A method review of the climate change impact on crop yield

Xinyao Feng†, Haoliang Tian†, Jiahui Cong and Chuang Zhao*

College of Resources and Environmental Sciences, China Agricultural University, Beijing, China

Climate change significantly impacts global agricultural production, giving rise to considerable uncertainties. To explore these climate impacts, three independent methods have been employed: manipulated experiments, process-based crop models, and empirical statistical models. However, the uncertainty stemming from the use of different methods has received insufficient attention, and its implications remain unclear, necessitating a systematic review. In this study, we conducted a comprehensive review of numerous previous studies to summarize the historic development and current status of each method. Through a method comparison, we identified their respective strengths, limitations, and ideal areas of application. Additionally, we outlined potential prospects and suggested directions for future improvements, including clarifying the response mechanisms, updating simulation technologies, and developing multi-method ensembles. By addressing the knowledge gap regarding method differences, this review could contribute to a more accurate assessment of climate impacts on agriculture.

KEYWORDS
crop yield, climate change, research methods, manipulated experiment, process-based crop model, empirical statistical model

1. Introduction

The impact of global climate change on agricultural production in the 21st century has been significant, many countries and regions worldwide have observed reduced yields in crops such as wheat, maize, rice, and oilseed rape (Luo et al., 2005; Arora, 2019; Ray et al., 2019; Sultan et al., 2019; Ortiz-Bobea et al., 2021; Lachaud et al., 2022; Chandio et al., 2023). It is expected that temperatures will continue to rise, leading to an increase in extreme weather events. This trend adds to the agricultural production uncertainty, particularly for major crops such as maize, rice, and soybeans (Vogel et al., 2019; Pörtner et al., 2022). Without adaptation measures, global yields of important food crops could decline by 12–20% by the end of this century (Lobell and Gourdji, 2012; Wheeler and Von Braun, 2013; Challinor et al., 2014; Aggarwal et al., 2019). Therefore, accurately assessing the impact of climate change on crop yields is crucial for ensuring global food security.

There are several methods are used for climate change's effects on agriculture research, such as manipulated experiments, process-based crop models, and empirical statistical models. Field experiments was the earliest commonly used to expose crops to different climatic conditions, either through natural variations or controlled climate factors, to study their impact on crop growth and yield. With technological advancements, process-based crop models and empirical statistical models have become more prominent. Process-based crop models utilize computer simulations to quantitatively analyze the physiological mechanisms and dynamic processes of crop growth and yield. Empirical statistical models...
establish mathematical relationships between climate change and crop yield. Over time, significant progress have been made in developing these methodologies. For example, Shi et al. (2013) identified uncertainties in statistical models related to research scale, collinearity of variables, and detrending. Similarly, White et al. (2011) and Röther et al. (2018) evaluated existing process based crop models, assessing their simulation effects and research standards while highlighting sources of error and limitations. However, many current studies tend to focus on specific research methods, which may introduce biases and uncertainties into climate change impact studies. This limitation reduces the comprehensiveness and reliability of individual approaches. For example, empirical statistical models, based on limited historical observations, face challenges in accurately predicting future yield-climate relationships (Lobell et al., 2006). Conversely, process-based crop models rely on empirical formulas to approximate internal crop growth processes (Wang et al., 2022). Therefore, it is crucial to understand the advantages and disadvantages of each method and explore avenues for future improvement.

This paper presents a comprehensive review of recent advancements in research methods used to study the impacts of climate change on agriculture and adaptation strategies. Its primary aim is to provide researchers with a deeper understanding of existing methods and serve as a reference for future methodological innovations and interdisciplinary collaborations. To achieve this goal, the paper systematically examines three major methods: manipulated experiments, process-based crop models, and empirical statistical models. It critically evaluates the advantages, disadvantages, and directions for future improvement for each method. Additionally, the paper explores the interconnectedness of multiple methodological approaches and their relevance to current research. It also discusses the challenges associated with current methods and highlights potential future research prospects.

