# MACGYVER IN GEOSCIENCES

EDITED BY: Rolf Hut, Theresa Blume and Peter M. Marchetto

PUBLISHED IN: Frontiers in Earth Science

frontiers Research Topics

**Frontiers eBook Copyright Statement**

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88963-715-7
DOI 10.3389/978-2-88963-715-7

## 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: researchtopics@frontiersin.org

# MACGYVER IN GEOSCIENCES

Topic Editors:
**Rolf Hut,** Delft University of Technology, Netherlands
**Theresa Blume,** German Research Centre for Geosciences, Germany
**Peter M. Marchetto,** University of Minnesota Twin Cities, United States

MacGyver science is the creative use of equipment for purposes that were not originally intended by the developer as well as the scientist's own development of sensors or technology for problems where commercially available solutions fall short.

Following the successful MacGyver conference sessions in the past years it is time to combine all our ideas, opinions and new research in an article collection. This is a call for papers for all MacGyver earth scientists– present your tools, processes, proof of concepts, designs, open source components, failures and successes, data sets, and emerging technologies, and contribute your part to this exciting collection.

Even if your new tools, prototypes or method has been described as part of the method section of a broader publication, we invite you to write a separate publication in our collection that focusses solely on the new tool, processes, proof of concepts, designs, open source components, etc.

**Citation:** Hut, R., Blume, T., Marchetto, P. M., eds. (2020). MacGyver in Geosciences. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-715-7

# Table of Contents


**04** ***Tree Sway Time Series of 7 Amazon Tree Species (July 2015–May 2016)***
Tim van Emmerik, Susan Steele-Dunne, Marceau Guerin, Pierre Gentine, Rafael Oliveira, Rolf Hut, John Selker, Jim Wagner and Nick van de Giesen

**09** ***Soda Bottle Science—Citizen Science Monsoon Precipitation Monitoring in Nepal***
Jeffrey C. Davids, Nischal Devkota, Anusha Pandey, Rajaram Prajapati, Brandon A. Ertis, Martine M. Rutten, Steve W. Lyon, Thom A. Bogaard and Nick van de Giesen

**29** ***Low-Cost, Open-Source, and Low-Power: But What to Do With the Data?***
Jeffery S. Horsburgh, Juan Caraballo, Maurier Ramírez, Anthony K. Aufdenkampe, David B. Arscott and Sara Geleskie Damiano

**43** ***Talking SMAAC: A New Tool to Measure Soil Respiration and Microbial Activity***
Ayush Joshi Gyawali, Brandon J. Lester and Ryan D. Stewart

**51** ***OPEnS Hub: Real-Time Data Logging, Connecting Field Sensors to Google Sheets***
Thomas DeBell, Luke Goertzen, Lars Larson, William Selbie, John Selker and Chet Udell

**57** ***Mojito, Anyone? An Exploration of Low-Tech Plant Water Extraction Methods for Isotopic Analysis Using Locally-Sourced Materials***
Benjamin M. C. Fischer, Jay Frentress, Stefano Manzoni, Sara A. O. Cousins, Gustaf Hugelius, Maria Greger, Rienk H. Smittenberg and Steve W. Lyon

**68** ***Mobile Monitoring—Open-Source Based Optical Sensor System for Service-Oriented Turbidity and Dissolved Organic Matter Monitoring***
Robert Schima, Stephan Krüger, Jan Bumberger, Mathias Paschen, Peter Dietrich and Tobias Goblirsch

**83** ***Particle Seeded Grains to Identify Highly Irregular Solid Boundaries and Simplify PIV Measurements***
William Basham, Ralph Budwig and Daniele Tonina

**92** ***Low-Cost Environmental Sensor Networks: Recent Advances and Future Directions***
Feng Mao, Kieran Khamis, Stefan Krause, Julian Clark and David M. Hannah

**99** ***Assessing the Sampling Quality of a Low-Tech Low-Budget Volume-Based Rainfall Sampler for Stable Isotope Analysis***
Benjamin M. C. Fischer, Franziska Aemisegger, Pascal Graf, Harald Sodemann and Jan Seibert

**107** ***Monitoring Atmospheric, Soil, and Dissolved $CO_2$ Using a Low-Cost, Arduino Monitoring Platform (CO2-LAMP): Theory, Fabrication, and Operation***
Joshua M. Blackstock, Matthew D. Covington, Matija Perne and Joseph M. Myre

**126** ***High-Resolution Bathymetry Mapping of Water Bodies: Development and Implementation***
Liah X. Coggins and Anas Ghadouani

**137** ***A User-Printable Three-Rate Rain Gauge Calibration System***
Jose M. Lopez Alcala, Chester J. Udell and John S. Selker

frontiers in Earth Science

DATA REPORT
published: 07 December 2018
doi: 10.3389/feart.2018.00221

# Tree Sway Time Series of 7 Amazon Tree Species (July 2015–May 2016)

*Tim van Emmerik[1]*, Susan Steele-Dunne[1], Marceau Guerin[2], Pierre Gentine[2], Rafael Oliveira[3], Rolf Hut[1], John Selker[4], Jim Wagner[5] and Nick van de Giesen[1]*

[1] Water Resources Section, Delft University of Technology, Delft, Netherlands, [2] Department of Earth and Environmental Engineering, Columbia University, New York, NY, United States, [3] Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, Brazil, [4] Department of Biological and Ecological Engineering, Oregon State University, Corvallis, OR, United States, [5] Oregon Research Electronics, Tangent, OR, United States

**Keywords: tree physiology, acceleration, Amazon, drag coefficient, turbulence, interception, water stress**

OPEN ACCESS

***Edited by:***
*Steven V. Weijs,*
*University of British Columbia, Canada*

***Reviewed by:***
*Joshua B. Fisher,*
*NASA Jet Propulsion Laboratory (JPL), United States*
*Ahmed M. ElKenawy,*
*Mansoura University, Egypt*

****Correspondence:***
*Tim van Emmerik*
*t.h.m.vanemmerik@tudelft.nl*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *30 June 2018*
***Accepted:*** *16 November 2018*
***Published:*** *07 December 2018*

***Citation:***
*van Emmerik T, Steele-Dunne S, Guerin M, Gentine P, Oliveira R, Hut R, Selker J, Wagner J and van de Giesen N (2018) Tree Sway Time Series of 7 Amazon Tree Species (July 2015–May 2016). Front. Earth Sci. 6:221. doi: 10.3389/feart.2018.00221*

## 1. INTRODUCTION

Trees are a crucial part of ecosystems through their important influence on the water and carbon cycles (Reichstein et al., 2013; Schlesinger and Jasechko, 2014; Patton et al., 2016). They play a key role in hydrological and ecological systems, and land-surface interactions. Ground measurements of tree properties and tree-related fluxes can quantify, for example, tree mass variations, $CO_2$ uptake, transpiration, rainfall interception and canopy drag, which are crucial parameters for understanding and modeling tree behavior and their roles in ecosystems. Unfortunately, tree measurement sensors are often based on invasive techniques (e.g. sap flow sensors or dendrometers), unable to withstand challenging field conditions (e.g. weathering or power/electronical failure due to climate), or cannot measure all parameters of interest.

Accelerometers offer an economic and robust way for obtaining data series of tree motion, which has been used to infer tree properties and tree-related fluxes. Mounted on a tree trunk, they measure tree sway, which is determined by its mass, elasticity, wood density and drag coefficient. These parameters can in turn be related to water-related mass changes (transpiration and water uptake) (Llamas et al., 2013), biomass-related mass changes (growth, leaf fall and leaf flush) (Selker et al., 2011), tree-atmosphere interactions (drag coefficient and momentum transfer) (van Emmerik et al., 2018a). Additionally, tree sway has shown to be related to tree throw failure mechanisms, and thereby help to reduce storm risks of trees and forests (Flesch and Wilson, 1999a,b; Wilson and Flesch, 1999), by assessing the influence of forest cutblocks.

The relation between tree sway and tree properties and tree-related fluxes can be derived from (1) Newton's second law, and (2) a momentum balance. First, it is often assumed that trees behave like damped harmonic oscillators (Gardiner, 1995; Peltola, 1996). In that case, the tree's natural frequency is linked to mass and elasticity through:

$$\omega_0 = 2\pi f_0 = \sqrt{\frac{k}{m}} \tag{1}$$

With natural frequency in [rad/s] $\omega_0$, natural frequency in [Hz] $f_0$, mass $m$ and elasticity $k$. From the frequency spectrum of tree acceleration the natural frequency can be determined. Second, following the momentum balance the frequency spectrum of tree sway $P_y(f)$ as a function of frequency can be expressed as (Amtmann, 1985; Mayer, 1987):

$$P_y(f) = H_m(f)^2 \rho_a^2 C_d^2 A_2 u^2 H_a(f) P_u(f) \tag{2}$$

With mechanical transfer function $H_m(f)$, air density $\rho_a$, drag coefficient $C_d$, wind speed $u$, aerodynamic transfer function $H_a(f)$ and wind load frequency spectrum $P_u(f)$.

These relations have been used in previous tree-related research. Moore and Maguire (2004) showed that factors such as branch removal and snow loading influences tree natural frequency. Mayer (1987) used tree sway data to demonstrate that the primary sway (i.e., the system as a whole) is related to tree throw. Flesch and Wilson (1999a,b) and Wilson and Flesch (1999) used tilt sensors to analyze tree sway, and assess the influence of management techniques on the possible reduction on wind throw. Later, Moore and Maguire (2005) empirically found a relationship between tree height and tree natural frequency. These, and many more, studies already demonstrated the use of tree sway data. In particular the derived estimation of tree natural frequency is useful for risk assessment of wind damage/throw, and quantification of tree properties (e.g., mass, elasticity) and responses (e.g., intercepted precipitation, mass changed in response to water stress).

Data in previous studies were mainly obtained for several hours, or in rare cases days. Our dataset presents a unique opportunity to use long-term tree sway data of 19 individual trees. This allows analysis of diurnal and seasonal changes, as well as comparison between individuals of the same species, and between different species as a whole. In this paper we describe the details of the obtained dataset, and discuss current and potential applications of the dataset.

## 2. DATA COLLECTION

### 2.1. Measurement Location

All data were collected around the K34 observatory (130 m above Mean Sea Level) in the Amazon rainforest (2.6085°S, 60.2093°W), 60km Northwest of Manaus, Brazil (see van Emmerik et al., 2017). A map of the measurement location, including the measured trees and the K34 observatory is presented in **Figure 1A**. The study area is characterized by a wet tropical climate. Additional meteorological, hydrological and plant physiological data may be obtained from the K34 observatory, managed by the National Institute of Amazonian Research (INPA).

### 2.2. Plant Material

Accelerometers were mounted on 19 individual trees, covering seven different species with one to four individuals per species. The trees were selected to cover a broad range of estimated wood density, tree height, and diameter at breast height (DBH). **Table 1** presents an overview of the measured trees. Tree species were classified by a taxonomist on site. Wood density was estimated using the Global Wood Density Database (Chave et al., 2009; Zanne et al., 2009). Tree height was measured using measurement tape during installation of the accelerometers.

### 2.3. Sensor Description

All accelerometers used were of the type Acceleration Logger Model AL100 (Oregon Research Electronics, Tangent, OR, USA), which is specifically designed to be physically robust and water proof. Robustness is achieved through a casing that prevents water and organisms to reach the sensor. The sensor has internal batteries and data storage, so no (wire) connections are required during measurements. Consequently, its dimensions (14.5 × 9.2 × 5.5 cm, 0.4 kg) are larger than other available accelerometers. The accelerometers can measure acceleration in three axis with a frequency up to 25 Hz. Depending on the sampling rate and environmental conditions, it can log for several months on two C-size cell batteries. Data are stored on regular microSD cards. For example, with a 8GB data card it can store up to 320 days of data measured with a frequency of 10 Hz. Data are written to a newly created file each day to minimize data loss in case of empty batteries, full data storage, or other potential failures. The casing has space to include several silica bags to absorb moisture when installed in humid environments.

### 2.4. Measurement Setup

The accelerometers were placed on the tree trunk, right below the point of main branching. This guaranteed the largest excitation of the sensor, and therefore largest acceleration values, that were still representative for the whole tree. Installing accelerometers on the primary of secondary branches comes with the risk of mainly measuring higher order effects. The accelerometers were mounted using a spring around the trunk. This allows some expansion of the tree when growing, and also assures a rigid attachment to the tree. All measurements were done with a sampling frequency of 10 Hz.

## 3. DATA DESCRIPTION

The available dataset includes raw acceleration values for all three dimensions as measured by the sensor with a frequency of 10 Hz. In **Table 1** it can be found which accelerometer (serial numbers SN1001-SN1019) was mounted on which specific tree. For each sensor, data are available per day. This reduces computational needs, as each daily file can be read and processed individually.

As most applications require transformation of the data from the time to frequency domain, an additional dataset of the acceleration data in the frequency domain can be downloaded. For each sensor, files are available that contain time series of the acceleration in the frequency domain. These were produced by applying a Fast Fourier Transform (FFT) on 30 min of horizontal (y-axis) acceleration with a moving window of 10 min. **Figure 1B**. illustrates a typical frequency spectrum estimated using tree acceleration. It can be seen that the natural frequency around 0.2 Hz can be estimated well. In this example, even higher order resonance peaks are visible.

When applying the FFT to the whole dataset, a time series of the frequency spectrum is created **Figures 1C–E**. show a typical time series of the frequency spectrum for two trees. It can be seen that when wind speed is higher than a certain threshold, the resonance peaks are activated. It can also be seen that there are slight differences in timing and magnitude of the frequency spectra, indicating that tree specific information can be deduced. Note that this figure only shows a subset of the total available data. The complete measurement period stretches from July 2015 to May 2016.

For all applications so far (van Emmerik et al., 2017, 2018a), data have been processed using a FFT. We show that using a simple FFT the natural frequency can already be determined.

