) review all focused on recycling, remanufacturing, and other EOL strategies rather than reuse and product life extension strategies. However, reuse may present a more sustainable option than remanufacturing, which in turn may be more sustainable than recycling (Stahel, 2013), and so the lack of prioritization with multidimensional indicators, and the lack of focus on reuse and life extension strategies, misses the potential hierarchy for value creation and sustainability in a CE.

Vegter et al. (2021) conducted a review of 18 performance measurement systems, each consisting of a set of metrics, for CSCM. The authors propose different criteria than Corona et al. (2019) for an effective measurement system in CSCM including:

- Considering multiple dimensions of performance.
- The perspectives of multiple relevant stakeholders.
- The dimensions of circularity (reducing, maintaining, and recovering resources).
- The economic, environmental, and social dimensions of sustainable development.
- Limiting the range of measures to only the critical performance measures.
- Connecting performance on strategic, tactical, and operational levels (vertical integration).
- Aligning measures along processes (horizontal integration) including plan, source, make, deliver, use, return, recover, and enable.
- Recognizing and allowing for trade-offs and synergies with insight into the interdependencies among measures.

They find that social considerations are underrepresented and only two measurement systems provide insight into the interdependencies between circularity and sustainable development. They also find the literature on metrics in CSCM sparse with most in early stages of development that have not been tested in practice. And while the authors assess the level of integration horizontally, vertically, and across the forward and reverse chains, they do not assess the circular strategies (e.g., the 9 R's) each measurement system considers.

Few CE and CSCM metrics focus on the reuse CE strategy, however, reusable packaging is a well-known strategy in supply chains and there is a robust literature discussing packaging materials, reuse operations, and indicators used to manage operations with reusable packaging. Meherishi et al. (2019) conducted a broad review of packaging sustainability in SCM and CE, and found that reuse and return practices are one of the most popular topics studied (22 out of 59). For studies that looked at the EES impacts of packaging in supply chains (35), they found that environmental impacts (22) are studied more than twice as often as economic impacts (10) with only three considering social impacts. Mahmoudi and Parviziomran (2020) focused on reusable packaging and reviewed 86 studies on the environmental and economic impacts, system design, and operations management including performance measurement. They found reusable packaging has been studied in the food and beverage, cold-chain, and automotive industries as well as generic studies applied to any logistic system. Almost half (37) of the studies focused on performance measurements or criteria for measuring, and factors affecting, the environmental and economic impact of reusable packaging. And many of these propose new indices and metrics systems. They also found that more studies looking at environmental and economic impact were published in the last decade compared to decades prior, in contrast to the trend seen for operations management and logistics system design.

Building on these recent reviews, we focus our literature review on metrics in CSCM to identify patterns and common metrics that can be used for reusable items in supply chains. As shown by Corona et al. (2019), Vegter et al. (2021), metrics can be classified based on whether they measure circularity or the EES impact with some measurement systems combining both assessments. Since few CE and CSCM metric systems focus on reuse, we draw on other multidimensional indicators and indicators focused on other R strategies. Additionally, Ghisellini et al. (2016) and subsequent authors classify CE into micro, meso, and macro levels. The assessment of reusable items in supply chains is primarily focused at the micro level, so we use a new classification based on product or supply chain system level of analysis. This classification connects the metrics found in the literature on CE with the literature on SCM where SC problems are focused on the systems level (Choi et al., 2001; Chan and Chan, 2005). It also extends the definition of a supply chain proposed by Stevens (1989) as "a system whose constituent parts include material suppliers, production facilities, distribution services and customers linked together via the feed-forward flow of materials and the feedback flow of information" with the concept of a closed-loop supply chain integrating both forward and reverse flows (Govindan and Soleimani, 2017).

Measuring circularity and impact

Product-level metrics

One of the first circularity indicators that received attention in both practice and academia was the Material Circularity Indicator (MCI) developed by Ellen MacArthur Foundation and Granta Design (2015). The MCI is based on Material Flow Analysis (MFA; Ayres and Simonis, 1994) and focuses on quantifying a single indicator for the circularity of material flows based on a product's bill of materials. The metric is based on a linear flow factor calculated with the quantity (by weight) of materials and components that come from recycled or reused (remanufacture/repurpose loop) sources, and recycling efficiency. It also uses a utility factory based on an estimated lifespan. By focusing on material flows in a production process, the MCI focuses on outer-loop circularity with remanufacturing, repurposing, and recycling strategies. It measures circularity at the product level, which can be aggregated to the company-level based on a weighted average. The MCI focuses on a single circularity metric; however, it also provides complementary risk and impact indicators such as price variation, toxicity, scarcity, energy use, and CO_2 emissions. These complementary indicators allow for consideration of EES impacts.

Linder et al. (2017) argued that potential drawbacks with the MCI are the difficulty differentiating recovery pathways based on mass (e.g., recycled material vs. a remanufactured part) and estimating product lifespan, which often can't be measured easily. Instead, Linder et al. (2017) proposed a new product-level circularity (PLC) metric based on MFA for the relative embedded cost of recirculated parts and materials. Their rationale is that costs are more readily available and distinguishable by source. An economic value metric based on market prices can account for material scarcity as well. Linder et al. (2020) then evaluated the relationship between the PLC score and environmental impacts (measured using LCA) for 18 different products. They found a strong and significant negative correlation between product circularity and all three environmental impact measures (global warming potential, abiotic depletion potential, and environmental priority strategy), meaning that as product circularity increases, its environmental impact decreases.

Mesa et al. (2018) extended the product-level assessment to product families, a group of related products that are derived from a common set of components, with a set of six circularity metrics. In the product development process, the product family approach is a widely-used strategy to satisfy a variety of customer requirements with a more efficient manufacturing process. The six metrics they proposed measure material flows (by mass), component reusability (for remanufacturing/repurposing), reconfigurability, and functional performance. Similar to the MCI and PLC, Mesa et al. (2018) focused on outer-loop circularity. They also didn't assess the EES impacts of their circularity metrics, but they did evaluate the use of their metrics in the product development process with a case study (prosthetic fingers) and found circularity metrics can be used to optimize product design for circularity and functionality.

More recently, Bracquené et al. (2020) proposed a product circularity indicator (PCI) taking into account different recovery pathways using an MFA for mass at the material and component level rather than the product level like the MCI and PLC. They proposed a utility factor similar to MCI as well. By accounting for material losses in feedstock and component production, the PCI better reflects the benefits of using remanufactured parts compared to recycled materials. Bracquené et al. (2020) demonstrated the application of the PCI with a case study (washing machines) and a LCA to measure the environmental impacts of different circularity scenarios. They found that the PCI was an effective measure for primary material flows and impacts from the production phase. However, it was not able to effectively account for the production impacts of complex components like printed circuit boards nor impacts during the use phase, which are significant for an energy-intensive product like a washing machine. Similar to the MCI and other product circularity metrics, the PCI focuses on outer-loop circularity.

Product reuse is the shortest loop in the 9R framework involving the flow of products and materials and highlighted as a key strategy in CE (Korhonen et al., 2018). As mentioned above, however, few indicators focus on product reuse. One tool that attempts to address this discrepancy is the Circularity Calculator (CC) developed by IDEAL & CO (de Pauw et al., 2021). The CC provides four separate product-level metrics. Circularity (1) is measured using mass flows similar to the MCI. Value capture (2) assesses the economic impact of product design and production strategies based on the economic value of the materials from open, closed, remanufactured, and refurbished loops. Cycled content (3) assesses the percentage of a product's mass that comes from recycled and/or rapidly renewable resources. Finally, reuse is assessed based on the potential number of cycles a product can be used compared to a single-use product (one use cycle).

System-level metrics

Similar to product-level indicators, system-level indicators focus on outer loop circularity. Graedel et al. (2011) defined metrics that measure the circularity of metal production and scrap supply chains including the collection rate, process efficiency, recovery rate for scrap metal, and the recycled content and quantity of metal from EOL products. These metrics are used to estimate global recycling statistics, but would scale by redefining system boundaries to a more specific metal production, scrap, and recycling chain.

Brown and Bajada (2018) proposed two performance indicators for a recycling network that incorporate multiple stakeholder effects with multi-stakeholder collaboration including the impact of the speed of recycling and the effectiveness of collection and conversion of recycled material on production, and the upper bound of production for a given amount of recyclable material in the system. Based on a theoretical analysis, they found that cycle velocity, a factor not considered by other indicators, is positively related to production performance and recycling activity among stakeholders is driven by stakeholder collaboration and engagement.

Olugu et al. (2011) developed a set of sixteen circularity and EES performance measures for the automotive supply chain. While the lifespan of automobiles is frequently extended with reuse and repair strategies, Olugu et al. (2011) focused on the circular production supply chain which involves disassembly of EOL vehicles for remanufacturing or repurposing parts, then shredding the remaining vehicle for material recycling. They considered both the forward and backward chain. They then validated the set of metrics with thirty-three experts from academia and industry and found customer commitment to be the most important metric for the forward chain followed by quality, supplier commitment, and management commitment. Management was found to be the most important metric for the reverse chain followed by material features and recycling efficiency.

Some studies at the system-level focus on the EES impacts of a circular supply chain. Haghighi et al. (2016) proposed a balanced scorecard of indicators that measure EES performance in a recycling supply chain. They combined both quantitative and qualitative indicators in a multi-echelon network model, and tested it with data from 40 plastic recycling companies. They found that at the supplier echelon, the model is most sensitive to eliminating hazardous materials, flexibility was the most important indicator for the manufacturer echelon, and delivery cost was the most important factor for the distributor and retailer echelons. The model was almost equally sensitive to delivery time and customer satisfaction at the retailer echelon as well.

Ansari et al. (2020) focused on measuring EES impact as well. They identified 20 different indicators classified along the management processes of the supply chain operations reference (SCOR) model for a remanufacturing supply chain. They tested the complex relationships between indicators using a case study (remanufacturing business), and found that indicators for consumer awareness, technological compatibility, and workforce skill-level may be the most important in remanufacturing supply chains.

Measuring the performance of reusable packaging in supply chains

Out of all of the general CSCM metrics we reviewed, only the Circularity Calculator (de Pauw et al., 2021) included a direct measure of strategies that extend product life (maintenance, reuse, and repair). However, reusable packaging is a well-known strategy in supply chains. It is also one of the most popular topics studied related to packaging sustainability with many focused on performance measurements or criteria for measuring, and factors affecting, the environmental and economic impact of reusable packaging (Meherishi et al., 2019; Mahmoudi and Parviziomran, 2020).

Packaging can be classified based on its proximity to a product and functionality in the supply chain (Pålsson, 2018). Primary packaging (e.g., a shampoo bottle) surrounds a product, and provides convenience and protection, until the product is used by the end consumer. A secondary package (e.g., a corrugated box used for e-commerce shipping) is used to protect primary packages and bundle products into case quantities. Finally, tertiary packages (e.g., pallets) are used for bulk handling.

One of the first studies to look at the environmental impact of reuse and recycling in packaging systems was Tsiliyannis (2005). They developed a combined reuse/recycle model that measured the environmental performance of packaging systems based on virgin material demand and discarded waste. Using a theoretical analysis, they found that the total amount of packaging material that flows to consumers was a better indicator than the total number of reuse cycles, reuse rate, and recycling rate.

Goudenege et al. (2013) later developed a network model to compare the total costs and life cycle GHG emissions CO_2-equivalent or (CO_2e) for reusable (plastic cartons) and single-use (cardboard) secondary packages. The model was applied to a retail supply chain with forward and reverse flows between retail stores and distribution warehouses. They found that both total cost and life cycle emissions were lower for reusable packages if the backhaul between stores and warehouses is utilized with transportation costs negotiated lower than for the forward flow. If the backhaul is not utilized then costs and life cycle emissions are higher.

Goellner and Sparrow (2014), in contrast, found that environmental impacts are not as sensitive to backhaul transportation costs in cold chain logistics. They used a LCA to evaluate the global warming potential, acidification emission, eutrophication emissions, photochemical ozone emissions, human toxicity emissions, and post-consumer waste of single-use and reusable secondary packages for pharmaceutical and biological materials. Given the thermal control requirements of secondary packages in the cold chain, the impact of packaging manufacturing (cradle-to-gate) was significantly higher than the use and EOL phases.

Single-use packaging is common in the food industry as well including plastic bags and plastic or cardboard boxes that serve as primary packages as well as the cardboard boxes, wooden boxes, and disposal plastic crates used as secondary and tertiary packages. Accorsi et al. (2014) evaluated the economic and environmental impact of using a reusable plastic container (RPC) to carry fresh produce between farms and caterers in a food catering supply chain. They used LCA to measure the carbon footprint of packages and Life Cycle Costing (LCC) to measure infrastructure, storage, manufacturing, transportation, operating, and disposal costs. Compared to single-use packages, they found the RPCs would reduce life cycle GHG emission but increase overall costs. Within the supply chain, farmers would likely benefit from lower costs with distributors bearing the net cost increase.

Zimmermann and Bliklen (2020) conducted a similar study for secondary packages used in e-commerce shipping. Using a LCA of CO_2e emissions, they found that the majority of emissions for a reusable plastic bag and a reusable plastic box come from the initial production of the package followed by the forward transportation from retailer to customer. This life cycle footprint is then translated into the number of times a reusable package has to be used in order to achieve a lower footprint compared to the single-use package it would replace (the breakeven point). They found that the breakeven point for a reusable plastic box was 61 cycles compared to a standard cardboard box, 81 cycles compared to a cardboard box made from post-consumer fiber, and 32 cycles if the reusable box is made from post-consumer plastic. For a reusable plastic bag, the breakeven point was 20 cycles compared to a single-use plastic bag and only three cycles if the reusable bag was made from post-consumer plastic. This suggests that the number of reuse cycles, and factors that affect it such as the return rate by customers, has a significant impact on the environmental performance of reusable secondary packages in e-commerce.

Similar factors can affect the economic impact of reusable packaging as well. Mollenkopf et al. (2005) showed that reusable packaging is more economical for larger, bulk packaging with a high daily demand. Return rate, transportation costs, cycle times, delivery distance, fluctuation in maximum daily volume, and other factors can affect the economic impact as well (Mollenkopf et al., 2005; Breen, 2006; Cobb, 2016).

Proposed set of metrics

As discussed in the previous section, we found that few CE and CSCM metrics systems take into account reuse, repair, and other CE strategies that extend product lifespans. There is, however, a robust literature discussing strategies and operations including performance metrics for reusable packaging in supply chains. The focus of this section will be to connect these two bodies of work with a new set of key performance indicators for reusable packaging in circular supply chains.

In this paper, we embrace the definition of CSCM proposed by Farooque et al. (2019) and discussed earlier. That definition is grounded in the goal of systematically restoring technical materials and regenerating biological materials toward a zero-waste vision. Achieving this goal requires engaging all stakeholders from initial product or service design to EOL and waste management. A key activity in circular supply chains is the reverse logistics connecting forward chains with the reverse flow of products and materials. The metrics we propose are designed to help managers and organizations work toward the vision of a CE by measuring the circularity and EES impact of the supply chains they engage with.

Similar to our review of CE and CSCM metrics, we categorize this new set of metrics based on whether they measure the circularity of SC initiatives or the impact of implementing a circular SC initiative. Within each of these groups, a hierarchy starting with the aim of zero waste (level 1) then classifies metrics based on whether the analysis is focused on the product or supply chain system (level 2). Each of the product or system metrics (level 3) are then often a composite of more granular indicators (level 4) focused on components of the product or system. An illustration of this hierarchy can be seen in Figure 1. The hierarchy is meant to clarify the design intent for each of these metrics.

Measuring supply chain circularity

The first set of metrics are focused on measuring the circularity of SC initiatives. As seen in Table 1, we propose three product-level and two system-level composite metrics for the circularity of reusable items in SCM. Each of these metrics is meant to be measured for a fixed period. Previous MFA-based studies have used either mass

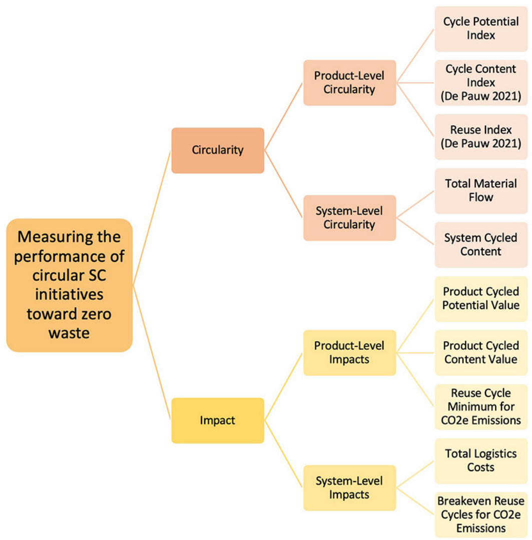


FIGURE 1
Proposed set of metrics to measure the circularity and impact of reusable packaging in a circular supply chain.