2. Manipulated experiments

Manipulated experiments involve setting different environmental conditions during crop growth to simulate the impact of climate change on crop yield. As an early research method, manipulated experiments has evolved from utilizing natural climate variations to artificial control. This method is simple and straightforward, with a high level of operability. Initially, artificial manipulated experiments were conducted in growth chambers or greenhouse, where temperature, light, water, and gas control experiments were carried out in adjustable but fully enclosed environments. Modern large-scale artificial greenhouse relies on facilities such as strip lights, removable platforms, and exhaust systems to achieve uniform distribution of meteorological factors such as light, temperature, and water. Real-time monitoring and precise control are achieved with the support of computers. Hatfield and Prueger (2015) set up warmer conditions in an artificial climate chamber and found maize yield was significantly reduced. Besides, temperature effects are increased by water deficits and excess.

However, artificial greenhouse block the exchange of water and gases between crops and the external environment, lacking the comprehensive effects of the natural environment. To avoid abnormal increases in humidity and temperature caused by fully enclosed environments, open-top chambers (OTCs) and open-air CO2 enrichment systems (Free-Air CO2 Enrichment, FACE) have been used. OTCs is a gas-enriched greenhouse without a cover on the top, this design ensures sufficient exchange of water and gases between crops and the external environment, significantly enhancing the ability to simulate realistic growth conditions. OTCs have been widely used in experiments investigating the effects of gases such as CO2 and O2 on crop growth (Rogers et al., 1994; Ziska et al., 1997; Ewert et al., 2002; Kakani et al., 2003; Ainsworth et al., 2012). However, there are still some differences in conditions such as wind speed and light within the OTCs compared to natural conditions. Over time, the outer film of the OTCs may undergo oxidation, yellowing, and dirt accumulation, leading to shading of solar radiation and affecting the experimental outcomes (Leadley and Drake, 1993).

To further enhance simulation realism, researchers have increasingly turned to FACE systems. FACE systems release CO2 or CO2-rich air from above the ground onto plant canopies and adjust CO2 flow rates through feedback mechanisms (Long et al., 2004), enabling studies on the effects of elevated CO2 and O3 concentrations on crops’ productivity (Long et al., 2006; Myers et al., 2014). Kimball et al. (2002) conducted experiments in different countries using the FACE system and found elevated CO2 increased crop yield substantially in C3 species, but little in C4. Combining an infrared heater with the FACE system to conduct experiments, known as T-FACE (Temperature-FACE), allows for the research of the combined effects of temperature and CO2 concentration on crops (Ruiz-Vera et al., 2013). Compared to OTC, FACE systems effectively reduce edge effects and cause minimal disturbance to farmland microclimates (Ainsworth et al., 2008). Nonetheless, the vertical gradient of CO2 concentration gradually decreases, which is also influenced by wind speed (Long et al., 2004). Currently, FACE systems are evolving toward genetic variation and transgenic technology research aimed at adapting agricultural planting systems to future climates (Ainsworth et al., 2020).

3. Process-based crop models

Process-based crop models are a type of models based on the internal physiological mechanisms of crops, which can comprehensively consider the relationships among the soil-crop-atmosphere system. They describe various physiological processes of crop growth as equations and incorporate various environmental factors (meteorological modules such as temperature, water, and light, as well as soil parameters, cultivars, and agronomic management). The earliest crop models can be traced back to the model on corn canopy photosynthetic rate developed by de Wit (1965). Subsequently, scientists from different countries have developed series of models, such as the DSSAT series (Jones et al., 2003), and the APSIM series (Keating et al., 2003), WOFOST (Van Diepen et al., 1989), CropSyst (Stöckle et al., 2003), SIMPLE (Zhao et al., 2019), etc. In recent years, through setting different climate inputs, crop models have been widely applied in studies on the impacts of climate change on crop yields, whether in long-term changes (Leng and Hall, 2019; Shahid et al., 2021) or in quantifying
extreme weather events (Xiao et al., 2022a; Júnior et al., 2023). Researchers have integrated and compared multiple crop models, consistently reached the conclusion that global warming has a significant negative impact on crop yields (Asseng et al., 2013; Bassu et al., 2014; Sultan et al., 2019; Zhao et al., 2022).