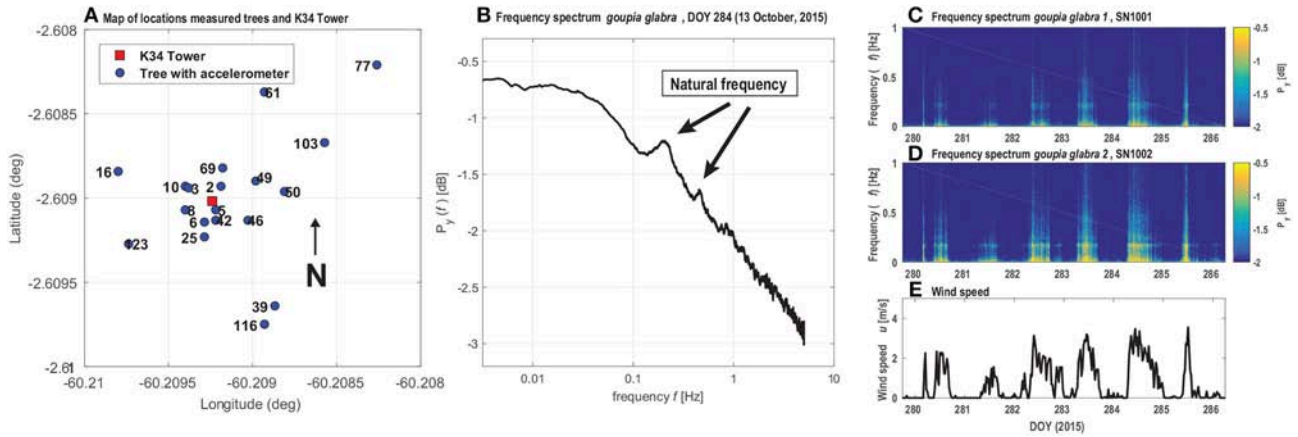



**FIGURE 1 | (A)** Map with the locations of the measured trees, includin gthe K34 tower, **(B)** Example of frequency spectrum estimated from acceleration data, **(C,D)** example of frequency spectrum time series of two trees in combination with **(E)** wind speed.

**TABLE 1 |** List of measured tree individuals, including acceleration sensor number, estimated wood density, estimated tree height and diameter at breast height (DBH).

| Accelerometer | Species | Estimated wood density [$10^3$ kg/m$^3$] | High - Low | Estimated height [m] | DBH [cm] |
|---|---|---|---|---|---|
| SN1001 | *Goupia glabra* | 0.7 | High | 25 | 242.5 |
| SN1002 | *Goupia glabra* | 0.7 | High | 32 | 208 |
| SN1003 | *Goupia glabra* | 0.7 | High | 32 | 135 |
| SN1004 | *Lecythis prancei* | 0.875 | Intermediate | 24 | 112 |
| SN1005 | *Lecythis prancei* | 0.875 | Intermediate | 35 | 116.5 |
| SN1006 | *Lecythis prancei* | 0.875 | Intermediate | 24 | 108.4 |
| SN1007 | *Scleronema micranthum* | 0.5 - 0.7 | Low | 38 | 127 |
| SN1008 | *Scleronema micranthum* | 0.5 - 0.7 | Low | 28 | 189.5 |
| SN1019 | *Scleronema micranthum* | 0.5 - 0.7 | Low | 26 | 81 |
| SN1010 | *Eschweilera* sp. | 0.8 | Intermediate | 20 | 93.5 |
| SN1009 | *Eschweilera coriacea* | 0.8 | Intermediate | 21 | 268 |
| SN1011 | *Eschweilera coriacea* | 0.8 | Intermediate | 18 | 92.4 |
| SN1012 | *Eschweilera coriacea* | 0.8 | Intermediate | 27 | 213.5 |
| SN1013 | *Dipteryx odorata* | 1.1 | High | 35 | 177 |
| SN1014 | *Dipteryx odorata* | 1.1 | High | 32 | 219.5 |
| SN1015 | *Pouteria anomala* | 0.7 | Low | 22 | 111 |
| SN1016 | *Maquira sclerophylla* | 0.5 | Low | 30 | 264 |
| SN1017 | *Maquira scherophylla* | 0.5 | Low | 18 | 201 |
| SN1018 | *Maquira scherophylla* | 0.5 | Low | 35 | 90.6 |
| SN1019 | *Pouteria* sp. | 0.7 | Low | 23 | 117.5 |

However, we also encourage further exploration with other spectrum estimation methods, such as Welch's method or the MUSIC algorithm.

## 3.1. Data Availability

All acceleration data are available on the 4.TU Repository (doi: 10.4121/uuid:c9974180-aa9b-40b4-8dbb06d5b1fce693) (van Emmerik et al., 2018b). All files are named "/YYYY/SN10XX/ALOGzzz.csv," with year YYYY, sensor number XX and day after start of logging zzz. Data are available from July 2015 to May 2016. We invite the community to explore the dataset and combine it with for example higher resolution wind data, hydro-meteorological measurements, tree physiological data and remote sensing observations to explore potential additional applications of long-term tree acceleration data.

# 4. EXAMPLES OF APPLICATIONS

## 4.1. Intercepted Precipitation

During rainfall events precipitation is intercepted by leaves of a tree. This increases the total tree mass. Selker et al. (2011) compared the natural frequency of a tree before and after a rainfall event and found a shift in frequency. A more detailed study by van Emmerik et al. (2017) found different relations between intercepted rainfall and shift in frequency of trees. For some trees, an almost linear relation was found between the

frequency shift and precipitation. For other trees, the relation was less obvious. This was hypothesized to be caused by (1) the location of the tree in the canopy (understory trees are less likely to intercept rainfall) and (2) by splashing of water drops during high intensity rainfall events. Accelerometers were demonstrated to be a promising new method for interception measurements. Current measurement techniques mainly focus on measuring throughfall, which requires a more extensive measurement setup with multiple rainfall sensors. However, in future measurements additional throughfall measurements might give additional insights in the accuracy of interception estimations using accelerometers. Moore and Maguire (2005) also demonstrated that tree sway measurements can be used to monitor snow cover of trees. This is not applicable to the dataset presented in this paper, but might be of interest to researcher who focus on snowy ecosystems.

### 4.2. Tree Mass Variations

Besides through rainfall and snow cover, tree mass also varies through other mechanisms, such as tree growth, changes in tree water content, or leaf fall and flush. A first analysis of the data presented in this paper, van Emmerik et al. (2017) demonstrated that, for most tree species, a relation can be found between aboveground tree biomass and natural frequency. Selker et al. (2011) presented a proof of concept that there is a difference in tree natural frequency when comparing the same tree with and without leaves. More recently, Jackson et al. (in review) demonstrated that there is a significant difference between trees during summer (with leaves) and winter (no leaves). van Emmerik (2017) and van Emmerik et al. (2018a) related changes in tree sway to water stress induced mass changes, which was hypothesized to be caused by decreasing tree water content and/or leaf fall. Future efforts might focus on using tree sway data to quantify tree mass changes on even lower time scales. With the current dataset, it might also possible to track diurnal changes in natural frequency, which could be related to changes in water content. Additional measurements of leaf water potential would allow further study of the relation between tree sway and tree mass variations.

### 4.3. Drag Coefficient and Roughness Length

Tree drag coefficient is a crucial parameter for the momentum transfer from the atmosphere to the trees. Accurate estimates of tree and canopy drag coefficient might allow better representation of the canopy in, for example, land-surface models. An approximation of tree drag coefficient was presented by van Emmerik (2017) and van Emmerik et al. (2018a), based on the assumption that during turbulent conditions, available wind energy can be estimated using Kolmogorov's theory (Kolmogorov, 1991). If actual high frequency wind speed data is available, drag coefficients can be determined under all wind conditions using equation 2. Here it was demonstrated that aggregating tree sway data over different time scales (weekly, monthly, complete monitoring period), can give different degrees of insights. When aggregated over the complete monitoring period, a clear relation between tree sway and tree mass was found. The weekly analysis in turn demonstrated shorter term mass variations in the individual trees. A similar approach might also allow for estimating the roughness length of a forest canopy if data from multiple sensors are combined, as the variation and evolution of individual tree drag and tree-atmosphere momentum transfer can be quantified.

### 4.4. Tree Damage

Using relatively short data records, it was already shown that tree sway is related to wind throw (Mayer, 1987). Longer-term records such as those presented in this paper might allow further analysis of tree failing mechanisms. Additionally, relatively unmeasured phenomena such as tree dormancy and mortality might be further investigated.

## 5. LIMITATIONS

Besides the many potential applications of tree sway data, fundamental investigations of the role of meteorological variables (e.g., temperature, humidity, wind speed and direction, precipitation) and tree properties (e.g., mass, elasticity) on tree sway is also crucial. Many applications assume tree sway can be described as mass-spring system, or through the momentum balance in Equation (2). However, a better fundamental understanding of tree sway might open doors to even more potential applications.

As applications of long-term tree sway data are still limited, further assessments are still very dependent on the availability of other data. With the current auxiliary data we hypothesized several relations between tree sway and other processes. However, more detailed data on e.g., leaf potential, wind speed, water uptake and intercepted precipitation are required to further test our hypotheses.

One of the key assumptions of the tree selection is that the chosen trees are free standing. If they are in contact with other trees, the mechanical system is more complex and the standard approach might not be applicable. The trees in this study were selected based on visual inspection, which might have been inaccurate. Also growth of the trees might have altered the position in respect to other trees.

The optimal installation location might also be different for other tree species. The selected trees were all relavtively straight, until the point of branching off. For tree species with more complex architecture and geometry, the optimal location should be reassessed.

The presented dataset only includes tree species in the Brazilian Amazon. For tree species in other climatic regions, different relations might be found between tree sway and e.g., mass (variations), tree-atmosphere interactions and intercepted precipitation. We encourage deployment in other climatic regions to further explore the possibilities of tree sway measurements.

## 6. CONCLUDING REMARKS

Tree acceleration data have been proven to give new insights in assessing risk of tree failure, intercepted rainfall, mass variations, and tree-atmosphere interactions. To date, only relatively short datasets have been used and published (hours

to days). Our dataset covers tree acceleration for a period of 10 months, including changing seasons, for 19 Amazon trees. We demonstrate that these data offer new opportunities of analyzing tree properties, and tree behavior in response to changing environmental conditions. We demonstrate various current applications. Yet, many potential new applications are to be explored. We encourage the scientific community to use our data to do so.

## AUTHOR CONTRIBUTIONS

TvE, SS-D, MG, and RH designed the study, TvE, MG, and RO installed the sensors, all authors contributed to data analysis and writing this manuscript.

## FUNDING

This work was funded by FAPESP GOAmazon project 2013/50431-2 and Vidi Grant 14126 from the Dutch Technology Foundation STW, which is part of The Netherlands Organisation for Scientific Research (NWO).

## ACKNOWLEDGMENTS

We thank Ricardo Dal'Agnol and Omar Chaparro Saaveedra for providing the GPS data of the measured trees. We also thank the reviewers for their constructive comments.

## REFERENCES

Amtmann, R. (1985). "Data acquisition system for wind induced tree vibration," in *The Forest-Atmosphere Interaction*, eds B. A. Hutchison and B. B. Hicks (Dordrecht: Springer), 149–159.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., and Zanne, A. E. (2009). Towards a worldwide wood economics spectrum. *Ecol. Lett.* 12, 351–366. doi: 10.1111/j.1461-0248.2009.01285.x

Flesch, T. K., and Wilson, J. D. (1999a). Wind and remnant tree sway in forest cutblocks.: I. Measured winds in experimental cutblocks. *Agricult. For. Meteorol.* 93, 229–242. doi: 10.1016/S0168-1923(98)00112-9

Flesch, T. K., and Wilson, J. D. (1999b). Wind and remnant tree sway in forest cutblocks. II. Relating measured tree sway to wind statistics. *Agricult. For. Meteorol.* 93, 243–258. doi: 10.1016/S0168-1923(98)00113-0

Gardiner, B. (1995). *The Interactions of Wind and Tree Movement in Forest Canopies*. Cambridge: Cambridge University Press.

Kolmogorov, A. N. (1991). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proc. Math. Phys. Sci.* 434, 9–13. doi: 10.1098/rspa.1991.0075

Llamas, R., Niemeier, J., Kruger, A., Lintz, H., Kleinknecht, G., and Miller, R. (2013). "Diurnal cycles of tree mass obtained using accelerometers," in *AGU Fall Meeting Abstracts*, Vol. 1 (San Francisco, CA) 1575.