(Ellen MacArthur Foundation and Granta Design, 2015; Bracquené et al., 2020; de Pauw et al., 2021) or economic value (Linder et al., 2017; de Pauw et al., 2021) for the unit of measure for circularity metrics. There are advantages and limitations to both strategies. In this paper we will define metrics in such a way that either unit of measure could be used. This allows the user to select the unit that best suits their circumstances such as the availability of data. We refer the reader to Bracquené et al. (2020) for additional methodological details on mass-based MFA and Linder et al. (2017) for a cost-based approach.

Product-level metrics

Product Reuse Index

The Product Reuse Index (C_{RP}) is a metric designed to help supply chain managers compare single-use products with reusables whose life can be extended with maintenance, reuse, repair, and similar strategies. In contrast to the C_{PP} and C_{CP} which assess outer loop (EOL) strategies, the C_{RP} is focused on inner loop strategies. It was proposed by de Pauw et al. (2021), and given by,

$$C_{RP} = \left(1 - \frac{1}{N}\right) \times 100, \tag{1}$$

where N is the number of potential use cycles for a given product.

Product Cycle Potential Index

The Product Cycle Potential Index (C_{PP}) is a metric designed to help supply chain managers assess the potential of a product without knowing future material flows. It is forward-looking and assesses the degree to which components within a product can be put to use in future products as components, materials, or nutrients through remanufacturing, repurposing, recycling, composting, and other EOL strategies. Based on many similar circularity metrics proposed by Ellen MacArthur Foundation and Granta Design (2015), Bracquené et al. (2020), de Pauw et al. (2021) and others, the C_{PP} is defined as the ratio of the mass (or economic value) of materials or components within a product that are potentially reusable and the total product mass (or economic value) given by,

$$C_{PP} = \frac{M_{PUM} + M_{PRM} + M_{PCM}}{M_{TP}} \times 100, \tag{2}$$

where M_{PUM} is the mass (or economic value) of components that could be remanufactured or repurposed for future products, M_{PRM} is the mass (or economic value) of materials that could be recycled, M_{PCM} is the mass (or economic value) of materials that could be composted, and M_{TP} is the total mass (or economic value) of the product. Additional elements (M_i) could be considered if there are other material flow strategies within the CSCN.

There are three elements of the C_{PP} and the applicability of each depends on the characteristics of a particular product.

Percent of reusable (remanufacture/repurpose) components: C_{PUM} is the ratio of the mass (or economic value) of components within a product that are potentially reusable and the total product mass (or economic value). This metric assesses the degree to which components within a product can be put to use in future products through remanufacturing, in which case new products are made from components of the same product, or repurposing, where new products are made from components of other products. Considerations for the reusability of components include durability, disassemble-ability, and refurbish-ability. Closed-loop systems for reusable components require reverse flows, disassembly processes, and production processes for new products with inputs of remanufactured or repurposed components. This metric is given by,

$$C_{PUM} = \frac{M_{PUM}}{M_{TP}} \times 100. \tag{3}$$

Percent of recyclable materials: C_{PRM} is defined as the ratio of the mass (or economic value) of materials within a product that are potentially recyclable and the total product mass (or economic value). This metric assesses the degree to which the materials within a product can be put to use in future products by recycling those materials. Considerations for the recyclability of materials include accessibility of the material through efficient disassembly or material extraction processes and existing market capabilities to recycle the specific material.

TABLE 1 Proposed set of metrics to measure the circularity of supply chain initiatives.

Level 1 (goal)	Level 2 (dimensions)	Eq #	Variable	Level 3 (composite metrics)	Eq #	Variable	Level 4 (detailed metrics)
Toward zero waste	Product-Level Circularity	2	C_PP	Cycle Potential Index	3	C_PUM	% of reusable (remanufacture/repurpose) components
					4	C_PRM	% recyclable materials
					5	C_PCM	% of compostable materials
		6	C_CP	Cycle Content Index (de Pauw et al., 2021)	8	C_PRU	% recycled materials used in product
					7	C_PUU	% of used (remanufacture/repurpose) components in product
		1	C_RP	Reuse Index (de Pauw et al., 2021)	-	-	Number of potential use cycles
	System-Level Circularity			Total material flow	12	C_OUT	Percentage of material not recovered from end consumers (out of system)
					11	C_BACK	Percentage of material brought back to the system (material recirculation)
		11	C_BACK	System cycled content	14	C_RR	Percentage of recyclable material recovered
					13	C_RU	Percentage of components recovered that can be remanufactured or repurposed
					9 and 10	C_RS	System Reuse Index (number of use cycles for reusable, inner loop, products)

This metric is given by,

$$C_{PRM} = \frac{M_{PRM}}{M_{TP}} \times 100. \tag{4}$$

Percent of compostable materials: C_{PCM} is defined as the ratio of the mass (or economic value) of materials within a product that are potentially compostable and the total product mass (or economic value). This metric assesses the degree to which the materials within a product will biodegrade with residential or industrial composting processes. Considerations for the compostability of materials include the conditions under which the material biodegrades, accessibility of the material through efficient disassembly or material extraction processes, and existing market capabilities to compost the specific material. This metric is given by,

$$C_{PCM} = \frac{M_{PCM}}{M_{TP}} \times 100. \tag{5}$$

Combining these three elements gives,

$$C_{PUM} + C_{PRM} + C_{PCM} = \frac{M_{PUM} + M_{PRM} + M_{PCM}}{M_{TP}} \times 100 = C_{PP}$$

Product Cycled Content Index

The Product Cycled Content Index (C_{CP}) is a backward-looking metric designed to help supply chain managers assess the content used in a product that came from post-consumer, EOL sources. It is based on similar circularity metrics as the C_{PP}. The C_{CP} is defined as the ratio of the mass (or economic value) of materials or components within a product that came from post-consumer sources and the total product mass (or economic value) given by,

$$C_{CP} = \frac{M_{PUU} + M_{PRU}}{M_{TP}} \times 100, \quad (6)$$

where M_{PUU} is the mass (or economic value) of components from post-consumer remanufactured or repurposed sources, M_{PRU} is the mass (or economic value) from post-consumer recycled (PCR), and M_{TP} is the total mass (or economic value) of the product.

There are two elements of the C_{CP} and the applicability of each depends on the characteristics of a particular product. Similar to C_{PP}, additional elements (M_i) could be considered if there are other material flow strategies within the CSCN.

Percent of used (remanufacture/repurpose) components in a product: $C_{PUU}$is defined as the ratio of the mass (or economic value) of post-consumer components within a product and the total product mass (or economic value). This metric assesses the degree to which the components within a product came from post-consumer remanufactured or repurposed sources. This metric is given by,

$$C_{PUU} = \frac{M_{PUU}}{M_{TP}} \times 100. \quad (7)$$

Percent of recycled materials used in a product: C_{PRU} is defined as the ratio of the mass (or economic value) of PCR materials within a product and the total product mass (or economic value). This metric assesses the degree to which the materials within a product came from PCR sources. This metric is given by,

$$C_{PRU} = \frac{M_{PRU}}{M_{TP}} \times 100. \quad (8)$$

Combining these two elements gives,

$$C_{PUU} + C_{PRU} = \frac{M_{PUU} + M_{PRU}}{M_{TP}} \times 100 = C_{CP}.$$

System-level metrics

System Reuse Index

Similar to the product level, a system level metric focused on inner loop strategies can help supply chain actors compare supply chains with single-use products to reusables where processes bring products back to the system and extend product life with maintenance, reuse, repair, and similar strategies. The System Reuse Index (C_{RS}) is similar to the product-level metric (C_{RP}) and given by,

$$C_{RS} = \left(1 - \frac{1}{N}\right) \times 100, \quad (9)$$

where N is the average number of use cycles for the reusable products within the system. If the system only contains a single reusable product (e.g., a single type of reusable secondary package) then,

$$C_{RS} = C_{RP}.$$

In other cases, the system may be defined for multiple (p) products where n is the number of use cycles for a given (i) product and,

$$N = \frac{1}{p}\sum_{i=1}^{p} n_i. \quad (10)$$

Total material flow

There are two metrics that are designed to help supply chain actors assess the materials that circulate and are disposed of by a circular supply chain network (CSCN) and the materials that are not recovered by the system from end consumers.

Percent of material brought back to the system: C_{BACK} is defined as the ratio of the mass (or economic value) of materials that return back to the CSCN in which they were produced after at least one use cycle by an end consumer and the total mass (or economic value) of production within the system. The boundaries of the CSCN should be clearly defined to identify production output, and closed-loops which return products and materials back to the system. If the end consumer is considered within the boundary of the system then these materials circulate within the system without leaving. This metric is given by,

$$C_{BACK} = \frac{M_{BACK}}{M_{TS}} \times 100, \quad (11)$$

where M_{BACK} is the mass (or economic value) of materials that return back to the system after one or more uses by an end consumer or circulate within the system, and M_{TS} is the total mass (or economic value) of production within the system.

Percent of material not recovered from end consumers: C_{OUT} is defined as the ratio of the mass (or economic value) of materials not recovered from end consumers and the total mass (or economic value) of production within the system. Since these materials are not recovered by the system, their fate will not be known to actors within the CSCN, and so from a systems perspective these are materials lost to the system. Similar to C_{BACK}, the boundaries of the CSCN should be clearly defined. This metric is given by,

$$C_{OUT} = \frac{M_{OUT}}{M_{TS}} \times 100, \quad (12)$$

where M_{OUT} is the mass (or economic value) of materials not recovered from end consumers.

Combining these two elements gives,

$$M_{IN} + M_{BACK} + M_{OUT} = M_{TS}.$$

where M_{IN} is the mass (or economic value) of material inputs to the systems from non-recovered sources (e.g., virgin materials).

System cycled content

While the pathway of materials that are not recovered by the system (C_{OUT}) will not be known, the pathways for recovered materials (C_{BACK}) are important indicators of the circularity of material flows including recyclable materials and components that can be remanufactured or repurposed.

Percent of components recovered that can be remanufactured/repurposed: C_{RU} is defined as the ratio of the mass (or economic value) of components that return back to the CSCN which can be remanufactured or repurposed for future products and the total mass (or economic value) of production within the system. The system boundary considerations described for C_{BACK} should be considered for this metric as well. This metric is given by,

$$C_{RU} = \frac{M_{RU}}{M_{TS}} \times 100, \tag{13}$$

where M_{RU} is the mass (or economic value) of components that return back to the CSCN which can be remanufactured or repurposed for future products.

Percent of recyclable material recovered: C_{RR} is defined as the ratio of the mass (or economic value) of materials that return back to the CSCN which can be recycled within the system for future products and the total mass (or economic value) of production within the system. The system boundary considerations described for C_{BACK} should be considered for this metric as well. This metric is given by,

$$C_{RR} = \frac{M_{RR}}{M_{TS}} \times 100, \tag{14}$$

where M_{RR} is the mass (or economic value) of materials that return back to the CSCN which can be recycled within the system for future products.

Combining these two elements gives,

$$C_{RU} + C_{RR} = \frac{M_{RU} + M_{RR}}{M_{TS}} \times 100 = \frac{M_{BACK}}{M_{TS}} \times 100 = C_{BACK},$$

and

$$M_{RU} + M_{RR} = M_{BACK}.$$

Measuring the impact of reusable packaging

Given the complexity of EES systems for different products and CSC configurations, the metrics proposed for the assessment of EES impact are more specific to packaging in supply chains. As seen in Table 2, there are three product-level and two system-level composite metrics for measuring the impact of packaging in supply chains. Similar to the previous section on circularity, these metrics are defined with a hierarchy that may include more granular indicators for components of the product or system. These metrics also are based on existing research discussed previously on the performance indicators used to manage reusable packaging in supply chains and the use of LCA to assess environmental impacts. The metrics we propose are not meant to be used as an alternative to, rather as a complement to, LCA. In fact, LCA may be the best methodology to determine the CO_2e emissions associated with the metrics proposed below. We did not address the gap in literature on social dimension for reusable packaging and CE assessment more broadly. This is a limitation with this study and opportunity for future research. In the following subsections, we will define each of these metrics and highlight methodological and other considerations.

Product-level metrics

Reuse Cycle Minimum for CO_2e Emissions

At the product level, there are many approaches to assessing the environmental impact of packages in supply chains. The Reuse Cycle Minimum for CO_2e Emissions indicator $\left(n_{min}^{reuse}\right)$ is designed to help supply chain managers compare the embedded life cycle GHG emissions of reusable and single-use packages at the product level without knowing characteristics of the reusable packaging system. For a single use, CO_2e emissions from the upstream fabrication and manufacturing (F&M) life cycle of a reusable package is typically higher due to more durable construction than for a disposable, single-use package (Coelho et al., 2020; Mahmoudi and Parviziomran, 2020; Zimmermann and Bliklen, 2020). However, the F&M emissions for a reusable package can be amortized over multiple use cycles. Therefore, one way to compare the F&M emissions footprint of a reusable package and a single-use package is to identifying the minimum number of use cycles a specific reusable package product will have to be used before the amortized emissions for a reusable package are equal to or lower than for a single-use package (Zimmermann and Bliklen, 2020). This metric is given by,

$$n_{min}^{reuse} \geq \frac{e_{FM}^{reuse}}{e_{FM}^{single}}, \tag{15}$$

TABLE 2 Proposed set of metrics to measure the impact of implementing circular supply chain initiatives.

Level 1 (goal)	Level 2 (dimensions)	Eq #	Variable	Level 3 (composite metrics)	Eq #	Variable	Level 4 (detailed metrics)
Toward zero waste	Product-level impacts	16	V_PP	Product Cycled Potential Value	-	-	Cost of single-use packaging
					-	-	Cost of reusable packaging
					-	-	Cost of recyclable packaging
					-	-	Cost of compostable packaging
					-	-	Cost of remanufactured/repurposed packaging
					-	-	Cost of 100% virgin materials for packaging
					-	-	Cost of renewable materials
		17	V_CP	Product Cycled Content Value	-	V_PRU	Cost and LC emissions of recycled materials in packaging
					-	V_PUU	Cost of remanufactured/repurposed packaging
					15	n_min	Reuse Cycle Minimum for CO_2e Emissions (GHG emissions per use)
	System-level impacts	19	V_TL	Total Logistics Cost	-	-	Cost of transportation
					-	-	Cost of sorting
					-	-	Cost of cleaning
					-	-	Cost of inspection
		18	n_BE	Breakeven Reuse Cycles for CO_2e Emissions	-	-	GHG emissions generated when taking back returns (waste, recyclables, and reuse articles)
					-	-	Other life cycle GHG emissions

where n_{min}^{reuse} is the minimum number of use cycles for the reusable package, e_{FM}^{single} is the upstream F&M life cycle CO_2e emissions for the single-use package, and e_{FM}^{reuse} is the upstream F&M life cycle CO_2e emissions for the reusable package. This metric is related to the system-level Breakeven Reuse Cycles for CO2e Emissions indicator.

Product Cycled Potential Value

Similar to the C_{PP}, the Product Cycled Potential Value (V_{PP}) is a forward-looking metric designed to help supply chain managers assess the potential value of a product without knowing future material flows. If economic value is used as the unit of measure for C_{PP} then,

$$C_{PP} \equiv V_{PP},$$

otherwise V_{PP} is defined as the ratio of the economic value of materials or components within a product that are potentially reusable and the total product economic value given by,

$$V_{PP} = \frac{V_{PUM} + V_{PRM} + V_{PCM}}{V_{TP}} \times 100, \qquad (16)$$

where V_{PUM} is the economic value of components that could be remanufactured or repurposed for future products, V_{PRM} is the economic value of materials that could be recycled, V_{PCM} is the economic value of materials that could be composted, and V_{TP} is the total economic value of the product. Similar to C_{PP}, additional elements (V_i) could be considered if there are other material flow strategies within the CSCN.