Process-based crop models based on site scale consider only the small-scale climate of individual locations. To incorporate large-scale regions, countries, or global scales, it is necessary to utilize crop models at a larger scale. One commonly approach is the grid-scale crop model. It involves dividing the geographic region into grids of specific resolutions and inputs gridded meteorological data, soil data, management data. For example, Deryng et al. (2011) used the PEGASUS model to simulate the response of major cereal crops to future climate change under different agronomic management. Rosenzweig et al. (2014) combined multiple global crop models to simulate the yield impact under historical climate conditions and found that the results were in good agreement with observed values, confirming the applicability of global grid-scale crop models. In recent years, the application of machine learning methods has made remarkable progress in regional or global-scale research, complementing traditional modeling methods through data-driven approaches (Reichstein et al., 2019). Studies have shown that machine learning-based crop modeling systems can accurately and rapidly predict crop yields in large regions at different spatial resolutions (Xiao et al., 2022b).

4. Empirical statistical models

Empirical statistical models is an approach used by establishing mathematical models that describe the relationship between climate factors and crop production. These models rely on both crop yield data and climate data to establish this correlation. Prior to developing empirical statistical models, it is necessary to separate the climate yield from trend yield and error terms. This is crucial because variations in crop yield are influenced not only by climate change but also by factors like technological advancements. Common methods employed to detrend the original yield data include differencing, multi-year moving averages, linear regression, and filtering analysis (Meza and Silva, 2009; Osborne and Wheeler, 2013; Troy et al., 2015; Kukal and Irmak, 2018). Additionally, considering the nonlinear relationship between economic factors and natural factors in grain production (Xu et al., 2021), econometric models have been introduced. These models include production functions (Just and Pope, 1978; Isik and Devadoss, 2006), economic-climate models (C-D-C models) (Chou and Ye, 2006), and the Ricardo model (Deressa and Hassan, 2009).

Empirical statistical methods have evolved from simple univariate linear regression models to more complex multivariate regression models, incorporating multiple influencing factors. The development has further advanced with the integration of machine learning and deep learning techniques, enabling the transition from univariate to multivariate and from linear to nonlinear modelling. The univariate linear regression model establishes a straightforward linear relationship between crop yield and a single climate factor (Parry and Martens, 1999). For example, Peng et al. (2004) constructed a simple univariate linear model using rice yield data and seasonal temperature data from observation stations in the Philippines. They found a significant negative correlation between rice yield and minimum temperature, with approximately a 10% yield reduction for each 1°C increase in minimum temperature. Univariate regression models can only consider the influence of a specific climate condition, such as temperature alone, and cannot account for all factors affecting yield (Carter et al., 1992). However, climate change is complex and often involves multiple simultaneous climate conditions impacting crop yield. Hence, the application of multivariate regression models has emerged. Multivariate regression models establish climate-yield correlation models with multiple climate conditions as independent variables and yield as the dependent variable. These models can be categorized based on temporal variations (time series models), spatial variations (cross-sectional models), or both (panel models) (Lobell and Burke, 2010). They have demonstrated good performance in simulating the impacts of climate change on crops like maize, wheat, and soybeans (Malikov et al., 2020; Ranjan et al., 2020). For instance, Lobell and Field (2007) established a multivariate linear regression model between temperature, precipitation, and yield, revealing a clear negative response of global wheat, maize, and barley yields to temperature increases. Schlenker and Lobell (2010) developed panel models and projected varying degrees of decline in crop yields such as maize, sorghum, and millet in Sub-Saharan Africa by the mid-century as climate change progresses.