Mayer, H. (1987). Wind-induced tree sways. *Trees* 1, 195–206. doi: 10.1007/BF01816816

Moore, J. R., and Maguire, D. A. (2004). Natural sway frequencies and damping ratios of trees: concepts, review and synthesis of previous studies. *Trees* 18, 195–203. doi: 10.1007/s00468-003-0295-6

Moore, J. R., and Maguire, D. A. (2005). Natural sway frequencies and damping ratios of trees: influence of crown structure. *Trees* 19, 363–373. doi: 10.1007/s00468-004-0387-y

Patton, E. G., Sullivan, P. P., Shaw, R. H., Finnigan, J. J., and Weil, J. C. (2016). Atmospheric stability influences on coupled boundary layer and canopy turbulence. *J. Atmospher. Sci.* 73, 1621–1647. doi: 10.1175/JAS-D-15-0068.1

Peltola, H. (1996). Swaying of trees in response to wind and thinning in a stand of scots pine. *Boundary-Layer Meteorol.* 77, 285–304. doi: 10.1007/BF00123529

Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., et al. (2013). Climate extremes and the carbon cycle. *Nature* 500, 287–295. doi: 10.1038/nature12350

Schlesinger, W. H., and Jasechko, S. (2014). Transpiration in the global water cycle. *Agricult. For. Meteorol.* 189, 115–117. doi: 10.1016/j.agrformet.2014.01.011

Selker, J., Lane, J., Rupp, D., Hut, R., Abou Najm, M., Stewart, R., et al. (2011). "The answer is blowing in the wind: using wind induced resonance of trees to measure time varying canopy mass, including interception," in *AGU Fall Meeting Abstracts*, Vol. 1 (San Francisco, CA) 1155.

van Emmerik, T. (2017). *Water Stress Detection Using Radar*. PhD thesis, Delft University of Technology.

van Emmerik, T., Steele-Dunne, S., Gentine, P., Oliveira, R. S., Bittencourt, P., Barros, F., et al. (2018a). Ideas and perspectives: tree–atmosphere interaction responds to water-related stem variations. *Biogeosciences* 15, 6439–6449. doi: 10.5194/bg-15-6439-2018

van Emmerik, T., Steele-Dunne, S., Guerin, M., Gentine, P., Oliveira, R. S., Hut, R., et al. (2018b). Tree sway of 19 Amazon trees. Available online at: https://doi.org/10.4121/uuid:c9974180-aa9b-40b4-8dbb-06d5b1fce693 (Accessed June 30, 2018).

van Emmerik, T., Steele-Dunne, S., Hut, R., Gentine, P., Guerin, M., Oliveira, R. S., et al. (2017). Measuring tree properties and responses using low-cost accelerometers. *Sensors* 17:1098. doi: 10.3390/s17051098

Wilson, J. D., and Flesch, T. K. (1999). Wind and remnant tree sway in forest cutblocks. III. A windflow model to diagnose spatial variation. *Agricult. For. Meteorol.* 93, 259–282. doi: 10.1016/S0168-1923(98)00121-X

Zanne, A., Lopez-Gonzalez, G., Coomes, D., Ilic, J., Jansen, S., Lewis, S., et al. (2009). *Global Wood Density Database. Dryad.* Available online at: http://hdl.handle.net/10255/dryad235

**Conflict of Interest Statement:** 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.

*Copyright © 2018 van Emmerik, Steele-Dunne, Guerin, Gentine, Oliveira, Hut, Selker, Wagner and van de Giesen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 22 March 2019
doi: 10.3389/feart.2019.00046

# Soda Bottle Science—Citizen Science Monsoon Precipitation Monitoring in Nepal

*Jeffrey C. Davids[1,2]*, Nischal Devkota[3], Anusha Pandey[3], Rajaram Prajapati[3], Brandon A. Ertis[2], Martine M. Rutten[4], Steve W. Lyon[5], Thom A. Bogaard[1] and Nick van de Giesen[1]*

[1] Water Management, Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands, [2] SmartPhones4Water, Chico, CA, United States, [3] SmartPhones4Water-Nepal, Lalitpur, Nepal, [4] Water Management, Institute for the Built Environment, Rotterdam University of Applied Sciences, Rotterdam, Netherlands, [5] The Nature Conservancy, Southern New Jersey Office, Delmont, NJ, United States

OPEN ACCESS

**Edited by:**
Theresa Blume,
German Research Centre for Geosciences, Helmholtz Centre Potsdam, Germany

**Reviewed by:**
Valentijn Pauwels,
Monash University, Australia
Ilja H. J. van Meerveld,
University of Zurich, Switzerland

***Correspondence:**
Jeffrey C. Davids
j.c.davids@tudelft.nl

**Specialty section:**
This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science

**Received:** 06 December 2018
**Accepted:** 27 February 2019
**Published:** 22 March 2019

**Citation:**
Davids JC, Devkota N, Pandey A, Prajapati R, Ertis BA, Rutten MM, Lyon SW, Bogaard TA and van de Giesen N (2019) Soda Bottle Science—Citizen Science Monsoon Precipitation Monitoring in Nepal. Front. Earth Sci. 7:46. doi: 10.3389/feart.2019.00046

Citizen science, as a complement to ground-based and remotely-sensed precipitation measurements, is a promising approach for improving precipitation observations. During the 2018 monsoon (May to September), SmartPhones4Water (S4W) Nepal—a young researcher-led water monitoring network—partnered with 154 citizen scientists to generate 6,656 precipitation measurements in Nepal with low-cost (<1 USD) S4W gauges constructed from repurposed soda bottles, concrete, and rulers. Measurements were recorded with Android-based smartphones using Open Data Kit Collect and included GPS-generated coordinates, observation date and time, photographs, and observer-reported readings. A year-long S4W gauge intercomparison revealed a −2.9% error compared to the standard 203 mm (8-inch) gauge used by the Department of Hydrology and Meteorology (DHM), Nepal. We analyzed three sources of S4W gauge errors: evaporation, concrete soaking, and condensation, which were 0.5 mm day$^{-1}$ ($n$ = 33), 0.8 mm ($n$ = 99), and 0.3 mm ($n$ = 49), respectively. We recruited citizen scientists by leveraging personal relationships, outreach programs at schools/colleges, social media, and random site visits. We motivated ongoing participation with personal follow-ups via SMS, phone, and site visit; bulk SMS; educational workshops; opportunities to use data; lucky draws; certificates of involvement; and in certain cases, payment. The average citizen scientist took 42 measurements (min = 1, max = 148, stdev = 39). Paid citizen scientists ($n$ = 37) took significantly more measurements per week (i.e., 54) than volunteers (i.e., 39; alpha level = 0.01). By comparing actual values (determined by photographs) with citizen science observations, we identified three categories of observational errors ($n$ = 592; 9% of total measurements): unit ($n$ = 50; 8% of errors; readings in centimeters instead of millimeters); meniscus ($n$ = 346; 58% of errors; readings of capillary rise), and unknown ($n$ = 196; 33% of errors). A cost per observation

analysis revealed that measurements could be performed for as little as 0.07 and 0.30 USD for volunteers and paid citizen scientists, respectively. Our results confirm that citizen science precipitation monitoring with low-cost gauges can help fill precipitation data gaps in Nepal and other data scarce regions.

**Keywords: citizen science (CS), recruitment, motivation, performance, low-cost rain gauge, smartphones, open data kit (ODK), cost per observation**

## INTRODUCTION

Precipitation is the main terrestrial input of the global water cycle; without it, our springs, streams, lakes, and communities would gradually disappear. Understanding the spatial and temporal distribution of precipitation is therefore critical for characterizing water and energy balances, water resources planning, irrigation management, flood forecasting, and several other resource management and planning activities (Lettenmaier et al., 2017). However, observing, and moreover understanding, precipitation variability over space and time is fraught with difficulty and uncertainty. Because of these challenges, there are persistent, but spatially heterogeneous, precipitation data gaps that need to be addressed (Kidd et al., 2017).

Accuracy is a primary concern, even for common precipitation measurement methods (Krajewski et al., 2003; Villarini et al., 2008) including: manual and automatic gauges, radar, and satellite remote sensing. Manual and automatic gauges are expensive to maintain and thus generally do not lead to adequate spatial representations of precipitation (e.g., Volkmann et al., 2010). For example, the total area of all the rain gauges in the world is less than half a football field (Kidd et al., 2017), or 0.000000002% of the global terrestrial landscape. Precipitation radars can provide meaningful data between gauges, but are subject to errors from beam blockage, range effects, and imperfect relationships between rainfall and backscatter (Kidd et al., 2017). Additionally, radars are expensive and operate by line of sight, so spatial cover of radar in mountainous terrains like Nepal can be limited. Satellite remotely sensed precipitation products have the benefit of global coverage, but can be impacted by random errors and bias (e.g., Koutsouris et al., 2016) arising from the indirect linkage between the observed parameters and precipitation and imperfect algorithms (Sun et al., 2018). Clearly, there remain precipitation data gaps and uncertainties that need to be filled.

Low-cost sensors and consumer electronics can play a role in closing these data gaps (Hut, 2013; Tauro et al., 2018). In general, the potential of low-cost sensors to improve understanding of a process depends on the interplay between (1) the spatial heterogeneity of the process being observed, (2) the impacts on accuracy of the low-cost sensor, and (3) the observational cost savings. The need for higher density observations increases as the spatial heterogeneity of the process being observed increases. So, if (1) the observed process has high spatial heterogeneity, and (2) the low-cost sensor provides high accuracy, with (3) high cost savings, the potential of the low-cost sensor to improve understanding of the process is considered high.

Citizen science has emerged as a promising tool to help fill data gaps. At the same time, citizen science can improve overall scientific literacy and reconnect people with their natural resources. McKinley et al. (2017) define citizen science as "the practice of engaging the public in a scientific project." They go on to clarify that crowdsourcing is another way for public participation in science through "… large numbers of people processing and analyzing data." Notable examples of citizen science precipitation monitoring include: the Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS: www.cocorahs.org); Weather Underground (www.wunderground.com); Met Office Weather Observation Website (WOW: wow.metoffice.gov.uk/); UK Citizen Rainfall Network (Illingworth et al., 2014); the NOAA Citizen Weather Observer Program (CWOP: wxqa.com); and an internet-connected amateur weather station network called Netatmo (www.netatmo.com) (Kidd et al., 2017).

Launched in the spring of 1998 by the Colorado Climate Center at Colorado State University, CoCoRaHS is a volunteer-led precipitation monitoring effort (Reges et al., 2016). Volunteers measure daily precipitation with a standardized 102 mm (4-inch) gauge (Sevruk and Klemm, 1989) and report their data via an online system. While CoCoRaHS was established in response to small scale flash floods, it has grown into the world's largest volunteer precipitation monitoring network, with over 20,000 active observers in the United States, Canada, the U.S. Virgin Islands, the Bahamas, and Puerto Rico (Cifelli et al., 2005).

In Nepal, three specific attempts have been made to launch citizen science precipitation measurement campaigns. The first was a single year effort in 1998 initiated by Nepali scientists Ajaya Dixit and Dipak Gyawali who partnered with community members to measure rainfall in the Rohini River watershed, a tributary to the Ganges, in south-central Nepal. The second was launched by Recham Consulting in 2003, and included 17 gauges similar to US National Weather Service 203 mm (8-inch) gauges in the Kathmandu Valley. However, the project stalled after only a few years of data collection. The third, Community Based Rainfall Measurement Nepal (CORAM-Nepal), was launched in 2015 with seven high schools in the Kathmandu Valley (Pokharel et al., 2016). CORAM's approach to obtain rainfall data is to partner with local high school science teachers and students, but other community members were also welcome to participate. CORAM-Nepal uses standard 102 mm (4-inch) CoCoRaHS gauges and collects data from schools monthly by phone call or site visits. All of these previous efforts grappled with the challenges of sustainable (1) funding, (2) human resources, and

(3) technological issues related to data collection, quality control, data storage, analysis, and dissemination of precipitation data.

What is needed is a sustained effort to monitor precipitation via citizen scientists. To achieve sustainability, such an effort needs to be both accurate and cost effective. The latter part may be attainable through leveraging low-tech MacGyver-type solutions—but only if they lead to accurate and reproducible observations (note that MacGyver was a popular television show in the late 1980s and early 1990s that often highlighted the ability of the protagonist—Angus "Mac" MacGyver—to make just about anything from commonly available materials). To this end, our research was conducted in the context of SmartPhones4Water (S4W), a California based non-profit organization investigating how young researchers, citizen scientists, and mobile technology can be mobilized to help close growing water data gaps (including precipitation). S4W's first pilot project in Nepal (S4W-Nepal; Davids et al., 2018a,b) was launched in early 2017. This paper focuses on the 2018 monsoon (May through September) precipitation monitoring efforts in Nepal using low-tech gauges (in contrast to high-tech approaches like Netatmo).

Our research questions can be organized into two primary categories: (1) low-cost S4W precipitation gauge analyses and (2) citizen scientist involvement.

1. **S4W precipitation gauge analyses**
   a. *What are the types and magnitudes of errors for S4W's low-cost precipitation gauges?*
   b. *How do precipitation measurements from S4W's low-cost gauges compare to other commonly used gauges, including the Department of Hydrology and Meteorology (DHM), Nepal standard gauge?*
2. **Citizen scientist involvement**
   a. *How effective were our methods to recruit citizens to join the monsoon precipitation monitoring campaign?*
   b. *How effective were our methods to motivate citizens to continue taking daily precipitation measurements?*
   c. *What were the types and frequencies of common citizen scientist observation errors?*
   d. *What were the average costs per observation for citizen scientists, and did this relate to citizen scientist performance?*

## CONTEXT AND STUDY AREA

To answer our research questions, S4W-Nepal launched a 2018 monsoon precipitation monitoring campaign; 154 citizen scientists generated 6,656 precipitation measurements using low-cost (<1 USD) S4W gauges constructed from repurposed soda bottles, concrete, and rulers. Measurements were recorded with smartphones using an Android-based application called Open Data Kit (ODK) Collect, and included GPS-generated coordinates, observation date and time, photographs, and citizen scientist reported readings. Measurements were primarily in the Kathmandu Valley and Kaski District of Nepal (**Figure 1**).

Precipitation in Nepal is highly heterogeneous, both spatially and temporally. Spatial variability of precipitation in Nepal is driven by (1) strong convection and (2) orographic effects (Nayava, 1980). Temporal fluctuations are mostly due to the South Asian summer monsoon (June to September)—a south to north moisture movement perpendicular to the Himalayas (**Figure 1**) along the southern rim of the Tibetan Plateau (Flohn, 1957; Turner and Annamalai, 2012). Roughly 80% of Nepal's (and South Asia's in general) precipitation occurs during the summer monsoon (Nayava, 1974; Shrestha, 2000). Annual precipitation in Nepal varies spatially by more than an order of magnitude, ranging from 250 mm on the northern (leeward) slopes of the Himalayas to over 3,000 mm around Pokhara in the Kaski District (Nayava, 1974). In general, both (1) the percentage of annual rainfall occurring during the summer monsoon rainfall and (2) total annual precipitation decrease from the center of the country westward. About 88% of our 2018 monsoon measurements were performed in Nepal's Kathmandu Valley. Within the Kathmandu Valley, average monsoon precipitation (42 years average) is 1,040 mm (Pokharel and Hallett, 2015), with average annual precipitation being roughly 1,300 mm at Tribhuvan International Airport. Thapa et al. (2017) state that average annual precipitation ranges from roughly 1,500 mm in the Valley floor to 1,800 mm in the surrounding hills.