Product Cycled Content Value

Similar to the C_{CP}, the Product Cycled Content Value (V_{CP}) is a backward-looking metric designed to help supply chain managers assess the value of content used in a product that came from post-consumer, EOL sources. If economic value is used as the unit of measure for C_{CP} then,

$$C_{CP} \equiv V_{CP},$$

otherwise V_{CP} is defined as the ratio of the economic value of materials or components within a product that came from post-consumer sources and the total product economic value given by,

$$V_{CP} = \frac{V_{PUU} + V_{PRU}}{V_{TP}} \times 100, \quad (17)$$

where V_{PUU} is the economic value of components from post-consumer remanufactured or repurposed sources, V_{PRU} is the economic value from PCR, and V_{TP} is the total economic value of the product. Similar to C_{CP}, additional elements (V_i) could be considered if there are other material flow strategies within the CSCN.

System-level metrics

Breakeven Reuse Cycles for CO_2e Emissions

Related to the product-level n_{min}^{reuse}, the Breakeven Reuse Cycles for CO_2e Emissions indicator $\left(n_{BE}^{reuse}\right)$ is designed to help supply chain managers compare the life cycle GHG emissions of reusable and single-use packages. This metric accounts for the full life cycle of packages in a reusable packaging system including F&M, forward and reverse transportation between the actors in the system, and processing packages between uses (Zimmermann and Bliklen, 2020). The life cycle emissions for a single-use package only includes F&M and forward transportation to the customer. This metric is given by,

$$\sum_{i=1}^{n_{BE}^{reuse}} e_i^{reuse} \leq \sum_{i=1}^{n_{BE}^{reuse}} e_i^{single} \quad (18)$$

which is evaluated iteratively and where n_{min}^{reuse} is the breakeven number of use cycles, e_i^{single} is the life cycle CO_2 equivalent (CO_2e) emissions for the single-use package at use cycle i, e_i^{reuse} is the life cycle CO_2e emissions for the reusable package at use cycle i.

Total Logistics Costs

Total Logistics Costs is a well-known metric designed to help supply chain managers assess the costs associated with the flow of products within a supply chain network. While accounting for materials considered flows outside of the system (e.g., M_{OUT}), cost accounting for reusable packaging is typically focused within the system (Mollenkopf et al., 2005). For a reusable packaging system, this metric is given by,

$$V_{TL} = \sum_{i=1}^{p} \left(V_i^{package} \times (n_i - 1)\right) + \sum_{i=1}^{d} \left(V_i^{deliv}\right) + V_{fixed} \quad (19)$$

where $V_i^{package}$ is the variable costs associated with each package (i) cycle (n_i) including sorting, inspecting, cleaning, storage, and any transportation and handling associated with managing the pool (p) of reusable packages in the system. V_i^{deliv} is the non-packaging variable costs associated with fulfilling and delivering a customer order including picking, packing, and transportation. Finally, V_{fixed} is the fixed costs associated with operating this system such as purchasing the reusable packages and the equipment and facilities needed to fulfill and deliver orders. Details for these costs will depend on the specific characteristics and configuration of a reusable packaging system.

Case study: Reusable packaging at an omnichannel retailer

To demonstrate the use of the proposed metrics for supply chain circularity and the impact of reusable packaging, we apply the metrics to a real case study. The purpose is to demonstrate the practical application of these metrics including testing different outcomes from decisions based on the metrics. This case study is based on an omnichannel retail company with hundreds of physical retail outlets and a growing fulfillment network for e-commerce. We interviewed subject matter experts at the company, reviewed internal strategy and operational documents, and analyzed internal data on sales, returns, packaging, products, and customers. The company name and related information will be anonymized and no proprietary information will be discussed in this paper.

The focus of this case is secondary packaging used for orders received online and either delivered to customer's homes by a local courier or picked up at stores by customers. Omnichannel retail refers to the sale of goods and services to end consumers through multiple sales channels that bridge digital and physical customer experiences. At the case company, online orders can be delivered to customer's homes or picked up at stores. Orders placed for same-day delivery are delivered by local couriers. Cardboard boxes are used to package products for these same-day deliveries to protect products during delivery. The same boxes can be used (at the customer's discretion) for orders picked up at stores. The company has ambitious goals to reduce the environmental impact of their supply chain including packaging. This case assesses the feasibility of replacing single-use cardboard boxes with reusable packages to work toward the company's goals. The two reusable packages that are considered in this assessment are a polypropylene box and a woven-polypropylene bag.

Operational considerations

Packaging is an important part of product value chains. It can be an integral feature of a product and protects products from damage, contamination, and other exposures prior to use. It also supports the efficient flow of products through supply chains from production to end consumers. The secondary packages used for local delivery and store pick up support the final leg of this journey. Cardboard boxes are assembled at stores and protect products in route to customer's homes. Once an order is delivered, how the cardboard box is disposed of is at the discretion of the customer. Replacing these single-use cardboard boxes with reusable packages has a number of important operational considerations. In order to close the loop, the packages need to be collected from customers and returned back to stores and other locations where they can be redeployed for future orders. This reverse flow includes transportation between customers and the company's facilities. It also includes processes to receive, inspect, repair (if needed), clean, and prepare the package for reuse. We will assess the circularity of this supply chain using the proposed metrics applicable to inner loop strategies including product-level and system-level indicators.

Measuring circularity

Selection of metrics for a specific application depends on the characteristics and configuration of the system as well as the availability of data. Data for total material flows were not available, but the focus for this case organization is a reusable packaging system and so all three product-level metrics can be used to support the assessment.

The Product Reuse Index (C_{RP}) assesses durability and other characteristics of a specific package product that influence the number of cycles it could be used for. Since a single reusable package is being considered for this system, the product-level and system-level reuse indexes are equivalent. The C_{RP} for a single-use cardboard box would be,

$$C_{RP}^{single} = \left(1 - \frac{1}{N}\right) \times 100 = \left(1 - \frac{1}{1}\right) \times 100 = 0.$$

In contrast, $C_{RP} = 99$ for a polypropylene box that could be used for 100 cycles and $C_{RP} = 95$ for a reusable woven-polypropylene bag that could be used for 20 cycles.

The Product Cycle Potential Index (C_{PP}) and Product Cycled Content Index (C_{CP}) can be used to inform selection of packaging type as well. There are many types of reusable packages used for e-commerce fulfillment including reusable plastic or fabric bags, semi-rigid totes, and plastic boxes (Coelho et al., 2020; Escursell et al., 2021). Specific package products can be evaluated based on the potential circularity for materials it is made with. For example, a 100-gram box made from polypropylene (90% post-consumer) could be recycled by the company. The Product Cycle Potential Index for this package would be,

$$C_{PP}^{A} = \frac{M_{PUM} + M_{PRM} + M_{PCM}}{M_{TP}} \times 100 = \frac{0 + 100 + 0}{100} \times 100 = 100.$$

Since the box is made from 90% post-consumer polypropylene, Product Cycled Content Index would be,

$$C_{CP}^{A} = \frac{M_{PUU} + M_{PRU}}{M_{TP}} \times 100 = \frac{0 + 90}{100} \times 100 = 90.$$

Similar calculations could be made for other types of reusable packages, including packaging with more complex construction mixing reused components and different recycled materials, providing standardized metrics to compare material circularity and support package selection by the company.

Measuring impact

The focus for this assessment is on the system-level economic and environment impacts of the reusable packaging system, but at the product level the Reuse Cycle Minimum for CO2e Emissions indicator can be used to assess the viability of a reusable package before data is available to evaluate the Reusable Index. In this case, the F&M emissions for a single-use cardboard box, polypropylene box, and woven-polypropylene bag are ~6 kg CO_2e, 0.3 kg CO_2e, and 0.4 kg CO_2e (Zimmermann and Bliklen, 2020). This means that the polypropylene box will have to be used for at least 15 use cycles before the amortized emissions are equal to or lower than for the single-use cardboard box $\left(n_{min}^{reuse}\right)$ while the woven-polypropylene bag has a F&M emissions. This minimum, however, does not take into account the full life cycle emissions of a reusable packaging system.

At the system level, we can use the Breakeven Reuse Cycles for CO_2e Emissions indicator to compare the full life cycle GHG emissions of reusable and single-use packages. This assessment takes into account the emissions from F&M as well as forward and reverse transportation between the actors in the system, and processing packages between uses. Life cycle emissions for a single-use package only includes F&M and forward transportation to the customer. For the logistics system considered by Zimmermann and Bliklen (2020), the breakeven

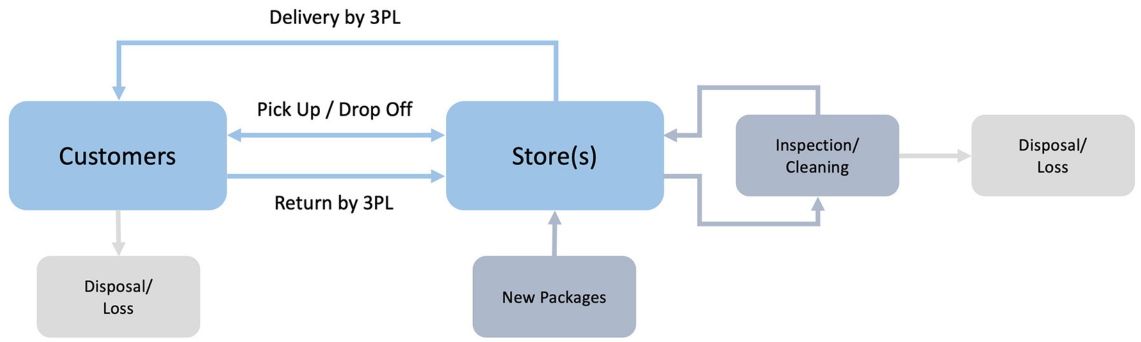


FIGURE 2
Flow of packages within this reusable packaging system.

TABLE 3 Parameters used for the simulation model.

Costs

- Cleaning: $0.5/package for each use cycle
- Handling: $0.2/package for each use cycle
- Storage: $0.1/package for each use cycle
- Transportation: $5/package for each extra trip
- Package: Scenarios with unit cost of $1, $2, $5, $10, etc. for each package acquired

Store conditions

- Simulation for 1 or 12 months of operations
- Demand of 5–30 packages/store/day (normal distribution)

Loss rates (package not reusable)

- 20% lost to customers or delivery service
- 10% damage and other losses

Reverse flow rates (bringing packages back)

- 90% either customer drop off or delivery backhaul (not an extra trip, $0/package)
- 10% to packages need to be brought back with an extra trip at $5/package

Cycle times

- Returned in 8 days on average
- Cleaned in 3 days on average

TABLE 4 Total Logistics Costs for each scenario.

Scenarios	$1 package	$5 package	$10 package
1 store, 1 month			
Variable costs	$0.4 K	$0.4 K	$0.4 K
Fixed costs	$0.3 K	$1.6 K	$3.2 K
Total costs	$0.7 K	$2.0 K	$3.6 K
Total cost per order	$1.22	$3.48	$6.30
1 store, 12 months			
Variable costs	$5.4 K	$5.4 K	$5.4 K
Fixed costs	$1.6 K	$8.2	$16.3 K
Total costs	$7.0 K	$13.6 K	$21.7 K
Variable cost per order	$0.87	$0.87	$0.87
Total cost per order	$1.13	$2.17	$3.47
400 stores, 12 months			
Variable costs	$2.2 M	$2.2 M	$2.2 M
Fixed costs	$640 K	$3.2 M	$6.5 M
Total costs	$2.8 M	$5.4 M	$8.7 M

point for a polypropylene box is 61 cycles compared to a standard cardboard box, 81 cycles compared to a cardboard box made from post-consumer fiber, and 32 cycles if the reusable box is made from post-consumer plastic. For the woven-polypropylene bag, the breakeven point is only two cycles and one cycle if the reusable bag is made from post-consumer polypropylene. Since the polypropylene box and woven-polypropylene bag can be used for up to 100 and 20 use cycles, the cumulative CO_2e emissions for the logistics system would be lower with either of these reusable packages than for single-use cardboard boxes.

In addition to the environmental impacts, we can use the Total Logistics Costs indicator to assess the system-level economic impacts. Cost details depend on the specific characteristics and configuration of the logistics system. In this case (Figure 2), retail stores serve as e-commerce fulfillment locations where orders are picked, packed, and staged for delivery. Orders are then picked up by customers at the store or delivered to customer's homes by a local courier. When customers pick up an order and/or shop in the store they can drop off any reusable packages they have from previous orders. Couriers can bring back reusable packages to the store from prior orders when they deliver new orders as well. At the store,

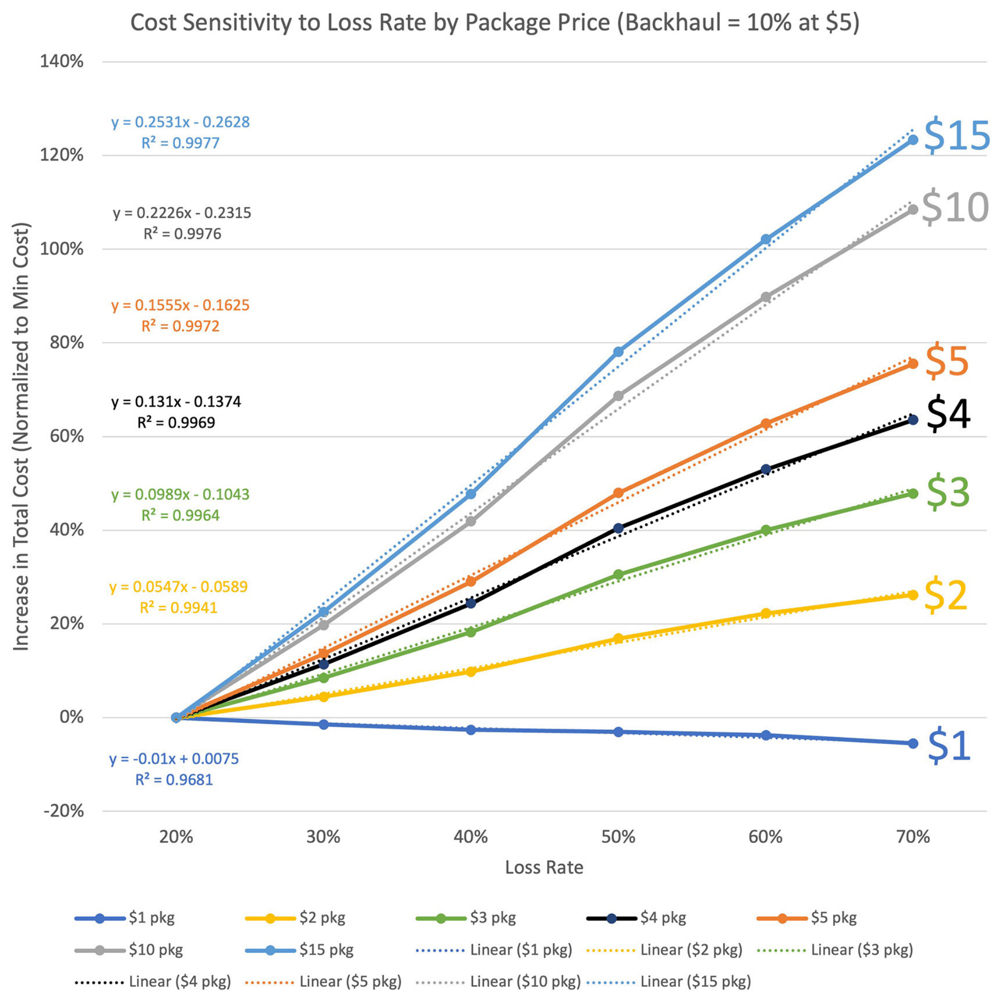


FIGURE 3
Cost sensitivity to loss rate for packages priced (color) between $1 and $15, and for 10% of packages subject to an extra backhaul cost of $5, showing total cost is sensitive to higher priced packages at a higher rate (slope of the line).

reusable packages are sorted and inspected. A separate cleaning service is used.

To evaluate the Total Logistics Costs, the variable costs for each reusable package use cycle include cleaning, handling, storage, and replacing any reusable packages that are damaged or lost. Transportation is the other variable cost associated with each order. Approximately 90% of retail sales for this case organization occur at physical retail stores, and so we assume that the local courier will incur costs to return packages to the store for only 10% of orders. Finally, since the case organization built and operates retail stores independently of the e-commerce orders delivered within this reusable packaging system, the only fixed cost is purchasing the inventory of reusable packages needed to meet demand for online orders fulfilled at stores. The full set of parameters used for the simulation model can be found in Table 3.