While regression models excel at capturing linear relationships, they face limitations when it comes to simulating nonlinear relationships. To overcome this limitation, machine learning techniques have gained widespread use in climate change impact assessment (Cao et al., 2021; Guo et al., 2021; Lischeid et al., 2022). Machine learning approaches employ semi-parametric variables based on deep neural networks (Crane-Droesch, 2018) or decision systems like support vector machines and fuzzy logic for yield modelling (Palanivel and Surianarayanan, 2019). They leverage algorithms such as decision trees and random forests for prediction (Jeong et al., 2016). Machine learning demonstrates significant potential in yield assessment, surpassing traditional regression models. Unlike traditional empirical statistical models that rely on specific-shaped response functions, machine learning compensates for their limitations by effectively capturing complex nonlinear relationships in high-dimensional datasets. For example, Leng and Hall (2020) compared the performance of machine learning and multivariate regression models and found that machine learning explained 93% of the yield variability, significantly higher than the 51% explained by multivariate regression. Moreover, under a global warming scenario of 2°C, the maize yield in the United States is projected to decrease by 13.5% (Leng and Hall, 2020).

5. Methods comparison

To date, these methods have been widely used. However, due to differences in their underlying principles, these methods have both advantages and disadvantages in practical applications (Table 1).
TABLE 1 Advantages and disadvantages of each method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manipulated experiments</td>
<td>• Strong environment controllability</td>
<td>• Hard to upscale to regional level</td>
</tr>
<tr>
<td></td>
<td>• Interpretability of growth and development processes</td>
<td>• Limited to some specific years</td>
</tr>
<tr>
<td></td>
<td>• High reliability of results</td>
<td>• Time and money expense</td>
</tr>
<tr>
<td>Process-based crop models</td>
<td>• The physiological mechanism of crop growth included</td>
<td>• Parameter uncertainties</td>
</tr>
<tr>
<td></td>
<td>• High simulation efficiency</td>
<td>• Large input data requirements</td>
</tr>
<tr>
<td></td>
<td>• Strict control on climatic variables</td>
<td>• Empirical physiological processes included</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Lack processes of extreme climate impacts</td>
</tr>
<tr>
<td>Empirical statistical model</td>
<td>• Clear and simple equations between yield and climate</td>
<td>• Strong empirical assumptions</td>
</tr>
<tr>
<td></td>
<td>• No calibration process</td>
<td>• Lack of complex nonlinearity relationships collinearity problems between different factors</td>
</tr>
<tr>
<td></td>
<td>• Rich input data sources</td>
<td>• Limitation of future scenario extrapolations</td>
</tr>
</tbody>
</table>

5.1. Manipulated experiments

Using manipulated experiments to precisely regulate the thresholds of climate factors can simulate the actual environmental conditions of crop growth and development, resulting in high reliability of the obtained results. Therefore, the results obtained from controlled experiments are often used to calibrate crop models (Asseng et al., 2004). However, the experimental period of manipulated experiments is dependent on the crop’s growth cycle, and it involves complex operations, a long experimental cycle, and substantial human and material resources. Consequently, it is challenging to conduct long-term studies on future climate changes spanning several decades to centuries. Furthermore, due to the heterogeneity of the climate terrain and soil in the experimental area, the site-based experimental results have poor representativeness to larger areas. To upscale to the regional scale, a large amount of experimental data might be needed. A plausible solution is to establish a unified research framework. For example, Coordinated Distributed Experiments (CDE) offers the advantage of addressing global problems while ensuring the inherent accuracy of control experiments (Fraser et al., 2013).

5.2. Process-based crop models

Process-based crop models can effectively simulate the physical mechanisms of crop growth and strictly control the impact of single variables, avoiding the need for long-term field experiments. However, operating such models requires a significant amount of parameter calibration work, especially for large scale and long term simulations, and the process is complex. Within the model, crop growth processes and growing environments are approximately described using empirical or descriptive formulas, which results in some deviations in the response of crop physiological processes. Moreover, this uncertainty varies among different models (Wang et al., 2022).