## METHODS AND MATERIALS

### S4W Rain Gauge

#### Construction and Use

S4W gauges were constructed from recycled clear plastic bottles (e.g., 2.2-liter Coke or Fanta bottles in Nepal) with a 100 mm diameter, concrete, rulers, and glue (**Figure 2A**). A tutorial video describing how to construct an S4W rain gauge is available on S4W's YouTube channel (https://bit.ly/2sItFTh; Nepali language only). The clear plastic bottles had uniform diameters for at least 200 mm from the base toward the top; bottles with non-uniform cross sections were not used. Concrete was placed in the bottom of the bottle up to the point where the uniform cross section begins. The concrete provided a level reference surface for precipitation measurements. The additional weight from the concrete also helped to keep the gauge upright during windy conditions. Bottle lids were cut off at the point where the inward taper begins. This lid was then inverted and placed on top of the gauge in an attempt to minimize evaporation losses—which can be a major source of rain gauge error (Habib et al., 2001). A simple measuring ruler of sufficient length with millimeter graduations was glued vertically onto the side of the bottle. The ruler was placed with the zero mark at precisely the same level as the surface of the concrete. In order to minimize variability and possible introduction of errors, all gauges used in this investigation were constructed by S4W-Nepal. Each S4W gauge costs <1 USD in terms of materials and takes roughly 15 min to make (assuming a minimum of 10 gauges are constructed at a time).

The S4W gauge design is similar to what Hendriks (2010) proposed as a low-budget rain gauge, except that the addition of a solid base and measuring scale enabled direct measurements of precipitation depths, thus eliminating the need to measure water volumes. Similar low-cost funnel-type gauges have also been used extensively in rainfall partitioning studies (Lundberg et al., 1997; Thimonier, 1998; Marin et al., 2000; Llorens and Domingo, 2007).

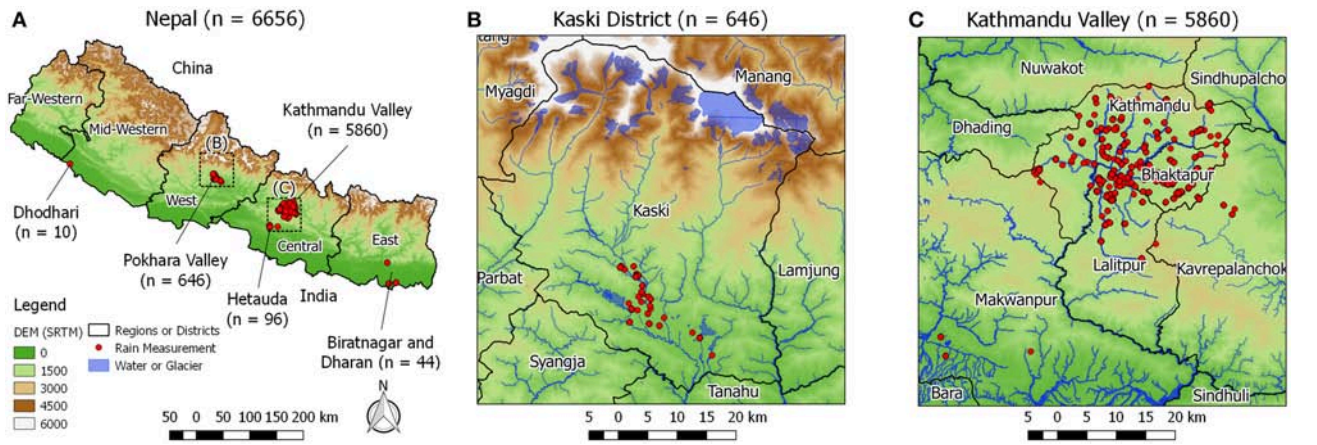



**FIGURE 1 |** Locations of 2018 Monsoon (May to September) precipitation measurements with the number of measurements shown in parentheses for **(A)** Nepal, with enlarged views of **(B)** the Kaski District, including the Pokhara Valley, and **(C)** the Kathmandu Valley. Topography shown from a Shuttle Radar Telemetry Mission (SRTM) 90-m digital elevation model (DEM) (SRTM, 2000).

**FIGURE 2 | (A)** Repurposed plastic bottle after placement of concrete, ruler, and inverted lid. **(B)** S4W Gauges installed on the roof of the S4W-Nepal office in Thasikhel, Lalitpur, Nepal. After selecting the parameter to measure, the citizen scientist **(C)** entered their observation of precipitation (mm) and **(D)** took a picture of the water level in the S4W rain gauge before emptying it. Each record was reviewed by S4W-Nepal staff to ensure that the numeric entry from the citizen scientist **(C)** matched the photographic record of the observation **(D)**. Any observed discrepancies were corrected, and records of edits were maintained.

Precipitation measurements were performed by citizen scientists using an Android smartphone application called Open Data Kit Collect (ODK Collect; Anokwa et al., 2009). Video tutorials of how to install and use ODK and perform S4W precipitation measurements are available on S4W's YouTube channel (https://bit.ly/2Rdtadx; Nepali language only). Citizen scientists collected the precipitation data presented in this paper by performing the following steps:

1. S4W gauges were installed in locations with open views of the sky (e.g., **Figure 2B**).
   a. Gauge heights above ground surface ranged from 1 meter (m) in rural areas to over 20 m (on rooftops) in densely populated urban areas.
2. An inverted lid without a cap (i.e., Cap1; see *Evaporation errors*) was used to minimize evaporation losses.
3. Measurements were performed as often as daily but sometimes less frequently.
4. Gauges were removed from their stands and placed on a level surface.
5. Precipitation readings were taken as the height of the lower meniscus of the water level within the bottle with the gauge placed on a level surface.
6. A numeric reading of precipitation level was entered into ODK in millimeters (mm; **Figure 2C**).
7. ODK was used to record a photograph of the water level with the smartphone camera level to the water surface (**Figure 2D**). ODK was also used to record date, time, and GPS coordinates.

8. Water was quickly dumped from the gauge to ensure that all ponded water above the concrete surface was removed but moisture within the concrete was retained.
9. The measurement was saved locally to smartphone memory and sent to the S4W-Nepal ODK Aggregate server running on Google App Engine.

   (a) ODK was designed to work offline (i.e., without cellular connection) and can be configured to automatically or manually send data after the connection is restored.

## Error Analysis

The World Meteorological Organization (WMO, 2008) identified the following primary error sources for precipitation measurements (estimated magnitudes in parentheses): evaporation (0–4%), wetting (1–15%), wind (2–10% for rain), splashing in or out of the gauge (1–2%), and random observational and instrument errors. The first three sources of errors are all systematic and negative (WMO, 2008). Because of the S4W gauge design, we separated wetting into concrete soaking and condensation on the clear plastic walls. The resulting categories of S4W gauge errors included: (1) evaporation, (2) concrete soaking, (3) condensation, and (4) other. Unlike some observation errors, which can be identified and corrected from photographs, gauge related errors must be understood and, if possible, systematically corrected. The following sections provide additional details regarding the first three sources of gauge errors related to the S4W gauge being low-cost and non-standard in nature. While all gauge errors were originally measured by differences in mass, all errors were converted to an equivalent depth (mm) for comparison. It should be noted that other rainfall gauge related errors, such as errors in construction of the gauge, errors related to placement of the gauge (e.g., a gauge installed too close to a building or below vegetation), or errors related to maintenance of the gauge (e.g., clogging) were not analyzed but are described in more detail below.

### *Evaporation errors*

For manually read gauges, evaporation errors occur when precipitation evaporates from the rain gauge prior to taking a reading. Gauge design, weather, and the duration between precipitation events and gauge readings all impact the magnitude of the evaporation errors. To assess evaporation errors for S4W gauges, we performed evaporation tests between June 5th and August 23rd, 2018. We evaluated the impact of the following three rain gauge cover configurations on evaporation losses: (a) Open (i.e., no lid), (b) Cap1 (i.e., lid without cap), and (c) Cap2 (i.e., lid with cap and 7 mm hole; **Figure 3**). We randomly selected three gauges for each of these cover configurations for a total of nine gauges. With these nine gauges, we performed eleven sets of 24 h evaporation measurements yielding a total of 99 evaporation observations (i.e., 33 for each cover configuration).

We performed an initial investigation to see if the depth of water in the gauge had a noticeable impact on evaporation losses. We investigated two water depths (i.e., 10 and 30 mm) that corresponded to commonly observed rainfall events in the Kathmandu Valley. Our initial results showed that evaporation losses were not noticeably different between the 10 and 30 mm depths, so we used 30 mm depths for the remainder of the tests.

During each 24 h period, all nine gauges were set on the roof of the S4W-Nepal office in Thasikhel, Lalitpur (https://goo.gl/maps/oq81TwPAZnk) in a place with full exposure to the sun and wind. If precipitation occurred during the 24 h period, the experiment was canceled and restarted the following day. We used an EK1051 [Camry] electronic weighing scale (accuracy $\pm$ 1 g $\approx \pm$ 0.08 mm) to determine evaporation losses by measuring the mass of the gauges before and after each successful (i.e., no precipitation) 24 h period.

### *Concrete soaking errors*

As previously described, S4W gauges have a concrete base. As a semi-porous media, concrete requires a certain amount of moisture prior to saturation and subsequent ponding or accumulation of water above the concrete surface. The amount of water absorbed prior to ponding is a function of the concrete mixture (e.g., type and ratio of materials, etc.), the volume of concrete, and the initial moisture content of the concrete. The depth of precipitation read from S4W gauges represents only precipitation that accumulates above the concrete surface. Any precipitation that soaks into the concrete itself was not included in gauge readings. Therefore, concrete soaking represented a systematic negative error.

To evaluate soaking, we used an EK1051 [Camry] electronic weighing scale to measure the mass of the nine gauges used in the evaporation tests in both dry and saturated conditions. For the first set of measurements, the concrete had cured and dried for 30 days and no additional water beyond the amount initially needed for making the concrete mixture had been introduced to the gauge. To saturate the concrete, $\sim$100 mm of water was added to the gauge and left for a period of 24 h. Subsequent soaking measurements were performed after drying the gauges in sunlight for periods ranging between one and 3 days.

### *Condensation errors*

For S4W gauges with Cap1 and Cap2 covers, condensation accumulated on the clear plastics sides of the rain gauge. Because we used weight as a measurement to quantify evaporation losses, condensation was not included as a loss; only water that fully exited the rain gauge was considered an evaporation loss. However, water that evaporates and subsequently condenses on the gauge walls causes a lowering of the ponded water level, or the amount of moisture within the concrete if no ponded water is present. Therefore, condensation constitutes a systematic negative error in S4W gauge readings.

To evaluate condensation, we filled the same nine gauges with roughly 5 mm of water and covered them with a Cap2 cover. The gauges were placed in the sun for $\sim$2 h to allow condensation to develop. Condensation was removed from gauges by wiping the inside of each gauge completely dry with a paper towel, ensuring that any remaining ponded water at the bottom was avoided. We determined condensation with an EHA501 [Camry] electronic weighing scale (accuracy $\pm$0.1 g $\approx \pm$0.008 mm) by measuring the mass difference between each saturated and dry paper towel.

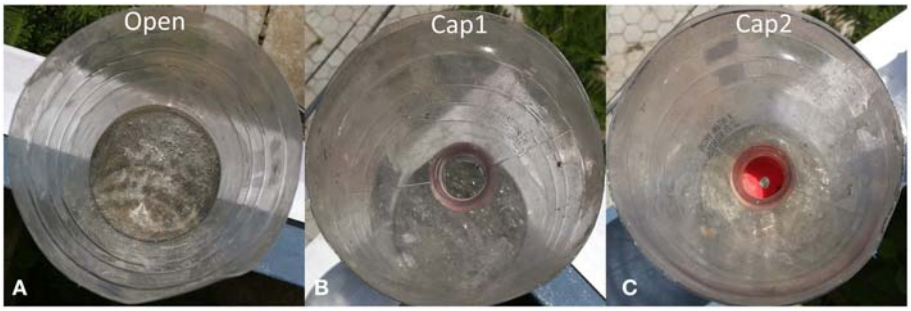



**FIGURE 3 |** Three different rain gauge cover configurations for evaporation measurements. Open **(A)** is completely open to the atmosphere. Cap1 **(B)** has the original top of the bottle inverted and placed back on top of the gauge. Cap2 **(C)** has the same cover but also includes the original soda bottle cap with a 7 mm punched or drilled hole in the center to allow precipitation to enter the gauge. The resulting areas open to evaporation were roughly 7,850, 530, and 40 $mm^2$ for Open, Cap1, and Cap2 covers, respectively. The diameters of the cover and the lower portion of the gauge are the same, but the thickness of the plastic material causes a tight connection between the cover and the gauge.

***Other errors not included in this analysis***

Differences in gauge installation can impact precipitation measurements. For example, gauge height can influence systematically negative wind-induced errors (Yang et al., 1998) or cause splash into the gauge. Wind-induced errors average between 2 and 10% and increase with decreasing rainfall rate, increasing wind speed, and smaller drop size distributions (Nešpor and Sevruk, 1999). Gauges that are not installed level will also cause an undercatch. The suitability of all gauge installation locations used in this paper were evaluated by S4W-Nepal staff by reviewing pictures of each gauge installation. Any issues identified from pictures were communicated directly to citizen scientists via personal communication (SMS, phone call, or site visit) and corrective actions were taken. However, installation errors are not the focus of this work and the data collected to date were insufficient to characterize these errors; therefore, gauge installation errors were not analyzed.

Gauge construction quality can also introduce errors. If future studies use gauges constructed by citizen scientists themselves (not the case in this study), the errors related to differences in construction quality should be considered.