We calculated the Total Logistics Costs for nine different scenarios including: one store for 1 month of operations, one store for 12 months, and 400 stores for 12 months, each for three different reusable package purchase prices. As seen in Table 4, the Total Logistics Costs for operating a circular supply chain with reusable packages could range from $2.8 million to $8.7 million in 400 stores over 12 months. The total cost per order is between $1.10 and $3.50, depending on the type of package used.

To determine the most important factors influencing the Total Logistics Costs of this reusable packaging system, we evaluated the sensitivity to several key parameters used in the model including the loss rate (reusable packages not recovered from customers), reusable package purchase price, and the cost and rate for transportation from customer's homes back to retail stores (backhaul).

As seen in Figure 3, we found that the total cost increases as the loss rate increases indicating total cost is sensitive to loss rate. Moreover, the sensitivity rate (indicated by the slope of the line) increases as the price of the package increases indicating that total cost is sensitive to both the loss rate and package price. For $1 packages, the total cost is insensitive or possibly inversely sensitive to the loss rate. For packages costing $2 or more, the sensitivity rate increases with the price of the package. However, while the cost sensitivity to the loss rate increases with package price, the slope of the sensitivity curves tapers as package price increases (Figure 4). This tapering suggests that other parameters are buffering cost sensitivity as package price increases.

Two other key parameters (Table 3) used in the model are the extra backhaul cost and the number of packages subject to the extra backhaul cost (backhaul rate). As seen in Figure 5, the sensitivity pattern for backhaul costs of $5 and $10 are similar. Packages >$2 are sensitive to the loss rate, and the sensitivity rate increases with package price. However, when we increase the number of packages subject to the backhaul cost from 10% to 50%, the additional backhaul costs buffer sensitivity to loss rate. As seen in Figure 6, the higher extra backhaul rate results in packages almost as much as $5 being insensitive to the loss rate—in which case it would be cheaper to accept the loss and purchase a new package.

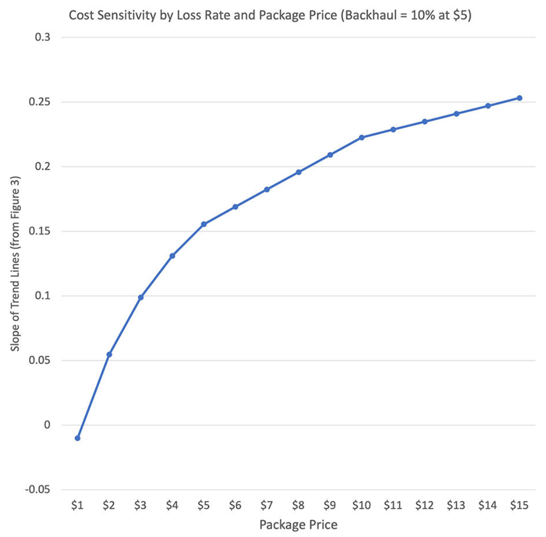


FIGURE 4
The trend in cost sensitivity rate to loss rate by package price showing that the rate of cost sensitivity decreases as package price increases, and suggesting other parameters influence cost sensitivity.

Discussion and conclusion

In this paper, we propose a new set of metrics to measure the circularity and impact of reusable packages in supply chains. Packaging plays an important role in supply chains, protecting products from production to end consumers. There is also a growing awareness that packaging uses material resources and results in waste that enters our environment. Reusable packaging is one solution that may address these challenges, and successful implementation and management of this and other CE strategies requires the ability to measure and report on progress across different functions and processes.

There is a growing body of research on metrics for circularity and the impact of CE strategies and CSCM. We found these studies focus on outer-loop circular strategies including remanufacturing and recycling with less attention to reuse strategies. At the same time, reusable packaging is a well-known strategy in supply chains. The contribution in this paper is a new set of metrics to measure circularity and the economic and environmental impacts of reusable items by connecting the existing research on CE metrics with reuse strategies in supply chains. We categorize these metrics as product-level or system-level based on the level of detail they incorporate. We also demonstrate their application with a case study from an omnichannel retail company.

With the case study, we found that the Product Reuse Index, Reuse Cycle Minimum for CO2e Emissions, and Breakeven Reuse Cycles for CO_2e Emissions indicators facilitate the comparison of different types of reusable and single-use packages. They show, for example, that a reusable system with either a polypropylene box and reusable woven-polypropylene bag would result in lower cumulative life cycle emissions than shipping with single-use cardboard boxes. The reusable box, however, would have to be used between 32 and 81 times before cumulative emissions are an improvement over cardboard. This has important implications for the durability of the box and recovery from customers. Some studies have found that recovery of reusable packages in a business-to-consumer system can be low (Accorsi et al., 2014; Zimmermann and Bliklen, 2020).

We also found that Total Logistics Costs is sensitive to the rate at which reusable packages are lost, and the impact grows with package cost. Packages can be recovered by asking customers to return them to physical retail locations. They also can be recovered using an owned or third-party transportation service such as local couriers. However, use of transportation services comes with a cost and we found Total Logistics Costs is more sensitive to the number of packages that have to be brought back rather than the unit transportation cost. This suggests that as the rate of package recovery using a transportation service increase, it may be more cost effective to accept a higher loss rate—recovery fewer packages—for higher priced packages, which would then reduce the environmental benefits of reusable packaging.

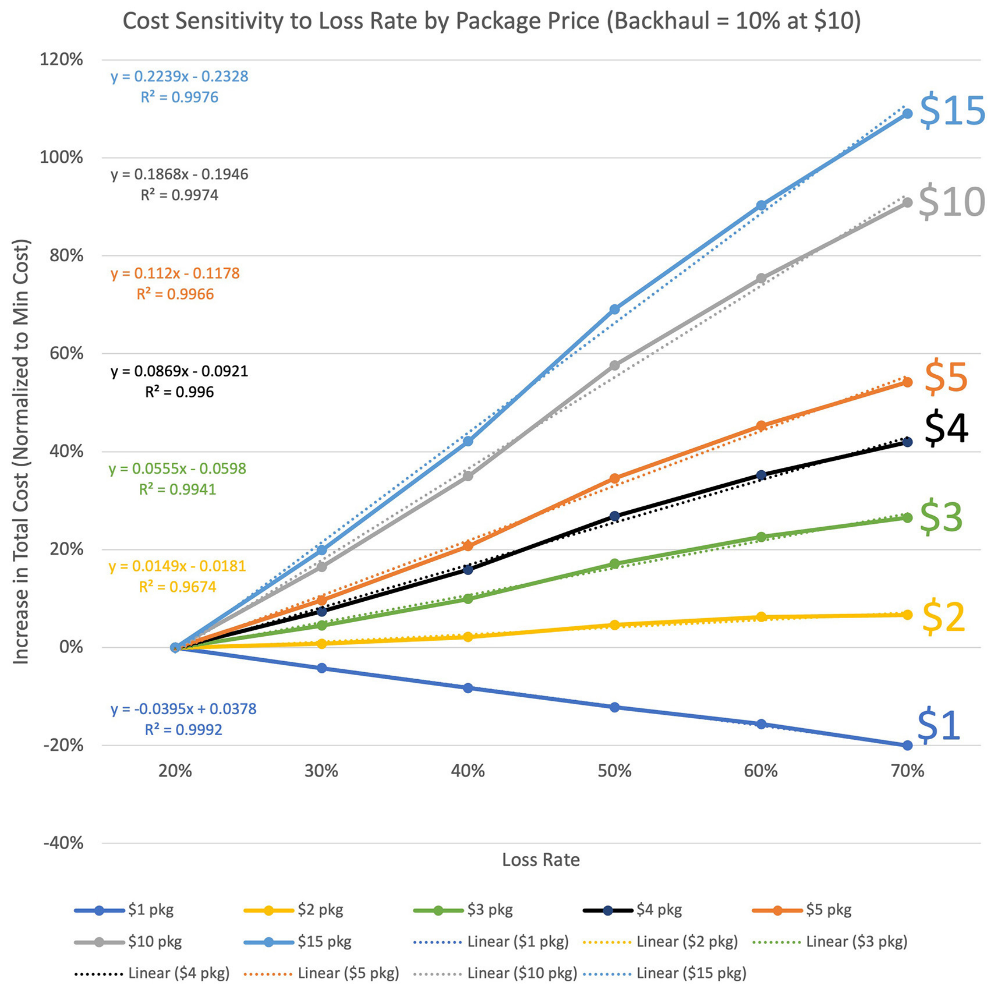


FIGURE 5
Cost sensitivity to loss rate for packages priced (color) between $1 and $15, and for 10% of packages subject to an extra backhaul trip cost of $10. Compared to Figure 3 ($5 backhaul trip cost) the sensitivity pattern did not change suggesting total cost is not sensitive to the backhaul trip cost when only 10% of packages are subject to the cost.

This case study demonstrates that the new set of metrics we propose for reusable items in SCM shed light on key financial and operational considerations for reusable packaging in an omnichannel retail environment. The same methodology could be applied to a wide range of business models considering SC initiatives for reusable items. Performance metrics inherently rely on the use of data, however, and the primary limitation that may be encountered is the availability of information about specific products and business operations. This limitation can be mitigated in some cases through the use of more general industry data, which may be more readily available in some cases while also reducing the specificity of the results.

While our aim was to address the gap in circularity and impact metrics focused on reuse strategies in supply chains, there are limitations to the new metrics we propose and the case study we used that provide opportunities for further research. The literature and our metrics are focused on economic and/or environmental impacts with limited consideration of the social dimension of sustainability. Further research could evaluate potential social considerations of a reusable packaging system including upstream impacts during the F&M of packages, the impacts on retail store employees and local delivery drivers, and the customer experience.

Another avenue for further research could be implementation and empirical testing of the metrics we propose. Vegter et al. (2021) found that the majority of metrics systems for CSCM are still in early phases of development with only 20% in implementation and 10% in use. We evaluated the application of the metrics within the constraints of a real organization, and found they provide insight into the potential impact of package recovery and other considerations. Further research could evaluate the implementation and effectiveness of these metrics for managing an operational reusable packaging system.

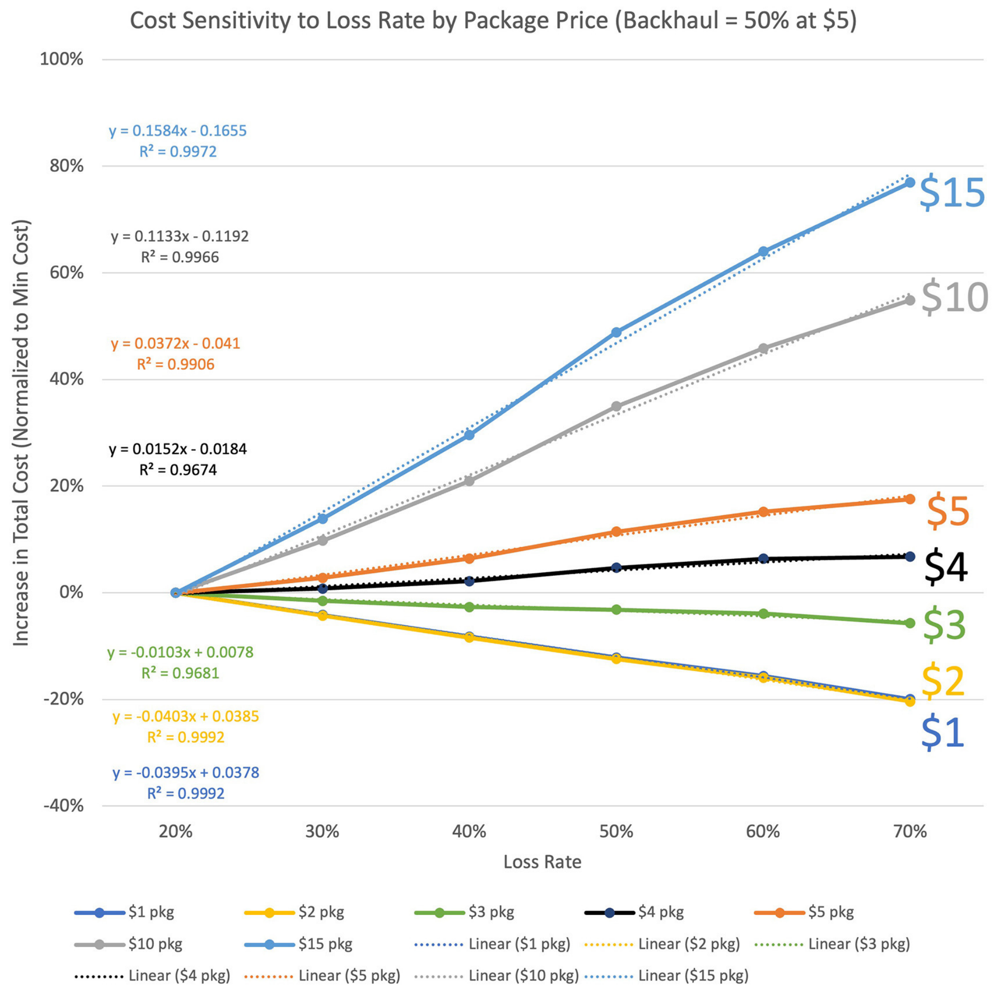


FIGURE 6
Cost sensitivity to loss rate for packages priced (color) between $1 and $15, and for 50% of packages subject to an extra backhaul cost of $5. Compared to Figure 3 (10% backhaul rate), the sensitivity pattern changed significantly suggesting that total cost is sensitive to the number of packages subject to a backhaul cost (backhaul rate).

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

KB: conceptualization, methodology, investigation, and writing—original draft. EG-F and EP-C: conceptualization, methodology, investigation, and writing—review and editing. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acampora, A., Preziosi, M., Merli, R., and Lucchetti, M. C. (2017). Environmental management systems in the wine industry: identification of best practices toward a circular economy. In *Proceedings of the 23rd International Sustainable Development Research Society Conference* (Bogotá, Colombia), p. 14–16.

Accorsi, R., Cascini, A., Cholette, S., Manzini, R., and Mora, C. (2014). Economic and environmental assessment of reusable plastic containers: a food catering supply chain case study. *Int. J. Prod. Econ.* 152, 88–101. doi: 10.1016/j.ijpe.2013.12.014

Ansari, Z. N., Kant, R., and Shankar, R. (2020). Remanufacturing supply chain: an analysis of performance indicator areas. *Int. J. Product. Perform. Manag.* 2020, 38. doi: 10.1108/IJPPM-01-2020-0038

Ayres, R. U., and Simonis, U. E. (1994). *Industrial Metabolism: Restructuring for Sustainable Development. Vol. 376*. Tokyo: United Nations University Press.

Batista, L., Bourlakis, M., Smart, P., and Maull, R. (2018). In search of a circular supply chain archetype – a content-analysis-based literature review. *Product. Plan. Control* 29, 438–451. doi: 10.1080/09537287.2017.1343502

Bracquené, E., Dewulf, W., and Duflou, J. R. (2020). Measuring the performance of more circular complex product supply chains. *Resour. Conserv. Recycl.* 154, 104608. doi: 10.1016/j.resconrec.2019.104608

Breen, L. (2006). Give me back my empties or else! A preliminary analysis of customer compliance in reverse logistics practices (UK). *Manag. Res. News* 29, 532–551. doi: 10.1108/01409170610708989

Brown, P. J., and Bajada, C. (2018). An economic model of circular supply network dynamics: toward an understanding of performance measurement in the context of multiple stakeholders. *Bus. Strategy Environ.* 27, 643–655. doi: 10.1002/bse.2069

Chan, F. T., and Chan, H. K. (2005). The future trend on system-wide modelling in supply chain studies. *Int. J. Adv. Manufact. Technol.* 25, 820–832. doi: 10.1007/s00170-003-1851-3

Choi, T. Y., Dooley, K. J., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *J. Operat. Manag.* 19, 351–366. doi: 10.1016/S0272-6963(00)00068-1

Cobb, B. R. (2016). Estimating cycle time and return rate distributions for returnable transport items. *Int. J. Prod. Res.* 54, 4356–4367. doi: 10.1080/00207543.2016.1162920

Coelho, P. M., Corona, B., ten Klooster, R., and Worrell, E. (2020). Sustainability of reusable packaging - current situation and trends. *Resour. Conserv. Recycl.* 6, 100037. doi: 10.1016/j.rcrx.2020.100037

Corona, B., Shen, L., Reike, D., Carreón, J. R., and Worrell, E. (2019). Towards sustainable development through the circular economy—a review and critical assessment on current circularity metrics. *Resour. Conserv. Recycl.* 151, 104498. doi: 10.1016/j.resconrec.2019.104498

Dai, T., and Tang, C. S. (2021). *Integrating ESG Measures and Supply Chain Management: Research Opportunities in the Post-Pandemic Era. SSRN Scholarly Paper ID 3939968*. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3939968

De Pascale, A., Arbolino, R., Szopik-Depczyńska, K., Limosani, M., and Ioppolo, G. (2021). A systematic review for measuring circular economy: the 61 indicators. *J. Clean. Prod.* 281, 124942. doi: 10.1016/j.jclepro.2020.124942

de Pauw, I., van der Grinten, B., Flipsen, B., and Alvarez-Mendez, L. (2021). *The Circularity Calculator, a Tool for Circular Product Development with Circularity and Potential Value Capture Indicators*. Available online at: https://hdl.handle.net/10344/10207 (accessed December 2, 2021).