5.3. Empirical statistical models

Empirical statistical models evaluate and predict crop yields by establishing a correlation between climate factors and historical yields, avoiding the complex tuning required by crop models and the need for inputting soil properties and management practices measured in the field. These models are relatively easy to apply and suitable for regional and global-scale studies. However, the climate factors used as input for these models are often monthly or seasonal averages, smoothing the impact of climate variability during the growing season and neglecting the effects of extreme weather events (Chen et al., 2004). Additionally, the accuracy of detrending methods used in empirical statistical models is difficult to evaluate. Moreover, empirical statistical models are based on limited historical observations, which introduce sampling uncertainty and simulation bias (Lobell et al., 2006). Because these models lack a reliable physical mechanism for extrapolation, they exhibit uncertainty in predicting yields beyond the historical range, and are more suitable for historical or near future studies.

6. Perspectives

6.1. Clarifying crop response mechanisms

The majority of existing yield prediction models operate as “one-way” mechanisms, failing to account for the fact that crops can alter their morphology and physiological functions to adapt to climate change. Moreover, current models are inadequate in assessing the chain reaction of crops to a series of weather events, their ability to adapt to the combined effects of multiple factors, and their response to the precursors of impact factors (Suzuki et al., 2014). In the future, more attention should be paid to the comprehensive influence of soil conditions, hydrological cycle, pest problems, and other factors in crop models (Newbery et al., 2016; Basso et al., 2018; Deutsch et al., 2018; Tomaz et al., 2020; Wei et al., 2021; Denissen et al., 2022). Additionally, the potential impact of climate change on crop production is not limited to the growth period, as environmental changes during non-growth periods can also indirectly affect crop production. Therefore, it is necessary to establish a model that comprehensively considers the basic knowledge of plant physiology and atmospheric science, including feedback mechanisms (Tonnang et al., 2022), in order to achieve a balance between the authenticity and controllability of the simulation.
6.2. Introducing emerging simulation technologies

The incorporation of emerging technologies can significantly enhance the research capabilities of traditional methods during the process of methodological development. For instance, the integration of remote sensing, big data, and artificial intelligence into existing approaches (Jiang et al., 2020) can address more complex data acquisition and processing requirements, enabling large-scale simulation and regulation. The utilization of remote sensing technology allows for weather and crop data with multiple spatial and temporal resolutions, enabling the assimilation of dynamic crop phenotype data provided by satellites to achieve closer real-time monitoring. This, in turn, enhances the ability of crop models to monitor and predict large-scale crop yields (Huang et al., 2019). However, the current stage of development of data assimilation technology and its application in this field is still in its early stages. As such, machine learning technologies may be key to enhancing the maturity of this approach (Cai et al., 2019).

6.3. Application of method ensemble

Many researchers have assembled multiple crop models to optimize simulation effectiveness and avoid systematic errors associated with a single model (Bassu et al., 2014; Martre et al., 2015), which have been shown to provide more reliable results (Asseng et al., 2013). However, integrating models does not fundamentally improve the underlying mechanisms, combining physiological principles and basic science can be essential (Yin et al., 2021). What is more remarkable is that combining multiple methods can effectively reduce uncertainty (Zhao et al., 2017). For instance, the integration of manipulated experiments with process-based crop models can supplement missing modules within existing models or establish more targeted new models by observing various physiological processes throughout the entire growth period of crops under particular conditions. In addition, developing a joint model that combines process-based crop models with statistical models (Roberts et al., 2017) can be beneficial. Such models use simple statistical models to summarize and statistically analyze simulation results generated by process-based crop models.

By using polynomials and limited weather variables, these models can accurately replicate process-based crop model results in global grid cells, avoiding the complex parameter adjustment process associated with process-based crop models while also predicting long-term trends (Blanc and Sultan, 2015).

Author contributions

CZ motivated the conception and review design. XF and HT wrote the draft of the manuscript. CZ and JC contributed to the manuscript revision. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Youth Foundation of China (42201032), the National Natural Science Fund for Excellent Young Scholars (Overseas), and the Fundamental Research Funds for the Central Universities (15053348).

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


Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol.* 274, 144–159. doi: 10.1016/j.agrformet.2019.03.010