Other possible maintenance or observation errors that may impact citizen scientists' measurements include: clogging of gauge inlets, incomplete emptying of gauges, and taking readings on unlevel surfaces. Effective training and follow-up is likely the key to minimizing such errors, so future work should explore different training approaches and their efficacy for various audiences. Training approaches should also consider scalability; for example, site visits become impractical if there are 1,000 participants.

### Comparison to Standard Rain Gauges

To evaluate the accuracy of S4W gauges, a comparison with three other gauges (within 5 meters) was performed in Bhaisepati, Lalitpur, Nepal from May 1st, 2017 to April 30th, 2018 (**Figure 4**). Measurements were generally taken within 12 h of the end of each precipitation event, and in the morning or evening to minimize condensation errors. The other gauges included an Onset Computer Corporation Hobo Tipping Bucket RG3-M Rain Gauge (Onset), a manually read Community Collaborative Rain, Hail, and Snow Network standard gauge (CoCoRaHS), and a manually read standard 203 mm (8-inch) diameter Nepali Department of Hydrology and Meteorology gauge (DHM; similar to US National Weather Service 203 mm (8-inch) gauges). The Onset gauge measured the date and time of every 0.2 mm of precipitation from June 3rd to November 23rd, 2017.

We used DHM gauge measurements as the reference or actual precipitation. Because Onset data were not available for the entire year period (i.e., May 1st, 2017 to April 30th), cumulative errors for the Onset gauge are not presented. Only fully overlapping data sets between DHM and Onset are used. Based on DHM measurements, we grouped the data into three precipitation event sizes (i.e., 0–5, 0–25, and 0–100 mm).

## Recruiting and Motivating Citizen Scientists

Citizen science projects rely on citizens. As such, the success of any citizen science project relies at least partly on successful citizen recruitment and engagement efforts. We decided to focus monitoring on a 5-month period from May through the end of September in 2018. Even though the monsoon usually does not start until the middle of June (Ueno et al., 2008), starting the campaign in May provided time to ramp up interest and participation. Interested and motivated citizen scientists were encouraged to continue measurements after the campaign. We recruited citizen scientists for the monitoring campaign with a variety of methods (the number of citizen scientists recruited with each method is shown in parentheses):

- **R1: Leveraging personal relationships ($n = 53$)**—At the time of the 2018 monsoon expedition, the S4W-Nepal team was comprised of nine young researchers (i.e., BSc, MSc, and Ph.D. researchers or recent graduates). Our first round of citizen science recruiting started with our personal connections. Each of us asked our family, friends, and colleagues to consider joining the S4W-Nepal monsoon monitoring campaign.
- **R2: Social media posts ($n = 11$)**—We made posts on S4W's Facebook page (https://www.facebook.com/

**FIGURE 4 |** Comparison between **(A)** four different gauges including: **(B)** Onset Computer Corporation Hobo tipping bucket, **(C)** Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS) standard 102 mm (4-inch) diameter gauge, **(D)** S4W gauge, and **(E)** the Nepali Department of Hydrology and Meteorology (DHM) standard 203 mm (8-inch) diameter gauge (similar to a US National Weather Service 203 mm (8-inch) gauge).

SmartPhones4Water) in order to explain the monsoon monitoring campaign and invite interested individuals to join as citizen scientists. S4W-Nepal's 2018 monsoon monitoring expedition titled "Count the Drops Before It Stops" included the main themes of "Join, Measure, and Change the way water is understood and managed in Nepal" (poster included as **Supplementary Material**).

- **R3: Outreach programs at schools/colleges ($n$ = 61)**—In order to reach larger groups of possible citizen scientists, we organized outreach events at four secondary schools and five colleges during the spring of 2018. The outreach programs typically included presentations about the global water cycle, the Asian South Monsoon, the Kathmandu Valley water crisis, the importance of measuring resources we are trying to manage, and how the S4W-Nepal project is trying to quantitatively "tell the story" of the Valley's water problems to citizens and policy makers alike, with the aim to improve understanding and management in the future. Outreach programs generally ended with a call for volunteers, practical training on how to measure precipitation, and the distribution of S4W gauges to interested individuals. In the case of secondary schools, S4W gauges were provided to the schools directly, along with large pre-printed canvas graphs for plotting both daily precipitation amounts and cumulative monsoon precipitation totals.
- **R4: Random site visits ($n$ = 29)**—The recruiting methods above mainly reached people living in the core urban areas of the Valley. However, our goal was to maximize the spatial distribution of our precipitation monitoring network, so it was important to include sites in the surrounding rural areas as well. In order to recruit citizen scientists in these areas, we made random site visits to strategic areas lacking citizen scientists. Sometimes during these random site visits, we would first talk to local community members to explain the vision and importance of the S4W-Nepal project. If community members responded positively, we would ask for references of individuals with a general interest in science and technology who had working Android smartphones. At other times, we started a dialogue directly with people we thought might be interested. In either case, once an individual with a working Android smartphone showed interest, we together installed an S4W gauge and performed initial training, including taking a first measurement together. In roughly 10 cases, we provided donated Android smartphones to individuals who were keenly interested in participating, but did not have a working smartphone.

To visualize recruitment progress, we developed a heatmap of the number of measurements performed showing time by week on the horizontal axis and (A) individual citizen scientists, (B) recruitment method, and (C) motivational method on the vertical axis. When computing grouped averages, zeroes were used for citizen scientists who did not take measurements in the respective weeks. We used the Mann-Whitney $U$ test (Mann and Whitney, 1947) for the entire 22-week period to determine if a significantly different number of measurements were taken for all possible pairs of recruitment methods and between paid (see motivation M7 below for details on payments) and volunteer citizen scientists. Citizen scientist composition was defined by four categories including: (A) volunteer or paid, (B) gender, (C) age, and (D) education. For education, citizen scientists were classified based on the highest level of education they had either completed or were currently enrolled in.

Once a citizen scientist has been successfully recruited it is critical to motivate their continued involvement. Previous studies have shown that appropriate and timely feedback is a key motivation factor for sustaining citizen science (Buytaert et al., 2014; Sanz et al., 2014; Mason and Garbarino, 2016; Reges et al., 2016). Essentially, there were two different combinations of motivations for the volunteers ($n$ = 117) and paid ($n$ = 37)

citizen scientists, respectively. Motivations M1 through M6 were applied to all volunteers; whereas, M1, M2, and M7 were applied to paid citizen scientists.

- **M1: Personal follow-ups**—At the end of each week, we reviewed the performance of each citizen scientist and developed plans for personal follow-ups for the subsequent week. Follow-ups focused on citizen scientists who had taken measurements in the last month but had not taken a measurement in the last 5 days, or citizen scientists making either unit or meniscus errors (see section Performance of Citizen Scientists). Personal follow-ups included (a) SMS messages, (b) phone calls, and (c) site visits. Roughly 20 site visits were made each week, amounting to an average of two visits per volunteer, and five (i.e., monthly) visits per paid citizen scientists during the 5-month campaign. During personal follow-ups, S4W-Nepal staff reiterated the importance of the work the citizen scientists were doing, and the difference that their measurements were making. Another purpose was to develop stronger personal relationships and develop a sense of being part of a larger community of people who are passionate about improving the way water resources are managed in Nepal.
- **M2: Bulk SMS messages**—At the end of each week we provided personalized bulk SMS messages to all citizen scientists who had taken measurements during the 2018 monsoon campaigns. The goals of the messages were to acknowledge the citizen scientists' contributions, to summarize their measurements in a meaningful way, and to reinforce that their data was making a difference. The personalized message read: "Hello from S4W-Nepal! From *StartDate* to *EndDate* you have taken *NumberOfMeasurements* totaling *AmountOfPrecipitation* mm. Your data is making a difference! https://bit.ly/2Rb15Uo" where StartDate was the beginning of the monsoon campaign, *EndDate* was the date of the citizen scientists' most recent measurement, *NumberOfMeasurements* was the number of measurements and *AmountOfPrecipitation* was the cumulative depth of presentation between *StartDate* and *EndDate*. The link at the end of the message was to S4W's Facebook page.
- **M3: Outreach and workshops**—Because Nepal is a collectivist or group society, we thought it was important to gather as an entire group at least once a year for a post-monsoon celebration. At this celebration, preliminary results from our efforts were presented and stories from the citizen scientists were shared. We also did follow-up visits to schools that measured precipitation.
- **M4: Use of the data**—S4W's aim is to share all of the data we generate, but our data portal is not finished yet. We encouraged citizen scientists to continue their participation by providing them with all the data generated by the monsoon monitoring campaign.
- **M5: Lucky draws**—We held a total of nine lucky draws (i.e., raffles) for gift hampers that included earphones, study lamp, wallet, movie ticket, and mobile balance credits. Only citizen scientists taking regular measurements (i.e., at least 50% of the time) were entered into the lucky draw.
- **M6: Certificates of involvement**—Especially for high school, undergraduate, and graduate students, certificates are important motivational factors because companies or organizations looking for new hires consider participation and employment certificates an important part of a candidate's resume. In order to get a certificate, citizen scientists had to take measurements for at least 50% of the days during the monsoon.
- **M7: Payments**—In some cases, especially in rural areas with limited employment opportunities, where the need for data was high, and the number of possible volunteers was low, S4W-Nepal compensated citizen scientists for measurements. For these citizen scientists, S4W-Nepal provided a small per observation transfer to their mobile phone account. Precipitation observations earned 25 Nepali Rupees (NPR; roughly 0.22 USD).

We used the number of measurements per citizen scientists as a simple indicator of the effectiveness of motivational efforts. For each group in each citizen scientist characteristic (i.e., volunteer or paid, gender, age, and education level), we used the Kruskal-Wallis $H$ test (Kruskal and Wallis, 1952) to see if there were statistically significant differences (alpha level = 0.01) between the number of measurements taken by citizen scientists per group in each category per month during the entire 5-month period. For example, for age, we tested if more measurements per month were taken by ≤18 compared to both 19–25 and >25, and so forth.

## Performance of Citizen Scientists

Using a custom Python web application, we manually reviewed pictures from every precipitation observation to ensure that values entered by citizen scientists (**Figure 2C**) matched photographic records (**Figure 2D**). Any observed discrepancies were corrected, and records of edits were maintained. Through this process we identified three categories of citizen science observation errors: unit, meniscus, and other errors. Unit errors caused an order of magnitude difference between original citizen scientist values and edited values due to citizen scientists taking readings in centimeters instead of millimeters. Meniscus errors were caused by citizen scientists taking readings of capillary rise instead of the lower portion of the meniscus, which was as much as 3 mm in some cases. Other observation errors were errors caused by unknown factors.

The combination of edit ratio and edit distance was used to determine the type of error for each corrected record. Edit ratio was calculated as:

$$ER_i = \frac{OV_i}{EV_i} \tag{1}$$

where $ER_i$ is the edit ratio, $OV_i$ is the original precipitation value, and $EV_i$ is the edited precipitation value for record i. Unit errors were defined as records with edit ratios between 8 and 12. Edit distance was calculated as:

$$ED_i = OV_i - EV_i \tag{2}$$

where $ED_i$ is edit distance for record i. Meniscus errors were defined as records with edit ratios <8 and edit distances between 0 and 3. The remaining edited records (neither unit nor meniscus errors) were classified as unknown observation errors.

On a weekly interval, we performed additional training and follow up (via SMS, phone, or in person) with citizen scientists who had made measurement errors during the previous week. Performance ratio was used to evaluate individual and group performance and was calculated as:

$$PR_{CS,\,t} = \frac{TNM_{CS,\,t} - NCM_{CS,\,t}}{TNM_{CS,\,t}} \times 100\% \tag{3}$$

where $PR_{CS,t}$ is the performance ratio for one or more citizen scientists (CS) during time period (t), $NCM_{CS,t}$ is the number of corrected measurements, and $TNM_{CS,t}$ is the total number of measurements for the same citizen scientist(s) (CS) and time period (t). Performance ratio (%) ranges from 0 to 100 with 100% being ideal.

We used the Mann-Whitney *U* test (Mann and Whitney, 1947) to evaluate if the interquartile range (IQR) of citizen scientists (in terms of the number of measurements they took) had worse performance ratios (PRs). After dividing citizen scientists into two groups based on the number of measurements they took during the 5 months campaign [i.e., (1) the IQR and (2) the remainder], we calculated the Mann-Whitney *U* on the PRs (alpha level = 0.01).

## Cost per Observation

In order to evaluate the cost effectiveness of our approach, and any relationships between cost and citizen science performance, we performed a reconnaissance-level cost per observation (CPO) analysis. For each citizen scientist, average CPO was calculated as:

$$CPO_{CS,\,t} = \frac{EC_{CS,\,t} + RC_{CS,\,t} + MC_{CS,\,t}}{TNM_{CS,\,t}} \tag{4}$$

where EC is equipment costs, RC is recruiting costs, MC is motivational costs, and TNM is the total number of measurements for each citizen scientist (CS) and time period (t). In this case, the time period was 5 months from May through September 2018. The following general assumptions were used for the CPO analysis:

- All costs (**Table 1**) are in Nepali rupees (NRP); an exchange rate of 114.3 NPR (November 22nd, 2018) per United States dollar (USD) was used for currency conversion
- All costs assume an hourly labor rate of 50 NPR per hour
- The full study period of 22 weeks was used for calculating costs unless stated otherwise

It is important to have a general sense of Nepal's economic context to properly interpret CPO results. Nepal's per capita gross domestic product (GDP) in 2018 was 1,004 USD or 114,800 NPR (CEIC, 2019). Assuming 2,080 working hours per year (i.e., 40 h work week for 52 weeks), the average hourly rate for 2018 was 0.48 USD or 55 NPR per hour.

All citizen scientists used the S4W gauge, so equipment costs were constant. RC was different for citizen scientists depending on which recruitment strategy (R1 through R4) was applied; we assumed that only one recruitment strategy was ultimately responsible for each citizen scientists' participation (recruitment methods per citizen scientist are included as **Supplementary Material**). **Table 1** details the assumptions used to develop recruitment and motivational costs.