Ellen MacArthur Foundation and Granta Design (2015). *Circularity Indicators. An Approach to Measuring Circularity. Methodology*. Available online at: https://ellenmacarthurfoundation.org/material-circularity-indicator (accessed December 12, 2021).

Ellen MacArthur Foundation. (2013). *Towards the Circular Economy, Vol. 1: An Economic and Business Rationale for an Accelerated Transition*. Available online at: https://ellenmacarthurfoundation.org/towards-the-circular-economy-vol-1-an-economic-and-business-rationale-for-an (accessed December 11, 2021).

Escursell, S., Llorach-Massana, P., and Roncero, M. B. (2021). Sustainability in E-commerce packaging: a review. *J. Clean. Prod.* 280, 124314. doi: 10.1016/j.jclepro.2020.124314

Fan, Y., and Fang, C. (2020). Circular economy development in china-current situation, evaluation and policy implications. *Environ. Impact Assess. Rev.* 84, 106441. doi: 10.1016/j.eiar.2020.106441

Farooque, M., Zhang, A., Thürer, M., Qu, T., and Huisingh, D. (2019). Circular supply chain management: a definition and structured literature review. *J. Clean. Prod.* 228, 882–900. doi: 10.1016/j.jclepro.2019.04.303

Fashion for Good (2021). *The Rise of Reusable Packaging: Understanding the Impact & Mapping a Path to Scale*. Available online at: https://reports.fashionforgood.com/report/the-rise-of-reusable-packaging/ (accessed December 2, 2021).

Geissdoerfer, M., Morioka, S. N., de Carvalho, M. M., and Evans, S. (2018). Business models and supply chains for the circular economy. *J. Clean. Prod.* 190, 712–721. doi: 10.1016/j.jclepro.2018.04.159

Geissdoerfer, M., Savaget, P., Bocken, N. M., and Hultink, E. J. (2017). The circular economy – a new sustainability paradigm? *J. Clean. Prod.* 143, 757–768. doi: 10.1016/j.jclepro.2016.12.048

Geyer, R., Jambeck, J. R., and Law, K. L. (2017). Production, use, and fate of all plastics ever made. *Sci. Adv.* 3, 25–29. doi: 10.1126/sciadv.1700782

Ghisellini, P., Cialani, C., and Ulgiati, S. (2016). A review on circular economy: the expected transition to a balanced interplay of environmental and economic systems. *J. Clean. Prod. Toward Post Fossil Carbon Soc.* 114, 11–32. doi: 10.1016/j.jclepro.2015.09.007

Goellner, K. N., and Sparrow, E. (2014). An environmental impact comparison of single-use and reusable thermally controlled shipping containers. *Int. J. Life Cycle Assess.* 19, 611–619. doi: 10.1007/s11367-013-0668-z

Goudenege, G., Chu, C., and Jemai, Z. (2013). Reusable containers management: from a generic model to an industrial case study. *Supply Chain For.* 14, 26–38. doi: 10.1080/16258312.2013.11517313

Govindan, K., and Hasanagic, M. (2018). A systematic review on drivers, barriers, and practices towards circular economy: a supply chain perspective. *Int. J. Prod. Res.* 56, 278–311. doi: 10.1080/00207543.2017.1402141

Govindan, K., and Soleimani, H. (2017). A review of reverse logistics and closed-loop supply chains: a journal of cleaner production focus. *J. Clean. Prod.* 142, 371–384. doi: 10.1016/j.jclepro.2016.03.126

Graedel, T. E., Allwood, J., Birat, J. P., Buchert, M., Hagelüken, C., Reck, B. K., et al. (2011). What do we know about metal recycling rates? *J. Indus. Ecol.* 15, 355–366. doi: 10.1111/j.1530-9290.2011.00342.x

Haghighi, S. M., Torabi, S. A., and Ghasemi, R. (2016). An integrated approach for performance evaluation in sustainable supply chain networks (with a case study). *J. Clean. Prod.* 137, 579–597. doi: 10.1016/j.jclepro.2016.07.119

Hazen, B. T., Russo, I., Confente, I., and Pellathy, D. (2021). Supply chain management for circular economy: conceptual framework and research agenda. *Int. J. Logist. Manag.* 32, 510–537. doi: 10.1108/IJLM-12-2019-0332

Hoekstra, A. Y., and Wiedmann, T. O. (2014). Humanity's unsustainable environmental footprint. *Science* 344, 1114–1117. doi: 10.1126/science.1248365

Kalmykova, Y., Sadagopan, M., and Rosado, L. (2018). Circular economy – from review of theories and practices to development of implementation tools. *Resour. Conserv. Recycl. Sustain. Resour. Manag. Circul. Econ.* 135, 190–201. doi: 10.1016/j.resconrec.2017.10.034

Kirchherr, J., Reike, D., and Hekkert, M. (2017). Conceptualizing the circular economy: an analysis of 114 definitions. *Resour. Conserv. Recycl.* 127, 221–232. doi: 10.1016/j.resconrec.2017.09.005

Korhonen, J., Honkasalo, A., and Seppäl,ä, J. (2018). Circular economy: the concept and its limitations. *Ecol. Econ.* 143, 37–46. doi: 10.1016/j.ecolecon.2017.06.041

Kristensen, H. S., and Mosgaard, M. A. (2020). A review of micro level indicators for a circular economy – moving away from the three dimensions of sustainability? *J. Clean. Prod.* 243, 118531. doi: 10.1016/j.jclepro.2019.118531

Linder, M., Boyer, R. H., Dahllöf, L., Vanacore, E., and Hunka, A. (2020). Product-level inherent circularity and its relationship to environmental impact. *J. Clean. Prod.* 260, 121096. doi: 10.1016/j.jclepro.2020.121096

Linder, M., Sarasini, S., and van Loon, P. (2017). A metric for quantifying product-level circularity. *J. Indus. Ecol.* 21, 545–558. doi: 10.1111/jiec.12552

Maestrini, V., Luzzini, D., Maccarrone, P., and Caniato, F. (2017). Supply chain performance measurement systems: a systematic review and research agenda. *Int. J. Prod. Econ.* 183, 299–315. doi: 10.1016/j.ijpe.2016.11.005

Mahmoudi, M., and Parviziomran, I. (2020). Reusable packaging in supply chains: a review of environmental and economic impacts, logistics system designs, and operations management. *Int. J. Prod. Econ.* 228, 107730. doi: 10.1016/j.ijpe.2020.107730

Meherishi, L., Narayana, S. A., and Ranjani, K. S. (2019). Sustainable packaging for supply chain management in the circular economy: a review. *J. Clean. Prod.* 237, 117582. doi: 10.1016/j.jclepro.2019.07.057

Mesa, J., Esparragoza, I., and Maury, H. (2018). Developing a set of sustainability indicators for product families based on the circular economy model. *J. Clean. Prod.* 196, 1429–1442. doi: 10.1016/j.jclepro.2018.06.131

Mhatre, P., Panchal, R., Singh, A., and Bibyan, S. (2021). A systematic literature review on the circular economy initiatives in the European Union. *Sustain. Prod. Consumpt.* 26, 187–202. doi: 10.1016/j.spc.2020.09.008

Mollenkopf, D., Closs, D., Twede, D., Lee, S., and Burgess, G. (2005). Assessing the viability of reusable packaging: a relative cost approach. *J. Bus. Logistic*. 26, 169–197. doi: 10.1002/j.2158-1592.2005.tb00198.x

Neely, A., Gregory, M., and Platts, K. (1995). Performance measurement system design. *Int. J. Operat. Prod. Manag.* 15, 80–116. doi: 10.1108/01443579510083622

Olugu, E. U., Wong, K. Y., and Shaharoun, A. M. (2011). Development of key performance measures for the automobile green supply chain. *Resour. Conserv. Recycl.* 55, 567–579. doi: 10.1016/j.resconrec.2010.06.003

Pålsson, H. (2018). *Packaging Logistics: Understanding and Managing the Economic and Environmental Impacts of Packaging in Supply Chains*. London: Kogan Page Publishers.

Pearce, D. W., and Turner, R. K. (1990). *Economics of Natural Resources and the Environment*. Baltimore, MD: JHU Press.

Saidani, M., Yannou, B., Leroy, Y., Cluzel, F., and Kendall, A. (2019). A taxonomy of circular economy indicators. *J. Clean. Prod.* 207, 542–559. doi: 10.1016/j.jclepro.2018.10.014

Sarja, M., Onkila, T., and Mäkelä, M. (2021). A systematic literature review of the transition to the circular economy in business organizations: obstacles, catalysts and ambivalences. *J. Clean. Prod.* 286, 125492. doi: 10.1016/j.jclepro.2020.125492

Schöggl, J. P., Stumpf, L., and Baumgartner, R. J. (2020). The narrative of sustainability and circular economy - a longitudinal review of two decades of research. *Resour. Conserv. Recycl.* 163, 105073. doi: 10.1016/j.resconrec.2020.105073

Scott, K., Roelich, K., Owen, A., and Barrett, J. (2018). Extending European energy efficiency standards to include material use: an analysis. *Climate Pol.* 18, 627–641. doi: 10.1080/14693062.2017.1333949

Sovacool, B. K., Ali, S. H., Bazilian, M., Radley, B., Nemery, B., Okatz, J., et al. (2020). Sustainable minerals and metals for a low-carbon future. *Science* 367, 30–33. doi: 10.1126/science.aaz6003

Stahel, W. R. (2013). Policy for material efficiency—sustainable taxation as a departure from the throwaway society. *Philos. Trans. Royal Soc. A* 371, 20110567. doi: 10.1098/rsta.2011.0567

Stahel, W. R. (2016). The circular economy. *Nature* 531, 435–438. doi: 10.1038/531435a

Stevens, G. C. (1989). Integrating the supply chain. *Int. J. Phys. Distribut. Mater. Manag.* 19 , 3–8. doi: 10.1108/EUM0000000000329

Tisserant, A., Pauliuk, S., Merciai, S., Schmidt, J., Fry, J., Wood, R., et al. (2017). Solid waste and the circular economy: a global analysis of waste treatment and waste footprints. *J. Indus. Ecol.* 21, 628–640. doi: 10.1111/jiec.12562

Tsiliyannis, C. A. (2005). A new rate index for environmental monitoring of combined reuse/recycle packaging systems. *Waste Manag.* 23, 304–313. doi: 10.1177/0734242X05056757

Van Buren, N., Demmers, M., Van der Heijden, R., and Witlox, F. (2016). Towards a circular economy: the role of dutch logistics industries and governments. *Sustainability* 8, 647. doi: 10.3390/su8070647

Vegter, D., van Hillegersberg, J., and Olthaar, M. (2020). Supply chains in circular business models: processes and performance objectives. *Resour. Conserv. Recycl.* 162, 105046. doi: 10.1016/j.resconrec.2020.105046

Vegter, D., van Hillegersberg, J., and Olthaar, M. (2021). Performance measurement systems for circular supply chain management: current state of development. *Sustainability* 13, 12082. doi: 10.3390/su132112082

Walker, S., Coleman, N., Hodgson, P., Collins, N., and Brimacombe, L. (2018). Evaluating the environmental dimension of material efficiency strategies relating to the circular economy. *Sustainability* 10, 666. doi: 10.3390/su10030666

WBCSD (2021). Circular economy. World Business Council for Sustainable Development. (2021). Available online at: https://www.wbcsd.org/CE (accessed December 11, 2021).

WEF (2021). *Circular Economy*. World Economic Forum. Available online at: https://www.weforum.org/projects/circular-economy/ (accessed December 11, 2021).

Zimmermann, T., and Bliklen, R. (2020). Single-use vs. reusable packaging in e-commerce: comparing carbon footprints and identifying break-even points. *Ecol. Perspect. Sci. Soc.* 29, 176–183. doi: 10.14512/gaia.29.3.8

frontiers | Frontiers in Sustainability

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/frsus.2022.943635

OPEN ACCESS

EDITED BY
Alexander Ryota Keeley,
Kyushu University, Japan

REVIEWED BY
Neha Gupta,
Amity University, India
Jingkuang Liu,
Guangzhou University, China

*CORRESPONDENCE
Ademilade Aboginije
aboginijeademilade@gmail.com

SPECIALTY SECTION
This article was submitted to Quantitative Sustainability Assessment, a section of the journal Frontiers in Sustainability

RECEIVED 13 May 2022
ACCEPTED 16 November 2022
PUBLISHED 09 January 2023

CITATION
Aboginije A, Aigbavboa C and Thwala W (2023) Modeling and usage of a sustainametric technique for measuring the life-cycle performance of a waste management system: A case study of South Africa. *Front. Sustain.* 3:943635. doi: 10.3389/frsus.2022.943635

Modeling and usage of a sustainametric technique for measuring the life-cycle performance of a waste management system: A case study of South Africa

Ademilade Aboginije[1]*, Clinton Aigbavboa[2] and Wellington Thwala[3]

[1]Department of Civil Engineering Science, University of Johannesburg, Johannesburg, South Africa, [2]Department of Construction Management and Quantity Surveying, University of Johannesburg, Johannesburg, South Africa, [3]School of Engineering, University of South Africa, Johannesburg, South Africa

The use of eco-friendly materials, waste prevention protocols, the support and participation of building construction stakeholders, polluter pays concepts, producer responsibility, life-cycle system thinking, and the application of cost-efficiency and cost minimization strategies are some of the guiding philosophies that are of extreme value when designing a waste management system *via* circular economy initiatives. However, it is crucial to measure the waste management strategy used in each building project. In order to measure the life-cycle performance of waste management systems and to assess how sustainable they are, this study offers a statistical methodology using a sustainametric technique to indicate how sustainable waste management system performance in emerging construction industries, particularly in South Africa. This study employs a sustainametric approach to evaluate the life-cycle performance of the waste management system of South Africa, with evidence of its sustainability performance measurement that can help advance the its waste minimization policy and implementation. The result indicates the viability of the measuring model and the findings of each metric utilized. The conclusion confirms that South Africa has not fully adopted and/or implemented a more sustainable waste management system for efficient waste minimization during its construction activities. Moreover, it is the reality that most emerging economies urgently need to expand and improve the waste reduction method employed in its construction building projects.

KEYWORDS

building materials, construction and demolition waste, life-cycle assessment, sustainametric technique, sustainable development

Introduction

The acceleration in construction activity has brought about urbanization and, consequently, a quick increase in populations in several countries. However, rapid urbanization and industrialization have increased production and consumption processes, resulting in waste generation. Furthermore, because no concrete waste disposal standards exist, the environment has been clogged with garbage in numerous developing countries (Aboginije et al., 2020). Every year, thousands of demolitions occur, all of which have significant environmental and economic consequences since building materials have become unrecoverable and must be disposed of in landfills (Akinade et al., 2018; Huang et al., 2018). Furthermore, construction and building operations consume 3 billion metric tons of raw materials per year, accounting for 40% of global consumption. Similarly, annual construction production requires 170 million metric tons of basic materials and goods, 125 million metric tons of mining products, and 70 metric tons of secondary recycled and recovered products. An estimated 6 million metric tons of energy are used, and 23 metric million tons of CO_2 are emitted from the process. According to global research, at least 9.0% of materials purchased for construction operations end up as waste due to on-site waste generation (Abioye and Rao, 2015).