Motivational costs (MCs) for volunteers (MCVol) were entirely fixed, and were solved for using Equation 5. For paid citizen scientists, MCs were a combination of fixed (MCPaid; Equation 5) and variable costs (M7; Equation 6). MCs were calculated with the following equation:

$$MC_{CS,t} = \begin{cases} M1_a + M1_b + M1_c V + M2 + M3 + M4 + M5 + M6, & \text{if CS is Volunteer} \\ M1_a + M1_b + M1_c P + M2 + M7_{CS,t}, & \text{if CS is Paid} \end{cases} \tag{5}$$

where the variables are defined above, with the exception of $M7_{CS,t}$ for paid citizen scientists. $M7_{CS,t}$ was calculated as:

$$M7_{CS,t} = TNM_{CS,t} * R_{Precip} \tag{6}$$

where $R_{precip}$ is the payment rate for each precipitation measurement. $TNM_{CS,t}$ was limited to a maximum of one measurement per day.

# RESULTS

## S4W Rain Gauge Results

Of the S4W gauge errors investigated (**Table 2**), initial (post-cure) concrete soaking errors ($n = 9$) and evaporation without lids (Open; $n = 33$) were the largest, averaging 3.9 mm and 3.7 mm day$^{-1}$, respectively. Subsequent concrete soaking requirements ($n = 99$) averaged 0.8 mm, or roughly five times smaller than the initial soaking requirement. S4W gauge evaporation was reduced from Open by an average of 86% (0.5 mm day$^{-1}$) and 92% (0.3 mm day$^{-1}$) for Cap1 and Cap2 configurations, respectively. Condensation errors were similar to Cap2 evaporation, and averaged 0.31 mm ($n = 49$).

Cumulative precipitation amounts for the 1 year of data collected were 900, 930, and 927 mm for the S4W, CoCoRaHS, and DHM gauges, respectively. Using DHM as the reference for the entire year of data, cumulative gauge error was −2.9% for S4W and 0.3% for CoCoRaHS. Measured precipitation amounts were linearly correlated for the three precipitation ranges, but the correlation decreased in strength as total precipitation decreased (**Figure 5**). Points near the horizontal axis of **Figure 5A** ($n = 9$) indicate that some small rain events ($n = 5$ for DHM less than the 0.8 mm soaking loss; $n = 4$ for DHM between 0.8 and 2 mm) were completely missed by the S4W gauge.

For S4W, the magnitude of the systematic underestimation increased for smaller measurements (**Figures 5A–C**). For example, for precipitation measurements between 0 and 5 mm (**Figure 5A**), the S4W gauge linear regression coefficient was 0.95 indicating that measurements were on average −5% from the DHM gauge. In contrast, linear regression coefficients for 0 to 25 and 0 to 100 mm ranges were 0.96 (−4%) and 0.98

**TABLE 1 |** Summary of the results from the evaporation, soaking, and condensation experiments (error type), including configuration, unit, sample size (*n*), mean, minimum (min), maximum (max), and standard deviation (stdev).

| Error type | Configuration | Unit | n | Mean | Min | Max | Stdev |
|---|---|---|---|---|---|---|---|
| Evaporation | Open | mm day$^{-1}$ | 33 | 3.7 | 2.1 | 5.8 | 1.0 |
| | Cap1 | mm day$^{-1}$ | 33 | 0.5 | 0.1 | 1.0 | 0.2 |
| | Cap2 | mm day$^{-1}$ | 33 | 0.3 | 0.1 | 1.3 | 0.3 |
| Soaking | Initial (post-cure) | mm | 9 | 3.9 | 2.0 | 4.7 | 0.9 |
| | Subsequent | mm | 99 | 0.8 | 0.4 | 1.9 | 0.5 |
| Condensation | – | mm | 49 | 0.31 | 0.04 | 0.51 | 0.11 |

**TABLE 2 |** Number and compositions of citizen scientists taking measurements from May through September 2018.

| Year-Month | Active CS | Paid | Volunteer | Female | Male | ≤18 | 19–25 | >25 | <Bachelors | Bachelors | >Bachelors |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2018-05 | 121 | 21 | 100 | 47 | 74 | 11 | 87 | 23 | 21 | 92 | 8 |
| 2018-06 | 106 | 26 | 80 | 39 | 67 | 11 | 76 | 19 | 20 | 79 | 7 |
| 2018-07 | 96 | 30 | 66 | 38 | 58 | 12 | 63 | 21 | 21 | 65 | 10 |
| 2018-08 | 93 | 30 | 63 | 35 | 58 | 11 | 64 | 18 | 21 | 63 | 9 |
| 2018-09 | 64 | 20 | 44 | 26 | 38 | 10 | 43 | 11 | 15 | 43 | 6 |

*Active citizen scientists (CS) took at least one measurement during the respective month. ≤18, 19–25, and >25 refers to the citizen scientist's age, and <Bachelors, Bachelors, and >Bachelors refers to the highest level of education the citizen scientist had either completed or was currently enrolled in.*

(−2%), respectively. Measurements from the CoCoRaHS gauge were strongly correlated with the measurements from the DHM gauge for all ranges with small biases (linear regression coefficients between 1.00 and 1.01; **Figures 5D–F**). For Onset, the magnitude of systematic overestimation increased for larger events (**Figures 5G–I**), from 1.07 (7%) at 0 to 5 mm, and up to 1.09 (9%) and 1.12 (12%) at 0 to 25 and 0 to 100 mm ranges, respectively.

## Recruiting and Motivating Citizen Scientists Results

A heatmap of citizen scientists' precipitation measurements per week illustrates the rate of recruitment along with the continuity of their measurements (**Figure 6A**). "Citizen science heroes" can be seen as the persistent dark blue rows (e.g., the second row down from the top). In contrast, inconsistent citizen scientists can be seen as the rows with large variations in blue (e.g., fifth and sixth rows down from the top). Unfortunately, several citizen scientists took only a few measurements during their first week, especially toward the end of the second week (e.g., 2018-19). At a 0.05 alpha level, the average number of measurements per week was significantly higher for citizen scientists recruited via social media (R2) vs. personal relationships (R1; **Figure 6B**; $p = 0.018$), recruited via outreach programs (R3) vs. personal relationships (R1; **Figure 6B**; $p = 0.033$), and motivated with payments vs. volunteers (**Figure 6C**; $p = 0.013$). At an alpha level of 0.01, the average number of measurements per week was significantly higher for recruitment by random site visits (R4) vs. personal connections (R1; **Figure 6B**; $p = 0.003$). No other statistically significant differences (alpha level = 0.05) were observed between the remaining possible pairs of recruitment methods.

The number of active citizen scientists peaked in May ($n = 121$) and decreased through the campaign until September ($n = 64$; **Table 3**). The ratio of female to male citizen scientists remained relatively stable throughout the period (mean = 63%). From May to September, the number of volunteer citizen scientists decreased by 66%, whereas the number of paid citizen scientists only decreased by 5%. The most stable age group was ≤18, followed by 19–25, and finally >25. In terms of education, <Bachelors and >Bachelors were more stable than Bachelors, which decreased by 53%.

From May through September 2018, the average citizen scientist took 42 measurements (min = 1, max = 148, std = 39). Sixteen citizen scientists took only one measurement. Based on results from Kruskal-Wallis *H* tests, paid citizen scientists took significantly more measurements than volunteers (**Figure 7**; alpha level = 0.01; $p = 0.005$). No other statistically significant differences in contributions were observed.

There were statistically significant correlations between the number of measurements taken and mean daily precipitation for the same day (**Figure 8A**; $r = 0.60$; $r$ critical = 0.21; alpha level = 0.01) and the previous day (**Figure 8B**; $r = 0.38$; $r$ critical = 0.21; alpha level = 0.01), but the strength of the same day correlation was stronger, explaining 36% of the variance, while the previous day precipitation explained only 14%. This suggests that the harder it rains the more likely citizen scientists are to take a measurement that same day (and the next but less so).

## Performance of Citizen Scientists Results

Citizen scientist observation errors were found for 9% ($n = 592$) of the total measurements ($n = 6656$). Meniscus errors ($n = 346$) (**Figure 9**; light blue area) accounted for 58% of observation errors. Unit errors ($n = 50$) (**Figure 9**; light red sector) comprised 8% of the errors. Finally, unknown errors ($n = 196$) accounted for the remaining 33% of observational errors.

Only six citizen scientists had Unit, Meniscus, and Unknown errors. 41 citizen scientists had both Meniscus and Unknown

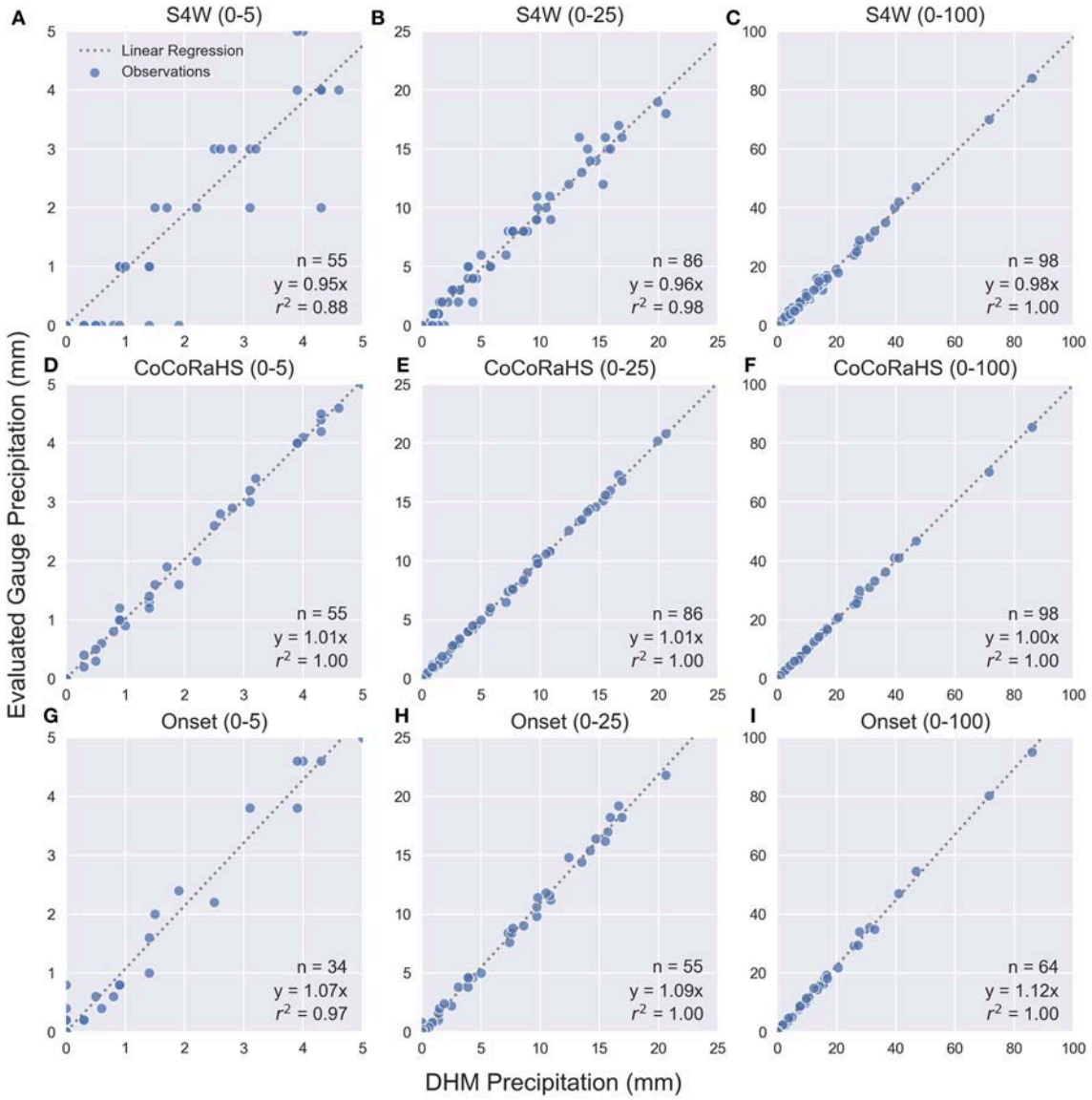



**FIGURE 5 |** Comparison of precipitation data from S4W, CoCoRaHS, and Onset gauges using DHM observations as the reference (i.e., horizontal axis). Per reference DHM measurements, data was filtered into three precipitation event ranges: 0–5 mm [i.e. panels **(A,D,G)**], 0–25 mm [i.e., panels **(B,E,H)**], and 0–100 mm [i.e., panels **(C,F,I)**]. No precipitation events above 100 mm were recorded. Data shown are from May 1st, 2017 through April 30th, 2018. The period of record for the Onset gauge was June 3rd, 2017 to November 27th, 2018; only fully overlapping data sets between Onset and DHM were used, resulting in decreased sample sizes for panels **(G–I)**.

errors; 10 had both Meniscus and Unit errors; and 8 had Unit and Unknown errors. The largest number of errors for a citizen scientist was 32, or 22% of their 143 records. The mean citizen scientist performance ratio (PR) was 93% (**Figure 10**). Stated alternatively, on average, there were errors on 7% of the measurements from citizen scientists. There were a total of 63 citizen scientists with perfect PRs (100%); 10 of these recorded more than the median number of measurements and 53 less (38 below Q1). Citizen scientists who took a moderate number of measurements (i.e., interquartile range (IQR) between Q1 and Q3; middle 50%) were significantly more likely to have a worse PR than those outside of the interquartile range (**Figure 10**; alpha level = 0.01; $p = 0.0001$).

## Cost per Observation Results

Fixed costs for equipment (S4W gauge) were 0.87 USD. Fixed costs for recruiting ranged from 0.66 to 5.02 USD, while for motivation they were 8.79 and 8.45 USD for volunteer and paid citizen scientists, respectively (**Table 4**; see **Table 1** for details). Variable costs were only applicable for paid citizen scientists, and were 0.22 USD per observation. Outreach programs recruited the largest number of citizen scientists ($n = 61$), but were also the most expensive recruitment method (5.02 USD per citizen scientists recruited). Leveraging personal relationships was the second most effective ($n = 53$) and cheapest approach (0.66 USD). Random site visits recruited 29 citizen scientists, of whom 27 were paid, and cost roughly 2.45 USD per recruited

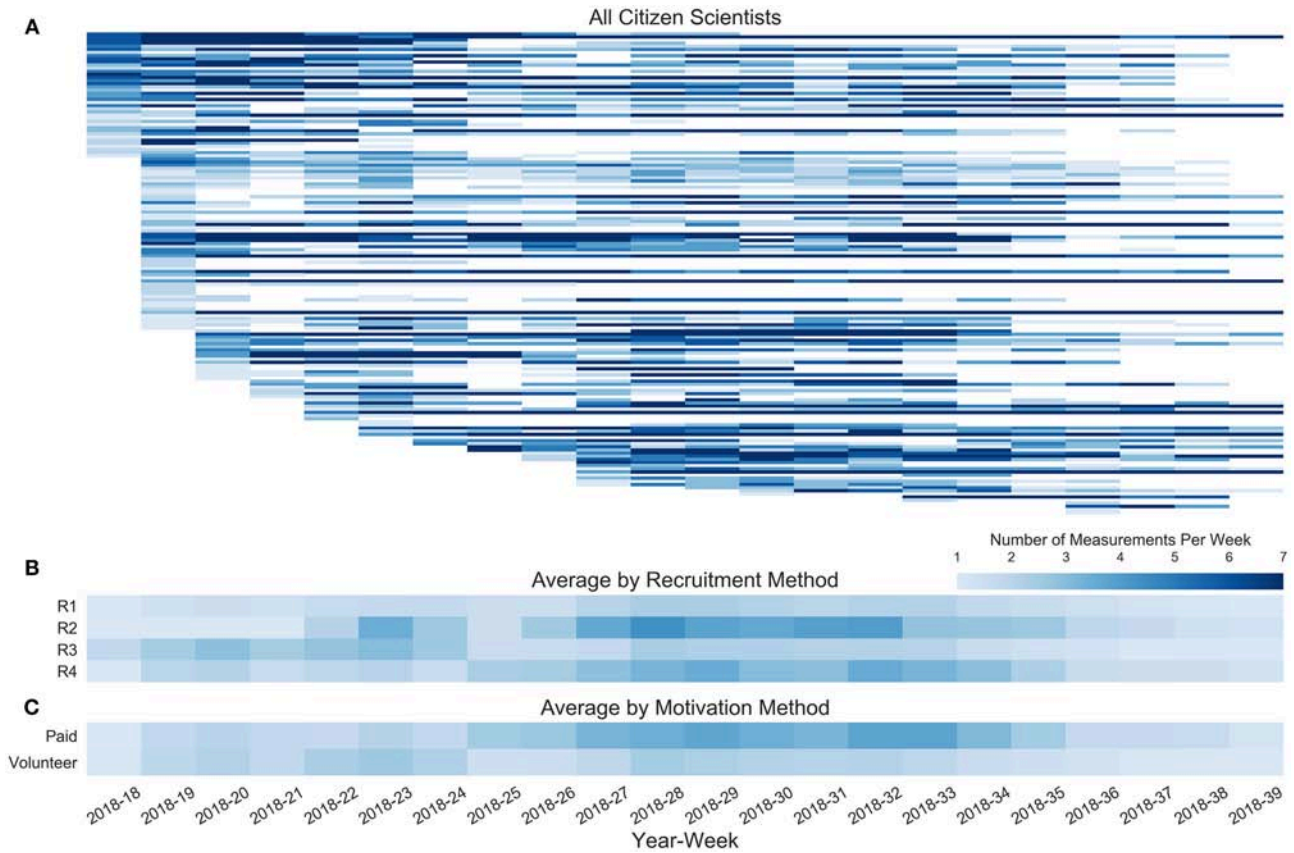



**FIGURE 6 |** Heatmap of the number of measurements per year-week for a 22-week period from the first week of May (i.e., 2018-18) through the end of September (i.e., 2018-39). Each column of pixels represents a single week. Each row of pixels represents **(A)** an individual citizen scientist ($n = 154$), **(B)** averages from the four recruitment methods [i.e., R1: Leveraging personal relationships ($n = 53$); R2: Social media ($n = 11$); R3: Outreach programs ($n = 61$); R4: Random site visits ($n = 29$)], or **(C)** motivation method group [i.e., paid ($n = 37$) or volunteer ($n = 117$)]; see sections Recruiting and Motivating Citizen Scientists and Recruiting and Motivating Citizen Scientists Results for details). The color of each pixel represents the number of measurements performed each week. Light and dark blue represent one and seven measurements, respectively; white means zero measurements were performed that week. For panel **(A)** citizen scientists are sorted vertically in reverse chronological order by the date of their first measurement; the rate of recruitment is shown by the slope of the left edge of pixels in the heatmap—larger negative slopes (i.e., 2018-18 and 2018-19) represent higher recruitment rates. When computing grouped averages for panels **(B,C)**, zeroes were used for citizen scientists that did not perform measurements in the respective weeks.

citizen scientist. Only 11 citizen scientists joined the monitoring campaign purely through social media, for a cost of 1.75 USD per recruited citizen scientist.

Estimated average costs per observation (CPO) for all citizen scientists ranged from 0.07 to 14.68 USD and 0.30 to 11.99 USD for volunteer and paid citizen scientists, respectively (**Figure 11**). Median CPOs where 0.47 USD for both volunteer and paid citizen scientists. Because all costs for volunteers are fixed, the number of observations per citizen scientist had the largest impact on CPOs. For example, volunteer citizen scientists (recruited with outreach programs) that took only one measurement had CPOs of 14.68 USD (**Figure 11A**). For paid citizen scientists, fixed costs were lower, but an additional variable cost of 0.22 USD (25 NPR) was added due to per observation payments. This resulted in a smaller range of CPOs, where (1) minimum CPOs approached per observation payment amount as the number of observations performed increased and (2) maximum CPOs approached fixed costs for paid citizen scientists as the number of measurements approached one (**Figures 11C,D**). Performance ratio (PR) did not appear to be related with CPO (**Figures 11A,B**).

Gauge cost had a large impact on fixed costs for all citizen scientists. For example, increasing gauge cost from 0.87 USD (S4W gauge) to 31.50 USD (CoCoRaHS gauge) increased median CPOs from 0.47 to 1.57 and 1.12 USD for volunteer and paid citizen scientists, respectively. Using DHM gauges, which cost 65.60 USD, increases median CPOs to 2.88 and 1.85 USD for volunteer and paid citizen scientists, respectively. This analysis was limited to 5 months, however, since the estimated lifespan of all three gauges is well over 5 months (perhaps 5 years or longer), CPOs will decrease as more measurements are taken. As gauge lifespan increases, CPOs approach the sum of annually recurring fixed costs plus per observation variable costs.

**TABLE 3 |** Assumptions and the resulting costs for each recruitment and motivational category.

| CPO category | Sub category | Assumptions | Cost (NPR) | Cost (USD) |
|---|---|---|---|---|
| R1 | | Leveraging personal relationships took four staff 10 h per week for 2 weeks, for a total of 4,000 NPR. Since 53 citizen scientists were recruited with this method, the cost was 75 NPR per citizen scientist recruited. | 75 | 0.66 |
| R2 | | For social media, an investment of 2 h per week at 50 NPR was made. Since 11 citizen scientists were recruited with this method, the cost was 200 NPR per citizen scientist recruited. | 200 | 1.75 |
| R3 | | Workshops and outreaches were organized at a total of four schools and five colleges/universities. Workshops at schools and colleges/universities were estimated to cost 2,500 and 5,000 NPR, respectively. Since 61 CS were recruited with this method, the cost was 574 NPR per citizen scientist recruited. | 574 | 5.02 |
| R4 | | Random site visits were used to recruit 29 citizen scientists in rural areas. Assuming a two-person team, working for 8 h, plus 40 km traveled per day, a daily subsistence allowance of 200 NPR/person, and recruitment of 5 citizen scientists per day, the average cost was 280 NPR per citizen scientist recruited. | 280 | 2.45 |
| M1 | | There are three types of personal follow ups: SMS (M1a), phone calls (M1b), and site visits (M1cV and M1cP) | | |
| M1 | M1a | For SMS, we assumed that each citizen scientist received eight SMS messages during the monsoon, and that each message cost 10 NPR, for a total of 80 NPR. | 80 | 0.70 |
| M1 | M1b | For phone calls, we assumed that each citizen scientist received eight phone calls, and that each call cost 15 NPR, for a total of 120 NPR. | 120 | 1.05 |
| M1 | M1c | Assuming a two-person team, working for 8 h, plus 40 km traveled per day, a daily subsistence allowance of 200 NPR/person, and visits of 10 citizen scientists per day, the average cost was 140 NPR per citizen scientist site visit. | | |
| M1 | M1cV | For site visits, we assumed that each volunteer citizen scientist received two site visits, for a total of 280 NPR per volunteer citizen scientist. | 280 | 2.45 |
| M1 | M1cP | For site visits, we assumed that each paid citizen scientist received five site visits, for a total of 700 NPR per paid citizen scientist. | 700 | 6.12 |
| M2 | | Bulk SMS messages were sent weekly, and cost roughly 3 NPR per message including the time to generate and load the necessary report(s), for a total of 66 NPR per citizen scientist. | 66 | 0.58 |
| M3 | | Outreach workshops focused on motivating volunteer citizen scientists, at an estimated cost of 40,000 NPR total, or with 117 volunteer citizen scientists, 342 NPR per volunteer. | 342 | 2.99 |
| M4 | | The motivation of data use was considered to have negligible cost, because of existing infrastructure necessary for other purposes. | 0 | 0.00 |
| M5 | | Lucky draws were used as a motivation for volunteer citizen scientists. A total of nine lucky draws were performed, with an estimated cost of 1,200 NPR each for 117 volunteers, or 92 NPR per volunteer. | 92 | 0.80 |
| M6 | | Certificates were used to motivate volunteer citizen scientists, and cost 25 NPR each. | 25 | 0.22 |
| M7 | | Payments were used to motivate paid citizen scientists, and cost 25 NPR per observation. | 25 | 0.22 |

*See section Recruiting and Motivating Citizen Scientists for more detailed descriptions of each category.*

# DISCUSSION

## S4W Rain Gauge Discussion

In the context of wind induced errors arising from using (or not using) wind shields or differences in gauge heights, which can be as large as 10% for precipitation gauges of the same type (Sevruk and Klemm, 1989), the S4W gauge errors related to evaporation, soaking, and condensation are relatively small. Nevertheless, our findings highlight the importance of (1) using covers to minimize evaporation (regardless of cap type), in addition to (2) effective training on how to properly install covers to minimize air gaps and evaporation losses. Since evaporation can be limited by the amount of time that ponded water is stored in the gauge, citizen scientists should be encouraged to take measurements as quickly as possible after precipitation events. Citizen scientists should also be specifically guided to minimize the other errors discussed in section Error Analysis by: (1) keeping gauge inlets free of clogging hazards, (2) fully emptying gauges after measurements, and (3) taking readings on level surfaces.

Average S4W gauge evaporation losses with Cap1 (mean $= 0.5$ mm day$^{-1}$) and Cap2 (mean $= 0.3$ mm day$^{-1}$) compared favorably with Tretyakov gauge summer evaporation losses reported by Aaltonen et al. (1993), which ranged from 0.3 to 0.8 mm day$^{-1}$. Interestingly, Golubev et al. (1992) found evaporation losses from US National Weather Service 203 mm (8-inch) gauges (similar to the DHM gauge used

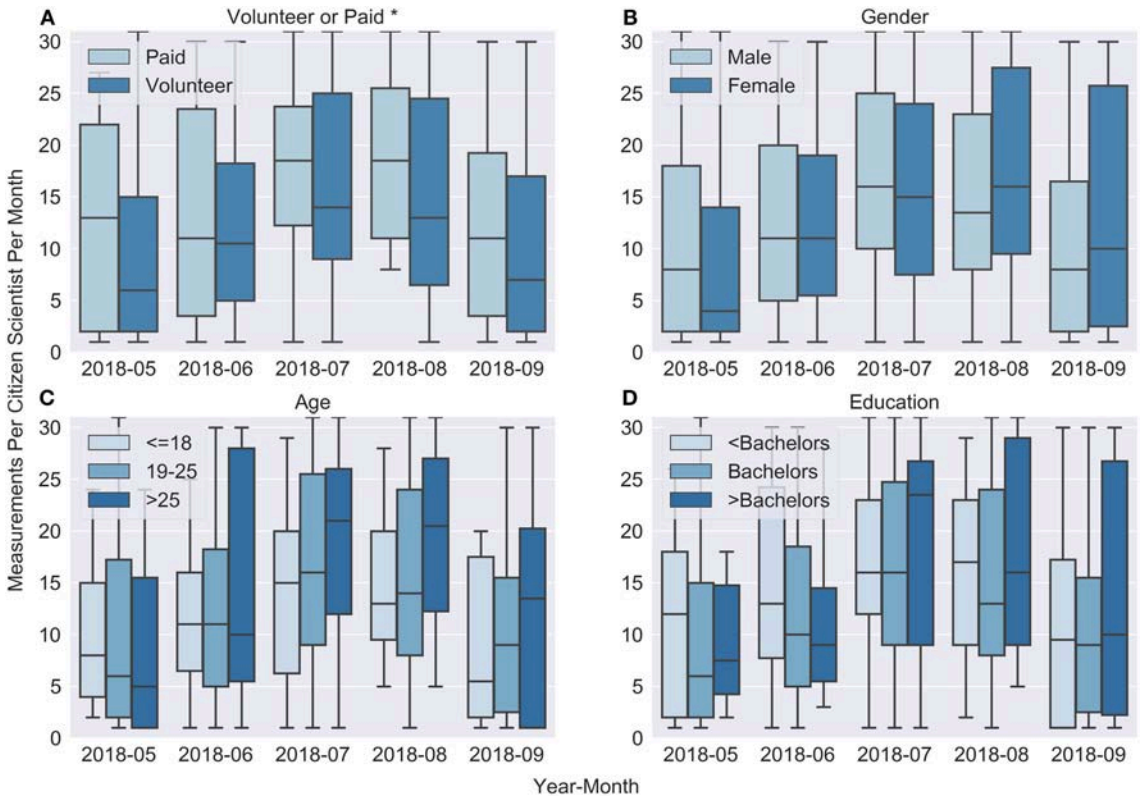



**FIGURE 7 |** Grouped box plots showing the medians and distributions of the number of citizen scientist precipitation observations per month. Box plot groups are shown for four different categories: **(A)** volunteer or paid; **(B)** gender, **(C)** age, and **(D)** education. For education, citizen scientists were classified into the highest education level that they had either completed or were currently enrolled in. An asterisk (*) in the subplot title indicates statistically significant differences (alpha level = 0.01) between the number of measurements performed by each group within that category during the entire 5-month period.