Furthermore, as shown in Figure 1, waste created by building, demolition, and remodeling activities account for up to 40% of total waste generated in most nations. As a result, building or demolition waste might be found on job sites. Most countries dispose of approximately 15–30% of their waste in landfills (Thomas and Lizzi, 2011; Aboginije et al., 2021). Waste may be efficiently controlled at its source. Also, the amount of waste generated fluctuates depending on population density and urban growth, with roughly 80% of on-site waste being recyclable and usable. As a result, every effort is made around the world to manage building waste in a more sustainable manner (Liu et al., 2019). A sustainable waste management system (WMS) is anticipated to prevent harmful effects of waste on the environment or aesthetics, according to the UN 2017, but its efficiency is visible when it is efficiently managed (United Nations Environment Programme, 2017; Islam et al., 2020). Several waste management strategies are used to minimize on-site waste, although some of the strategies are unsuitable. Thus, a sustainable waste management system following circular economy principles is essential to reduce or eliminate waste in the construction sector, but the circular economy principles are painstakingly implemented (Ginindza and Muzenda, 2013; Aboginije et al., 2021).

Similarly, sustainability, utilization of eco-friendly materials, waste preventive protocol, support and involvement of building construction stakeholders, polluter pays concepts, producer responsibility, life-cycle system thinking, and the implementation of cost-efficiency and cost minimization strategies are some of the important guiding principles to follow when designing a waste system using circular economy initiatives (Nagapan et al., 2012; Velenturf et al., 2019). Furthermore, there should be legislative laws and guidance to enable the application of such sustainable WMSs. For instance, European waste legislation and prevention strategies aim to reduce waste before construction begins; this is accomplished by detailed design and material-use plans, which are critical in lowering purchase prices and the volume of recyclable materials. Although waste prevention and reduction begin with the manufacture of building materials, it is necessary to improve waste generation throughout the production process to avoid waste later in the construction process (Nagapan et al., 2012; Jingkuang and Yue, 2022). The major characteristic of a sustainable waste management system (WMS) is that it uses waste as an input material to create new value products. The goal is to reduce waste generation through reuse and recycling, minimizing the need for landfill space, extracting the maximum value from waste, and limiting the environmental effects of unavoidable wastes.

This means that by recovering materials, the volume of waste dumped in landfills may be minimized, and a sustainable waste management system can recover 90.0% of building waste (Kumar et al., 2017). Contractors also employ a variety of reuse techniques when building. For example, broken bricks and stones can be used as a subgrade to enable access to the construction site, and timber or plywood can be used to build temporary structures on site. According to Shen et al. (2004), reusing and recycling of construction materials greatly reduce landfill areas. Furthermore, storage equipment must be developed to meet the requirements for proper waste storage. Following waste storage guidelines, it should be ensured that necessary actions are taken after waste has been stored (Begum et al., 2010; Udawatta et al., 2015; Wu et al., 2016; Jingkuang et al., 2020). Unfortunately, most construction companies, especially those in developing countries, do not prioritize proper waste storage in any of their projects. Waste components must be minimized in product and building materials, or the quantity of material used, and the potential toxicity of waste generated during manufacturing and after utilization must be decreased (Jingkuang et al., 2022; Yuan, 2017).

Therefore, a sustainable waste management system is required, and while developing a waste management system, the volume or size of the trash and the composition of the waste should be considered. As many major towns are intending to close their landfills, these considerations will aid project

Abbreviations: EPI, Environmental Performance Index; ESI, Environmental Sustainability Index; GGEI, Global Green Economy Index; SPM, sustainability performance measurement; SPI, sustainability performance indicator; TBL, triple bottom line; WMS, waste management system.

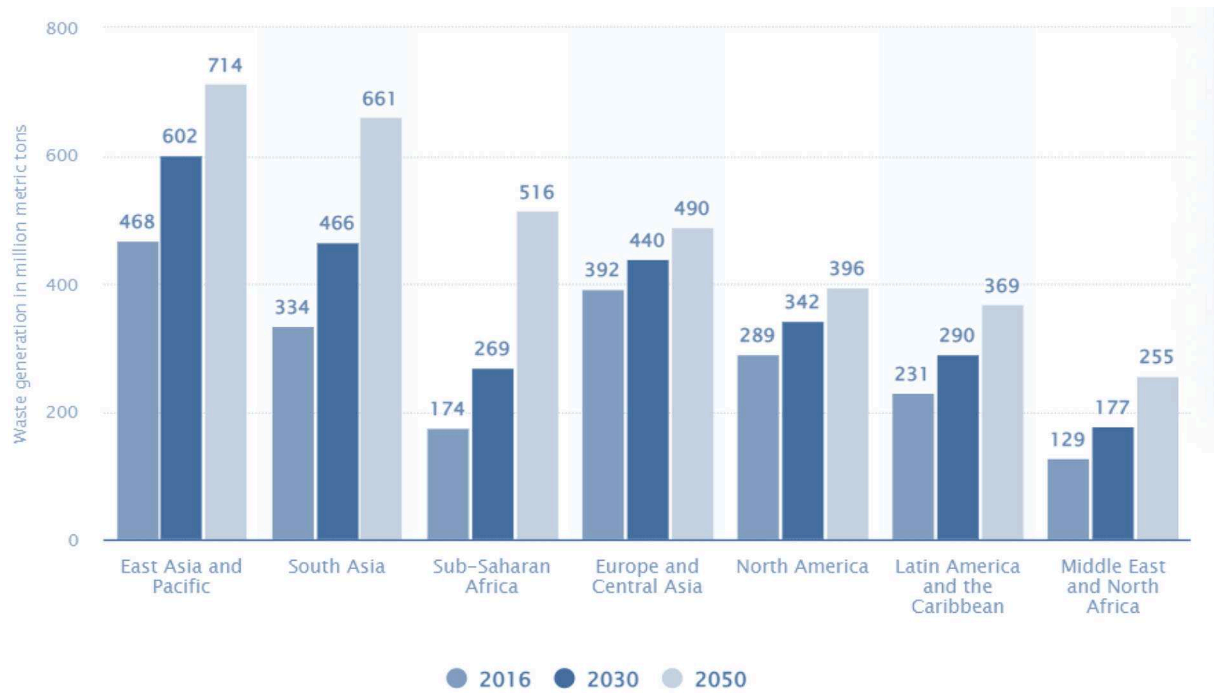


FIGURE 1
Projection of waste generation worldwide in 2016, 2030, and 2050 by region in million metric tons (Statista, 2022).

managers in allocating the appropriate volume of landfills for each waste concern. Furthermore, this will take long time to eradicate garbage generated in construction sites each year (Bojan et al., 2017). The purpose of the WMS design is to develop a sustainable environment by meeting the waste management mandates of countries. Other suitable standards such as waste avoidance, total waste generation reduction, and creation of a product reuse system should also be an integral part of the waste management system (Aboginije et al., 2021). The requirement for sustainability measurement, which can be applied in determining any waste management system performance, including improvement in operations, performance benchmarking, progress tracking device, and process evaluation, has gained the attention of researchers.

Therefore, this research aims to provide new knowledge and understanding by providing a mechanism that can be used to assess the sustainability of any country's WMS. The scoring mechanism is used to specify whether the system is sustainable and/or whether there is a consequential exigency to optimize the system. In addition, the solution provided will be beneficial to the construction sector of several economies, especially developing countries. The research objective is to design a mechanism to measure the performance of the waste management system of any country using sustainametric techniques. The sustainametric techniques are a set of measurement variables that obey sustainability principles.

Research methodology

The aim of any sustainable construction in the construction sector is to achieve sustainable development, which entails integrating sustainable principles into effective strategic frameworks. The goal of this research is to provide an indicator-based framework for measuring and evaluating the sustainability of any construction and demolition waste management system. As a result, developing a good sustainability indicator is essential, and an indicator-based framework is created to accomplish this goal. The performance of social, environmental, and economic aspects is used to evaluate sustainability, according to the U.S. Department of Transportation (Moldan et al., 2012; Singh et al., 2012). Some of the most well-known and widely used sustainability measures, according to Singh et al. (2012), are corporate sustainability reporting, triple bottom line accounting, and estimates of the quality of sustainability governance for individual countries using the Global Green Economy Index (GGEI), Environmental Sustainability Index (ESI), and Environmental Performance Index (EPI).

The TBL principle

The TBL concept results from a paradigm of sustainable development that is usually used to measure any performance,

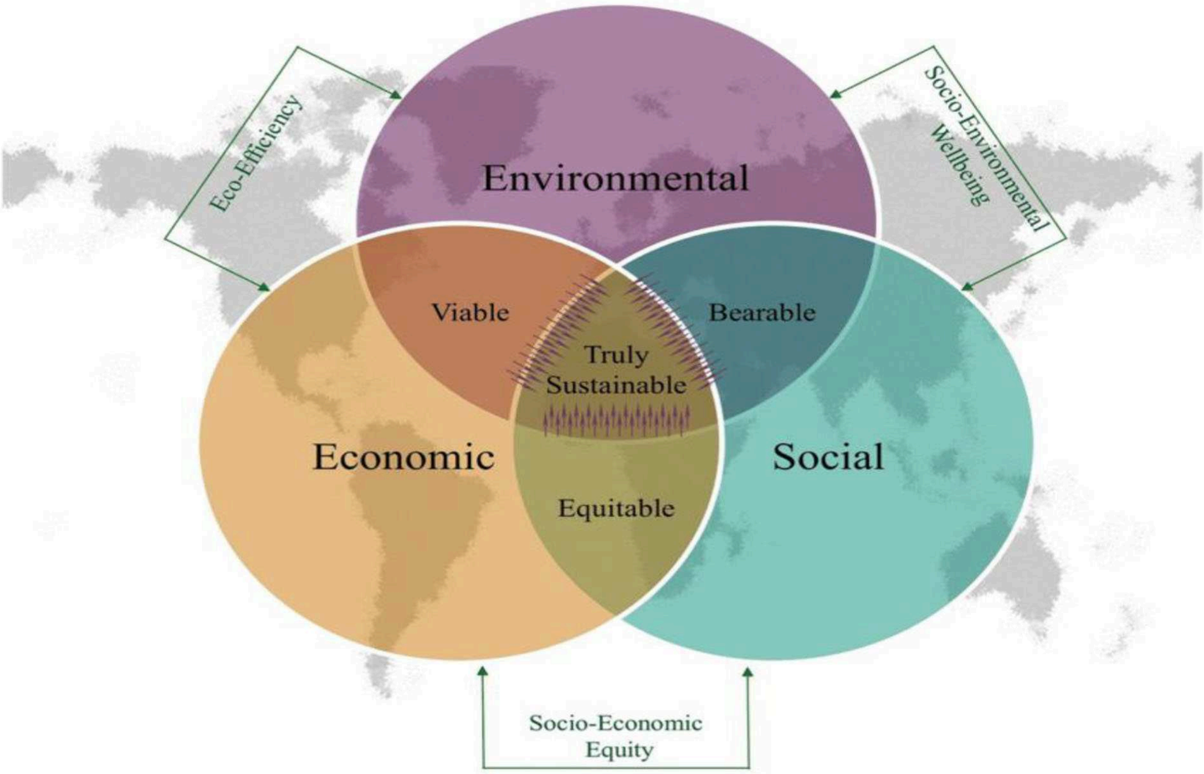


FIGURE 2
Triple bottom line concept (Tasdemir et al., 2020).

but there is a need to find a balance between the three dimensions, as illustrated in Figure 2. The TBL is described as a framework that produces nonpolluting goods and services while preserving energy and natural resources. It is also economically viable, safe, and healthy. Furthermore, it enables an organization to review its actions by considering not only the economic values generated but also the environmental and social values that might be multiplied or diminished. While achieving sustainability by balancing the triple bottom line principles is an ideal objective that can assist and guide decision-making, it will not be possible in every project. It is certainly possible to measure and report the environmental bottom line, albeit it can be a time-consuming and challenging procedure depending on the size of the company (Scerri and James, 2010; Sridhar, 2012).

Executing a sustainable WMS in achieving a green economy will support mitigating the climate crisis in terms of pollution prevention, among other things (Xiao et al., 2018). Since waste management is an integral part of the TBL of sustainability, companies should aim to address these issues, which require strong commitment and leadership as well as drastic changes. Several countries do not stop at merely making it viable, equitable, or bearable but, instead, aim for its sustainability. In this study, the TBL is used to understand the indications of the various impacts of sustainable waste management across the three sustainability patterns (Bell and Morse, 2008; Dalal-Clayton and Sadler, 2009; Dahl, 2012; Singh et al., 2012).

Sustainability measurement criterion

Sustainability can be measured in following ways: accounts of quantitative data, the use of narrative assessments, and the use of indicator systems. Accounting of quantitative data involves changing quantitative data into common units, like money or energy, the use of narrative assessments includes the use of graphics maps and tables, and the use of indicator systems involves organizing information from narrative assessments around indicators. Indicator-based systems can be measured easily, can be compared easily, and are more objective; hence, it is reported to be able to perform better than other methods (Dalal-Clayton and Sadler, 2009). As a result, decision-makers and stakeholders must be involved in the development of indicators in order for their values and concerns to be considered. However, the system must be both technically and scientifically sound. The system must first be specified, with an appropriate system boundary drawn, before it can be studied

TABLE 1 Criteria for assessing the performance of WMS (Aboginije et al., 2020).

Metrix	Variables	Phases	Sources
Rethink or redesign	Complex design and detailing avoidance	Plan and design	Nagapan et al., 2012; Akinade et al., 2018
	Waste-reduction contract		Aboginije et al., 2021
	Reusable, recycled, or renewable materials maximization		Xiao et al., 2018; Aboginije et al., 2020
Reuse	Choice of materials that have a long service life and can be used repeatedly.	Procurement and construction	Muzenda et al., 2012; Nagapan et al., 2012
	Reduction of material quantity while increasing quality		Huang et al., 2018; Islam et al., 2020
	Implementation of sustainable material procurement.		Daylath, 2011; Udawatta et al., 2015
Recycling	Resilience secondary materials markets optimization	Initiation and construction	Abioye and Rao, 2015; Xiao et al., 2018
	Utilization of recycled materials		Ginindza and Muzenda, 2013; Aboginije et al., 2020
	Provide incentives for transactions on secondary materials		Muzenda et al., 2012; Aboginije et al., 2020
Material recovery	Material recovery maximization	Construction	Velenturf et al., 2019
	Recovering resources and energy if possible		Rahim and Kasim, 2017; Xiao et al., 2018
Residual management	Minimization of harmful environmental effect	Operation and maintenance	Begum et al., 2010; Nagapan et al., 2012
	Encourage the preservation of resources		Begum et al., 2010; Abioye and Rao, 2015
	Landfill sites conservation		Abioye and Rao, 2015; Akinade et al., 2018
	Optimization of the waste management system		Akinade et al., 2018; Velenturf et al., 2019
Policy implementation	The government enforcement of a landfill tax	Procurement and construction	Muzenda et al., 2012; Bojan et al., 2017; Huang et al., 2018
	Enaction anti-incineration legislation		Ginindza and Muzenda, 2013; Aboginije et al., 2021
On-site waste management plan	Awareness among clients and contractors	Initiation and construction	Bojan et al., 2017; Zhang et al., 2019
	Waste expertise involvement on site		Huang et al., 2018; Aboginije et al., 2021

further. The constituents of the system include the complete input, output, emissions, energy, and other secondary aspects that should be thoroughly investigated (Dong and Hauschild, 2017).

The first step involves indicator selection. This step establishes operating circumstances, process parameters, and characteristics. The indications for which measurement is required are chosen. This serves as the system metric, which will be examined in the following steps. An assessment or measurement is carried out using proper assessing tools that have been confirmed and tested, or experiments for pre-defined indicators. This is carried out to offer a value for the indicator measurement (George and Mallery, 2003; Høgevold and Svensson, 2012). After the results have been gathered, the data are properly analyzed and interpreted, and tools are utilized to improve and change the system procedures. Because of the interdisciplinary character and complexity of the challenges that this topic embodies, measuring sustainability is difficult (Troschinetz et al., 2007; Ferro et al., 2017). Methods have emerged from various fields that are focused on ecological, economic, and social considerations. First, one must know what should be done with the results of a sustainability measure, what are the major concerns, and what are the system limitations. It is often more informative to track the growth of the entity—is it more sustainable now than it was previously? It is challenging to compare similar things due to the data complexity and diversity (countries, companies, institutions, and even products), rather than trying to explain the status of sustainability in one number or a table of numbers. The usage of imagining to portray the data is a useful way to do it (Gasparatos et al., 2008; Garcia-de-Vinuesa DL, 2018).

The ideal technique for measuring sustainability would display a tripod paradigm of pollution prevention, social equity, and economic benefit, which determines actual sustainability, and what the indicators measure must be linked through the metrics. A good indicator will track how a system becomes more or less sustainable over some time (Mayer et al., 2004; Sahely et al., 2005). The work of measuring sustainability is value-laden and socially charged, which makes studying sustainability as an objective science very difficult. According to Hammond et al. (1995) and Lele and Norgaard (1996), if the aim of the analysis is known, a multidisciplinary approach to problem conceptualization and study methodologies can

FIGURE 3
Map of South Africa showing the nine provinces (Bundy et al., 2022).

be employed. Sustainability metrics are employed to measure and quantify sustainability beyond general ideas. Different international groups have their various disciplines or policies and political views, and they disagree on how sustainability should be measured.

Although sustainability metrics like reporting systems are popular among public and private sectors, they are unable to influence actual policies and practices in a society (Hermann et al., 2007; Milne and Gray, 2013). Strategies from sustainable waste management in environmental, social, and economic areas help draw the metrics used for sustainability measurement in this work. These metrics include indicators, benchmarks, audits, indexes, and accounting, as well as assessment, appraisal, and other reporting systems that are applied over a wide range of spatial and temporal dimensions, albeit they are continuously evolving. Recently, testing of intents and behaviors that are normally distributed and that pursue goals of sustainability was proposed as a methodology of sustainability monitoring. The selection of sustainability benchmark indicators was founded on sustainable principles and the life-cycle impacts of its implementation in construction projects across the construction phases. Each of the indicated metrics can be used to analyze the degree of application of sustainable strategies in the WMS of any country (Ekanayake and Ofori, 2004; Rahim and Kasim, 2017).

Sustainability performance scoring system

The demographic factors employed are grouped. Each was given a code number (e.g., 1 and 2). The objective was to show the degree to which sustainable waste reduction strategies are executed in the building construction project. This can be obtained through the contribution of respondents in a semi-structured survey put together on a scale of 1–100 with the highest score indicating very high (i.e., more than 70% execution rate). The correspondence is required to be construction professionals with vast experience and expertise with track records. From the data collected, a TBL dimension was developed to show the sustainability-based reason that includes all the dimensions of sustainability that should be considered in any project scope (Lozano, 2006; Milne and Gray, 2013; Montabon et al., 2016). This shows the three vital aspects of sustainability (social, economic, and environmental) and the variables under each as utilized by the construction industry. In terms of the distribution of each variable under an indicator across the construction life-cycle phases is tabulated. Furthermore, the set of indicators for each stage of the construction life cycle, from planning to feasibility testing to refurbishment, is described. However, certain variables

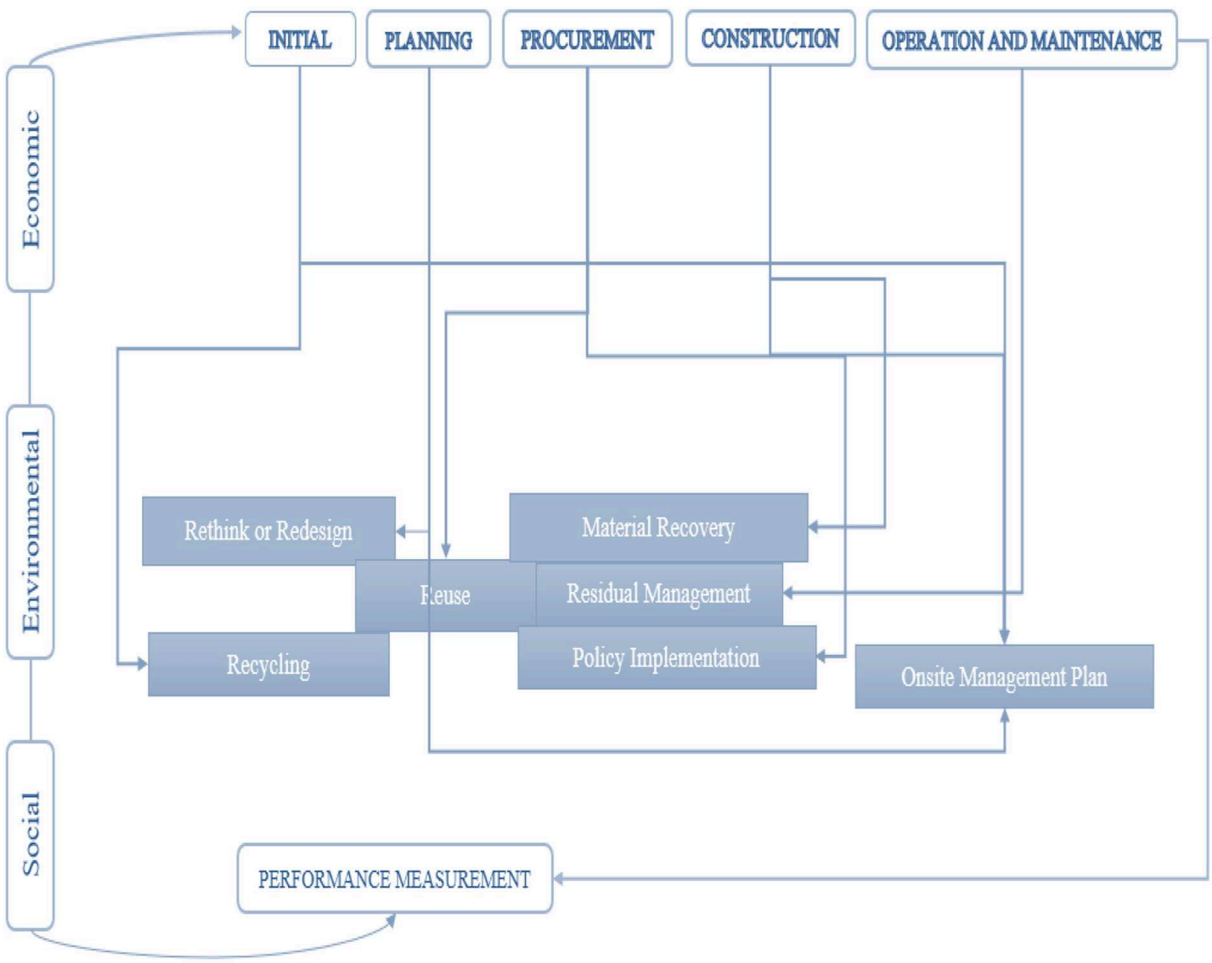


FIGURE 4
Flowchart of LCA material flow (author).

cover one or more stages, indicating that they can be used in many phases.

Data collected in this study were captured, extracted, and analyzed. The last step developed a material flowchart using life-cycle mapping. This makes the sustainable performance measurement using the triple bottom line dimension possible to indicate each variable execution rate across the construction life-cycle phases. A decisive step to provide the SPM as the requirement to measure the sustainability performance of the WMS was taken. As a result, the sustainability measurement is intended to support a decision-making mechanism by providing significant information for planning future actions prioritized in any waste sector, but it is only classified as "dimensionless," which means it is expressed in relative (percentage) measures (Nardo et al., 2005). There are also important evidence and asymmetry between the number of SPI for each triple bottom line dimension and their combination in tri-dimensional indicators (Pontius and Mclntosh, 2020). Table 1 presents the phases of the construction life cycle as construction projects advance (from the initial to the finish).

Case study area and rationale

South Africa is the second largest economy in Africa, with a growth of 1.25% predicted in 2022 from 0.98% in 2019. The country is located in the southernmost part of the African continent and covers a total size of 1.2 million square kilometers. It is noted for its cultural variety. It is bordered by Namibia, Botswana, Zimbabwe, and Mozambique. South Africa is the largest country in Southern Africa, with three capital cities, namely, Pretoria (executive), Cape Town (legislative), and Bloemfontein (administrative; judicial). It is a multicultural society with many cultures, languages, and religions. Afrikaans, English, Ndebele, Northern Sotho, Swati, Tswana, Tsonga, Venda, Xhosa, and Zulu are the 11 official

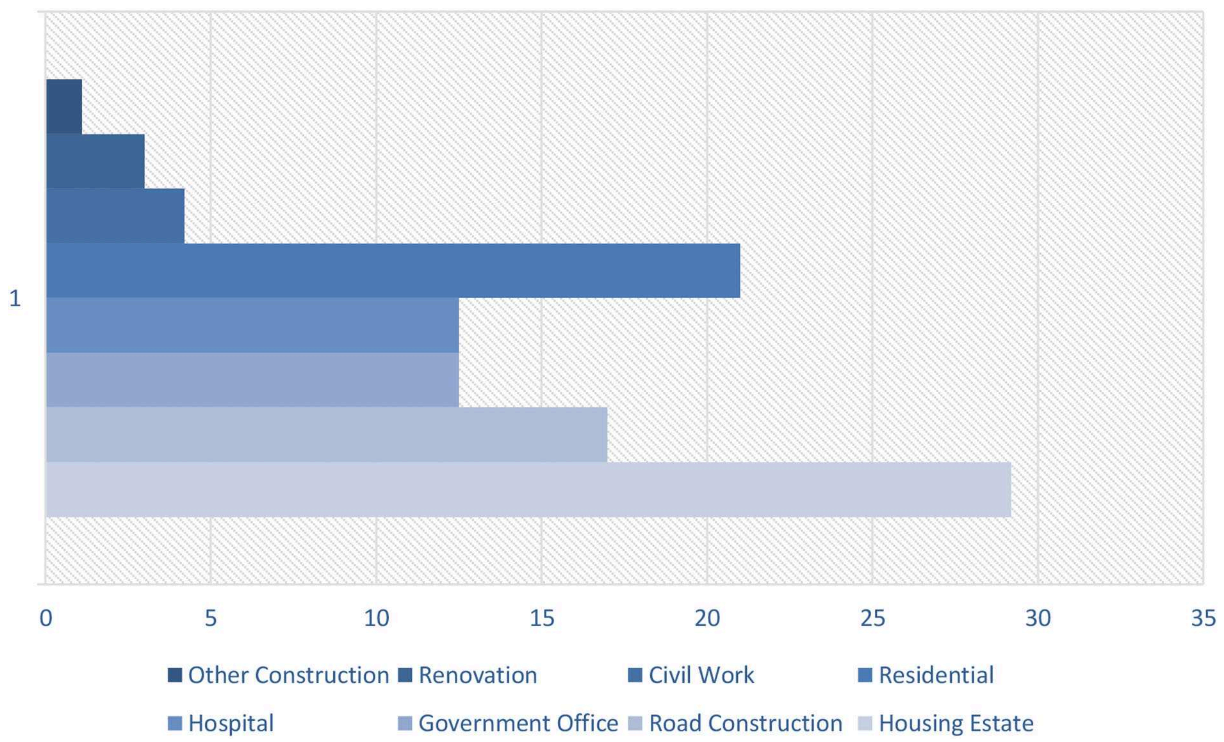


FIGURE 5
Respondents' demography according to the project executed (author).

languages of South Africa, which are spoken by a diverse ethnic population. The country is divided into nine provinces, as illustrated in Figure 3 (Bundy et al., 2022). South Africa now possesses a comprehensive legislative framework because it is still a relatively emerging economy.

However, the significant waste management problem of the country requires rapid care. Population expansion, urbanization, a lack of compliance, and general waste management behavior are some of the predominant waste sources. The population of South Africa was predicted to reach 60.14 million in mid-2021, up around 604 281 (1.01%), from mid-2020. The country is quickly urbanizing, with one of the fastest urbanization rates in the world (DEAT, 2001; Aboginije et al., 2021). As a result, the 'trash creation rate of the country is increasing daily, and attempts are being made to reduce it to a negligible level. Figure 4 illustrates the flowchart of LCA material flow.

Results and discussion

In this study, men formed a large proportion of respondents. From a total of 150 data samples that were retrieved, 73.8% were men and 25% were women, while 1.2% preferred not to identify their gender. In addition, construction stakeholders were evenly distributed to avoid any form of bias and to prevent any among the professionals from constituting a larger proportion of the population unnecessarily. On average, the respondents' years of experience were more than 15 years, and they had a bachelor's degree as their minimum qualification. The preponderance of respondents works in public consulting and contracting firms, followed by private firms, with government employees accounting for the least percentage of respondents. Also, it can be seen that 29.2% of respondents have worked on house estate projects, 17% have constructed roads, 12.5% have built government offices, 4.2% have experiences in civil works, 3.0% have worked on renovations, and only 1.1% have worked on other projects in construction. This is shown in Figure 5. A total of 83.9% of the respondents have had strong experience in CWM for the past 2 years.

George and Mallery (2003) indicated that internal consistency is a statistic and research metric that is based on the correlations between distinct test items, and it determines whether many items used to measure the same fundamental construct produce similar findings. Cronbach's alpha is used to determine the internal consistency of any collected sample. The complete variation was accounted for. An exploratory factor analysis (EFA) was also performed using the SPSS statistical tool (Peters, 2014). As shown in Table 2, principal axis factoring

10.3389/frsus.2022.943635

TABLE 2 Scores of variables.

Clusters	Variables	Grading score (%)
Rethink or redesign	Complex design and detailing avoidance	0.666
	Waste-reduction contract	0.232
	Reusable, recycled, or renewable materials maximization	0.332
Reuse	Choice of materials that have a long service life and can be used repeatedly.	0.334
	Reduction of material quantity while increasing quality	0.562
	Implementation of sustainable material procurement.	0.44
Recycling	Resilience secondary materials markets optimization	0.232
	Utilization of recycled materials	0.226
	Provide incentives for transactions on secondary materials	0.306
Material recovery	Material recovery maximization	0.203
	Recovering resources and energy if possible	0.67
Residual management	Minimization of harmful environmental effect	0.35
	Encourage the preservation of resources	0.442
	Landfill sites conservation	0.348
	Optimization of the waste management system	0.21
Policy implementation	The government enforcement of a landfill tax	0.414
	Enaction anti-incineration legislation	0.4
On-site waste management plan	Awareness among clients and contractors	0.551
	Waste expertise involvement on site	0.449

TABLE 3 Normality test.

Clusters	Variables	Grading score (%)
Rethink or redesign	Complex design and detailing avoidance	0.666
	Waste-reduction contract	0.232
	Reusable, recycled, or renewable materials maximization	0.332
Reuse	Choice of materials that have a long service life and can be used repeatedly.	0.334
	Reduction of material quantity while increasing quality	0.562
	Implementation of sustainable material procurement.	0.44
Recycling	Resilience secondary materials markets optimization	0.232
	Utilization of recycled materials	0.226
	Provide incentives for transactions on secondary materials	0.306
Material recovery	Material recovery maximization	0.203
	Recovering resources and energy if possible	0.67
Residual management	Minimization of harmful environmental effect	0.35
	Encourage the preservation of resources	0.442
	Landfill sites conservation	0.348
	Optimization of the waste management system	0.21
Policy implementation	The government enforcement of a landfill tax	0.414
	Enaction anti-incineration legislation	0.4
On-site waste management plan	Awareness among clients and contractors	0.551
	Waste expertise involvement on site	0.449

separates the variables into seven factorial components. Each cumulative deviation was calculated in percentage, and the total deviation was derived for the life-cycle phase of each building.

In the descriptive statistics, there were seven extracted items loaded into seven clusters. The variables used for this study were obtained from previous studies and primary literature sources reviewed by the researchers. In cluster 1, *"Rethink and Redesign"*, three factors were loaded. Avoidance of a complex design and detailing was scored the highest, with a 66.6% rating in the application, while contractual agreement on waste reduction, and design and purchase of reusable, recycled, or sustainably renewable materials were rated 33.4%, respectively, which implies that the latter was a barely used sustainable waste management strategy in the South African construction industry. In cluster 2, *"Reuse"*, three factors were loaded with the indication that 56.2% used a selection of materials that maximize the usable lifespan and opportunities for continuous use, 23.2% minimizing the quantity and maximizing the quality of materials, and 22.6% implementation of sustainable procurement.

In cluster 3, *"Recycling"*, three factors were loaded, indicating that 40.1% for use of resilience secondary materials market optimization, 30.6% utilization of recycled materials, and 20.3% provision for incentives in transactions on secondary materials. In cluster 4, *"Material Recovery"*, two factors were loaded, with 67.0% for material recovery maximization and 33.0% recovery of resources and energy if possible. In cluster 5, *"Residual Management"*, three factors were loaded, with 44.2% for the encouragement of natural resources, 34.8% minimization of negative impact on the environment, and 21.0% landfill site conservation. In cluster 6, *"Policy Implementation"*, two factors were loaded, with 41.4% for imposture of landfill tax by the government and 40.0% institution of laws against incineration. In cluster 7, *"On-site waste management plan"*, two factors were loaded, with awareness among clients and contractors scored 55.1% and waste expertise involvement on sites scored 44.9%. Table 2 shows the grading score of each of the variables.

While the normality test was used to evaluate if variables were regularly distributed or not. The normality test was conducted with 0.05 as the lowest value. For sample sizes < 50, statistical results were based on the "Kolmogorov–Smirnov" test, while results for sample sizes < 50 were based on the "Shapiro–Wilk" test. Because our sample size was greater than 50 in this study, the "Kolmogorov–Smirnov" was used. The p-value was < 0.05, according to the normality test, which makes it a suitable analysis. In each cluster, there are signs that some of the variables indicate a high performance of the waste management system, while others show a very low implementation rate. Table 3 shows the grading score for the normality test. The result indicates that the most common waste minimization strategy achieved is avoidance of complex design and detailing, material recovery maximization, and selection of materials that maximize the usable lifespan, while landfill site conservation was found to be the least of the waste minimization strategy in operation in South Africa. There is obviously a poor procurement mechanism, inadequate landfill site conservation, and lack of provision made for incentives in transactions on secondary materials.

Conclusion and recommendation

In South Africa, there is a noteworthy advance in the implementation of the sustainable WMS. However, the governments and other building stakeholders must ensure that a sustainable WMS is in operation from the feasibility study through project completion to decrease waste to the lowest possible level. In this study, the model applied for the grading system can be validated, and the result of each metric used for the measurement is viable. In a nutshell, the construction industry of South Africa is yet to fully adopt and implement a sustainable waste management system for effective waste minimization, although the overall performance shows that the construction sector is thriving and improving in its approach to waste management. There is an imperative requirement to upscale and upgrade the current waste management system applied to minimizing waste in the construction industry of South Africa. At the moment, there is an increase in research on sustainametric application in sustainability performance measurement, but further sustainametric/and statistical mechanisms can be used to model a pattern for the optimization of the WMS in any construction sector.

Furthermore, the waste management system implemented in South Africa has the potential to be much more evident in terms of job creation and possibilities, cost savings, and resource conservation, especially when integrated into the process of recycling and reusing waste materials. In addition, despite government taxes and penalties for unlawful dumping, many municipalities in South Africa still dispose of their waste in landfills. Therefore, there is still room for improvement in the waste management sector operations, given the low compliance rate with the sustainable waste management policy and framework. If appropriately applied, the sustainability assessment approaches mentioned in this study can aid in understanding any waste management systems in place and determining their sustainability.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

AA pioneers the research focus and does the literature background study. CA and WT give advice, correction on the technicality, and significance of the study. All authors contributed to the article and approved the submitted version.

Funding

This research was fully funded by the National Research Foundation, South Africa (Grant No. UID214365).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abioye, A., and Rao, B. (2015). Sustainable approach to managing construction and demolition waste: an opportunity or a new challenge. *J. Innov. Res. Sci. Eng. Tech.* 4, 10368–10377. doi: 10.15680/IJIRSET.2015.0411007

Aboginije, A. J., Aigbavboa, C. O., and Thwala, W. (2021). A holistic assessment of construction and demolition waste management in Nigerian construction projects. *Sust. J. Sust.* 13, 6241. doi: 10.3390/su13116241

Aboginije, A. J., Aigbavboa, C. O., Thwala, W., and Samuel, S. (2020). "Determining the impact of construction and demolition waste reduction practices on green building projects in Gauteng province, South Africa," in *10th Annual International Conference of Industrial Engineering and Operational Management,* Dubai, United Arab Emirate, 2428–2437.

Akinade, O. O., Oyedele, L. O., and Ajayi, S. O. (2018). Designing out construction waste using BIM technology: Stakeholders' expectations for industry deployment. *J. Clean. Prod.* 180, 375–385. doi: 10.1016/j.jclepro.2018.01.022

Begum, R. A., Satari, S. K., and Pereira, J. J. (2010). Waste generation and recycling: comparison of conventional and industrialized building systems. *Am. J. Environ. Sci.* 6, 383–388. doi: 10.3844/ajessp.2010.383.388

Bell, S., and Morse, S. (2008). *Sustainability Indicators: Measuring the Immeasurable?* London: Earthscan.

Bojan, R., Neven, V., and Branka, I. (2017). Concept of sustainable waste management in the city of Zagreb: towards the implementation of the circular economy approach. *J. Air Waste Manage. Assoc.* 67, 241–259. doi: 10.1080/10962247.2016.1229700

Bundy, C. J. Thompson, L. M., Gordon, D. F., Vigne, R., Cobbing, R. D., Hall, M. N. (2022). *South Africa Encyclopedia Britannica*, Available online at: https://www.britannica.com/place/South-Africa (accessed 11 May 2022).

Dahl, A. L. (2012). Achievements and gaps in indicators for sustainability. *Ecol. Indic.* 17, 14–19. doi: 10.1016/j.ecolind.2011.04.032

Dalal-Clayton, B., and Sadler, B. (2009). *Sustainability Appraisal. A Sourcebook and Reference Guide to International Experience*. London: Earthscan publisher.

Daylath, P. M. (2011). *Contractual obligations analysis for construction waste management. Master thesis submitted to the University of British Columbia*, Vancouver, BC, Canada.

DEAT (2001). White paper on integrated pollution and waste management for South Africa. A policy on pollution prevention, waste minimization, impact management, and remediation. Available online at: https://www.dffe.gov.za/sites/default/files/legislations/integrated_pollutionand_wastemanagement_1.pdf (accessed February 12, 2022).

Dong, Y., and Hauschild, M. Z. (2017). Indicator for environmental sustainability. *Procedia CIRP.* 61, 697–702. doi: 10.1016/j.procir.2016.11.173

Ekanayake, L. L., and Ofori, G. (2004). Building waste assessment score: design-based tool. *Build. Environ.* 39, 851–861. doi: 10.1016/j.buildenv.2004.01.007

Ferro, C., Padin, C., Svensson, G., Sosa Varela, J. C., Wagner, B., and Høgevold, N. M. (2017). Validating a framework of stakeholders in connection to business sustainability effort in supply chains. *J. Bus. Ind. Market*. 32, 124–137. doi: 10.1108/JBIM-12-2015-0253

Garcia-de-Vinuesa DL, Aguayo-Gonzalez F, and Cordoba-Roldan, A. (2018). A proposal of a framework for the evaluation of the sustainability of products from the paradigm of the circular economy based on industry 4.0 (2nd part). *DYNA* 93, 488–492. doi: 10.6036/8718

Gasparatos, A., El-Haram, M., and Horner, M. (2008). A critical review of reductionist approaches for assessing the progress towards sustainability. *Environ. Impact Assessment Rev*. 28, 286–311. doi: 10.1016/j.eiar.2007.09.002

George, D., and Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference.* 11, 0. *Update*. Boston: Allyn and Bacon.

Ginindza, B., and Muzenda, E. (2013). "Community Perception on Waste Management and Minimization: a Case Study for Mogale City and Westonaria Municipalities" in *Paper presented at the International Conference on Integrated Waste Management and Green Energy Engineering*, April 15-16, Johannesburg, South Africa 2013.

Hammond, A., Adriaanse, A., Rodenburg, E., Bryant, D., and Woodward, R. (1995). *Environmental Indicators: a Systematic Approach to Measuring and Reporting on Environmental Policy Performance in the ' Context of Sustainable Development)*. Washington, DC: World Resources Institute.

Hermann, B. G., Kroeze, C., and Jawjit, W. (2007). Assessing environmental performance by combining life cycle assessment, multi-criteria analysis, and environmental performance indicators. *J. Clean. Prod.* 15, 1878–1796. doi: 10.1016/j.jclepro.2006.04.004

Høgevold, N. M., and Svensson, G. (2012). A business sustainability model: a European case study. *J. Bus. Indu. Market*. 27, 142–151. doi: 10.1108/08858621211197001

Huang, B., Wang, X., Kua, H., Geng, Y., Bleischwitze, R., and Renf, J. (2018). Construction and demolition waste management in China through the 3R principle. *Res. Conserv. Recycl.* 129, 36–44. doi: 10.1016/j.resconrec.2017.09.029

Islam, M. D., Rahman, S. H., and Hassan, M. (2020). Municipal solid waste management using GIS application in Mirpur area of Dhaka city, Bangladesh. *Pollution* 2, 141–145. doi: 10.7508/pj.2016.02.004

Jingkuang, L., Yedan, L., and Xuetong, W. (2020). An environmental assessment model of construction and demolition waste based on system dynamics: a case study in Guangzhou. *Environ. Sci. Pollut. Res.* 27, 37237–37259. doi: 10.1007/s11356-019-07107-5

Jingkuang, L., Yixuan, C., and Xuetong, W. (2022). Factors driving waste sorting in construction projects in China. *J. Clean. Prod*. 336, 130397. doi: 10.1016/j.jclepro.2022.130397

Jingkuang, L., and Yue, T. (2022). Evolution game analysis on behavioral strategies of multiple stakeholders in construction waste resource industry chain. *Environ. Sci. Pollut. Res.* 12, 1–7. doi: 10.1007/s11356-022-23470-2

Kumar, A., Holuszko, M., and Espinosa, D. C. R. (2017). E-waste: An overview on generation, collection, legislation and recycling practices. *Resour. Conserv. Recycl.* 122, 32–42. doi: 10.1016/j.resconrec.2017.01.018

Lele, S., and Norgaard, R. B. (1996). Sustainability and the scientist's burden. *Conserv. Biol.* 10, 354–365. Available online at: http://www.jstor.org/stable/2386852

Liu, J., Li, Q., Gu, W., and Wang, C. (2019). The impact of consumption patterns on the generation of municipal solid waste in China: evidences from provincial data. Int. J. Environ. Res. Publ. Health. 16, 1–19.

Lozano, R. (2006). A tool for graphical assessment of sustainability in universities. *J. Clean. Prod*. 14, 963–972. doi: 10.1016/j.jclepro.2005.11.041

Mayer, A. L., Thurston, H. W., and Pawlowski, C. W. (2004). The multi-disciplinary influence of common sustainability indices. *Ecol. Environ.* 2, 419–426. doi: 10.1890/1540-9295(2004)002<0419:TMIOCS>2.0.CO;2

Milne, M., and Gray, R. (2013). W(h)ither ecology? The triple bottom line, the global reporting initiative, and corporate sustainability reporting. *J. Bus. Ethics* 118, 13–29. doi: 10.1007/s10551-012-1543-8

Moldan, B., Janoušková S., and Hák, T. (2012). How to understand and measure environmental sustainability: indicators and targets. *Ecol. Ind.* 17, 4–13. doi: 10.1016/j.ecolind.2011.04.033

Montabon, F., Pagell, M., and Wu, Z. (2016). Making sustainability sustainable. *J. Supp. Chain Manage.* 52, 11–27. doi: 10.1111/jscm.12103

Muzenda, E., Ntuli, F., and Pilusa, T. J. (2012). "Waste management strategies and the situation in South Africa," in *International Conference on Chemical Engineering and Technology*, Oslo, Norway, 13–14 August 2012, 309–312.

Nagapan, I. S., Abul-Rahman, I. A., and Aziz, A. A. (2012). Construction waste management: Malaysian perspective. *Int. Conf. Environ. Eng. Sust.* 2012, 299–309.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). *Handbook on constructing composite indicators: Methodology and user* *guide (STD/DOC (2005)3). Paris: OECD Statistics Directorate*. Available online at: http://www.oecd.org/dataoecd/37/42/42495745.pdf (accessed February 26, 2022).

Peters, Y. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *Eur. Health Psychol.* 16, 2.

Pontius, J., and Mclntosh, A. (2020). *Sustainability Science. Critical Skills for environmental professionals. Springer Textbooks in Earth Sciences, Geography, and Environment*. Cham: Springer.

Rahim, M. A., and Kasim, N. (2017). *Conceptual Model for Systematic Construction Waste Management. MATEC Web of Conference*. Paris: EDP Sciences.

Sahely, H. R., Kennedy, C. A., and Adams, B. J. (2005). Developing sustainability criteria for urban infrastructure systems. *Can. J. Civil Eng.* 32, 72–85. doi: 10.1139/l04-072

Scerri, A., and James, P. (2010). Accounting for sustainability: Combining qualitative and quantitative research in developing 'indicators' of sustainability. *Int. J. Soc. Res. Methodol*. 13, 41–53. doi: 10.1080/13645570902864145

Shen, L. Y., Tam, V, W., Tam, C. M., and Drew, D. (2004). Mapping approach for examining waste management on construction sites. *J. Const. Engg. Manage*, 130, 472–481.

Singh, R. K., Murty, H. R., Gupta, S. K., and Dikshit, A. K. (2012). An overview of sustainability assessment methodologies. *Ecol. Indic.* 15, 281–299. doi: 10.1016/j.ecolind.2011.01.007

Sridhar, K. (2012). Corporate conceptions of triple bottom line reporting: an empirical analysis into the signs and symbols driving this fashionable framework. *Soc. Res. J.* 8, 312–326. doi: 10.1108/17471111211247901

Statista. (2022). *Projection of waste generation worldwide in 2016, 2030, and 2050, by region*. Available online at: https://www.statista.com/statistics/233613/waste-generation-worldwide-by-region/

Tasdemir, C., Gazo, R., and Quesada, H. J. (2020). Sustainability benchmarking tool (SBT): theoretical and conceptual model proposition of a composite framework. *Environ. Dev. Sustain.* 22, 6755–6797. doi: 10.1007/s10668-019-00512-3

Thomas, H. C., and Lizzi, A. (2011). *Solid Waste Technology and Management*. Hoboken, NJ: Wiley

Troschinetz, A. M., Mihelcic, J. R., and Bradof, K. L. (2007). Developing sustainability indicators for a university campus. *Int. J. Eng. Educ.* 23, 231–241.

Udawatta, N., Zuo, J., Chiveralls, K., and Zillante, G. (2015). Improving waste management in construction projects: an Australian study. *Res. Conserv. Recycl.* 101, 73–83. doi: 10.1016/j.resconrec.2015.05.003

United Nations Environment Programme (2017). *Guidelines for national waste management strategies: Moving from challenges to opportunities*. Available online at: https://www.unep.org/about-un-environment (accessed February 12, 2022).

Velenturf, A., Jopson, P. M., and Juliet, S. (2019). Making the business case for resource recovery. *Sci. Total Environ.* 648, 1031–1041. doi: 10.1016/j.scitotenv.2018.08.224

Wu, Z., Shen, L., Yu, A. T. W., and Zhang, X. (2016). A comparative analysis of waste management requirements between five green building rating systems for new residential buildings. *J. Clean. Prod.* 112, 895–902. doi: 10.1016/j.jclepro.2015.05.073

Xiao, S., Dong, H., Geng, Y., and Brander, M. (2018). An overview of China's recyclable waste recycling and recommendations for integrated solutions. *Res. Conserv. Recycl.* 134, 112–120. doi: 10.1016/j.resconrec.2018.02.032

Yuan, H. (2017). Barriers and countermeasures for managing construction and demolition waste: a case of Shenzhen in China. *J. Clean. Prod.* 157, 84–93. doi: 10.1016/j.jclepro.2017.04.137

Zhang, N., Zheng, L., and Duan, H. (2019). Differences of methods to quantify construction and demolition waste for less-developed but fast-growing countries: China as a case study. *Environ. Sci. Pollut. Res*. 26, 25513–25525. doi: 10.1007/s11356-019-05841-4