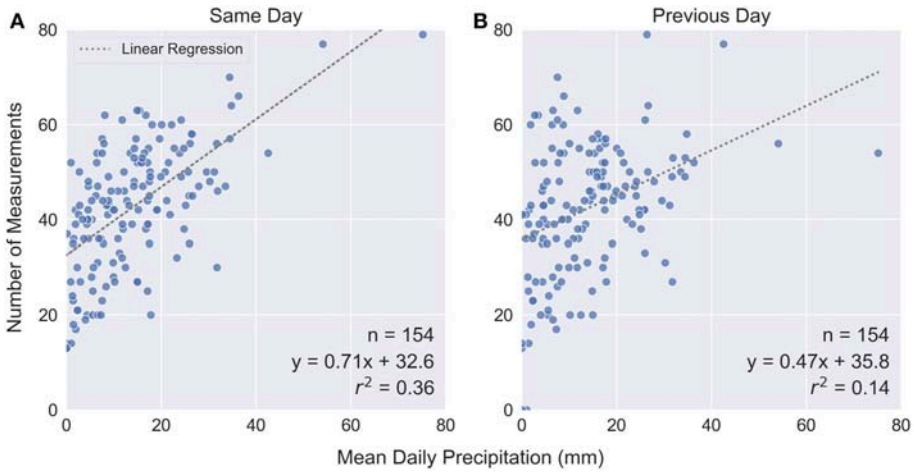



**FIGURE 8 |** Scatter plot of the number of measurements per day as a function of mean daily precipitation for the **(A)** same day and **(B)** previous day. Mean daily precipitation was taken as the average of all citizen scientists' measurements. There were statistically significant correlations (Pearson's *r*) for the **(A)** same day ($r = 0.60$; *r* critical = 0.21; alpha level = 0.01) and the previous day ($r = 0.38$; *r* critical = 0.21; alpha level = 0.01).

in this investigation) to be "negligible" (e.g., 0.2 mm day$^{-1}$). While variability in evaporation can be partially explained by differences in solar radiation, wind speed, temperature, and relative humidity (Sevruk and Klemm, 1989), it is also possible that small differences in cover installation could also explain part of the observed variability in evaporation losses.

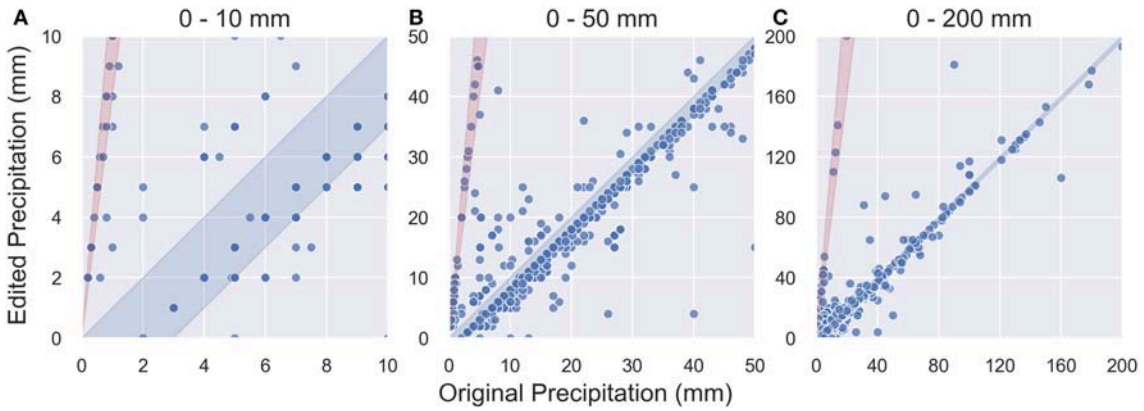



**FIGURE 9 |** Scatter plot of corrected records ($n = 592$) with original (i.e., raw) precipitation entries on the horizontal axis and edited (i.e., after quality control) values on the vertical axis. Data is shown for three different scales: **(A)** 0–10 mm, **(B)** 0–50 mm, and **(C)** 0–200 mm. Meniscus error range ($n = 346$) is shown as light blue area, while Unit error range ($n = 50$) is shown as light red sector. Points outside of the light blue and light red areas are unknown errors ($n = 196$).

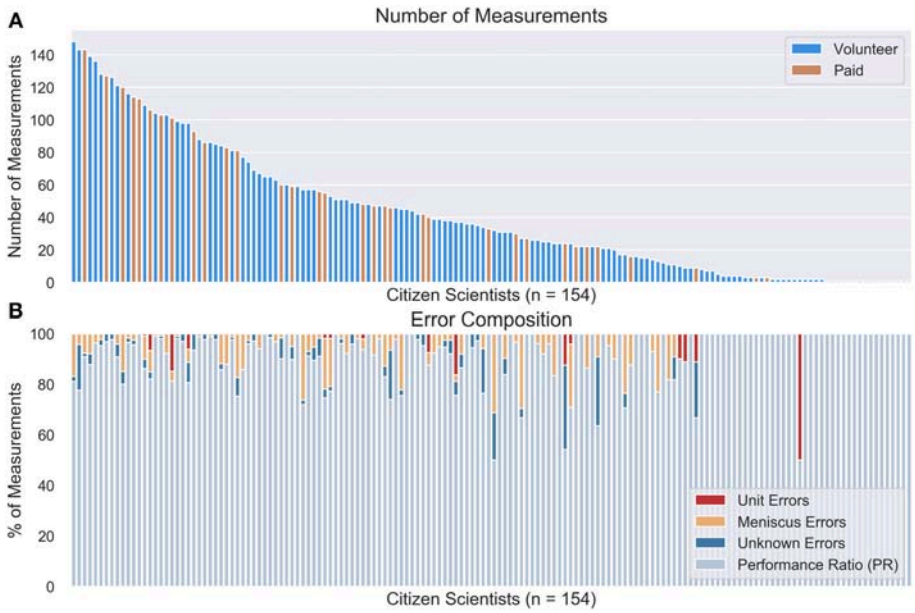



**FIGURE 10 |** Summary of **(A)** the number of measurements collected from May through September with volunteer and paid citizen scientists distinguished and **(B)** the corresponding error composition for all 154 citizen scientists. Citizen scientists sorted in descending order by their total number of measurements. Performance ratio (PR) becomes less informative as the total number of measurements for each citizen scientist decreases, especially at or below two.

For example, if a cover is installed at an angle, or not firmly pressed down, a small opening between the lid and the inside of the gauge can remain. These small openings could account for some of the high evaporation rates observed with Cap1 (max = 1.0 mm day$^{-1}$) and Cap2 (max = 1.3 mm day$^{-1}$) cover configurations (**Table 2**).

S4W gauges should be manually saturated prior to data collection to avoid the first roughly 3.9 mm of rain going to concrete saturation (**Table 2**). While subsequent saturation took only 0.8 mm, if not corrected for, this could introduce systematic negative bias into S4W gauge measurements. In order to reduce the need for corrections, alternative lower-porosity materials for filling the bottom of S4W gauges should be investigated.

Citizen scientists should be encouraged to take measurements at a consistent time in the morning (e.g., 07:00 LT; Reges et al., 2016) to minimize condensation errors and to simplify data processing. S4W gauge condensation averaged 0.31 mm, which is 61% of observed average daily Cap1 evaporation rates (0.5 mm day$^{-1}$) and 39% of concrete saturation requirements (0.8 mm). While percentage-wise, condensation errors were smaller than evaporation and concrete saturation, taking measurements in the morning (or evening) when condensation accumulations are low

**TABLE 4 |** Summary of fixed and variable costs for equipment, recruitment, and motivation per citizen scientist, including the number of applicable citizen scientists.

| Cost type | Description | Number of citizen scientists | Per citizen scientist fixed costs (USD) | Per observation variable costs (USD) |
|---|---|---|---|---|
| Equipment | S4W Gauge | 154 | 0.87 | – |
| Recruitment | R1: Personal relationships | 53 | 0.66 | – |
| | R2: Social media | 11 | 1.75 | – |
| | R3: Outreach programs | 61 | 5.02 | – |
| | R4: Random site visits | 29 | 2.45 | – |
| Motivation | MCVol: Volunteer motivations | 117 | 8.79 | – |
| | MCPaid: Paid motivations | 37 | 8.45 | 0.22 |

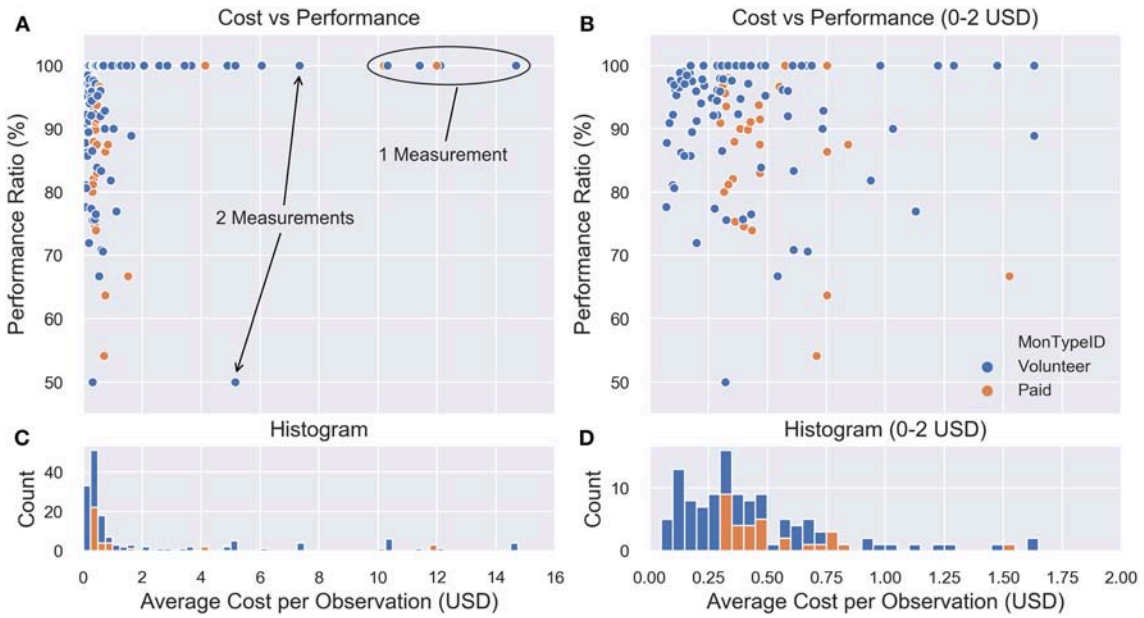



**FIGURE 11 |** Scatter plots of performance ratio (PR) as a function of average cost per observation for costs from **(A)** 0–16 USD and **(B)** 0–2 USD ranges, respectively. Each point represents the performance ratio and average cost per observation for a single citizen scientist. Histograms below show the total number of citizen scientists in each cost bin for **(C)** 0–16 USD and **(D)** 0–2 USD ranges, respectively.

can reduce these errors. A correction for condensation errors could be added if the time of a measurement is during peak daylight hours.

While S4W gauge error was relatively small (−2.9%) compared to the DHM standard, it is still possible to apply corrections for the systematic S4W gauge errors. We suggest that corrections could be based on either an (1) error correction factor (ECF) or (2) evaporation (EVAP). The ECF uses cumulative precipitation values for S4W and DHM gauges to develop a constant correction, which is our case was 1.03. After adjusting S4W gauge records with the ECF approach, corrected cumulative S4W precipitation matched the DHM total of 927 mm. Alternatively, the EVAP approach is based on average daily evaporation (i.e., 0.5 mm) with soaking requirements (i.e., 0.8 mm) as an upper limit. After applying the EVAP approach, corrected cumulative S4W precipitation was 943 mm, or roughly 1.8% higher than DHM. Additional details regarding both of these approaches are included as **Supplementary Material**.

It is important to note that gauge errors, or systematic measurement differences, arising from differences in gauge installations were not evaluated. While standardizing gauge installation criteria like gauge height could help to minimize these differences, it may not be practical to apply such standards to citizen science projects in urban areas. For example, in the densely populated mid-rise core urban areas of Kathmandu, installing precipitation gauges at 1 m would only be possible in large courtyards. In these cases, it is likely more practical (and accurate) to install rain gauges on roof tops.

S4W gauge evaluation results should be considered the likely errors for "ideal" citizen scientists. Other possible errors that may impact citizen scientists' measurements include: (1) clogging of gauge inlets, (2) incomplete emptying of gauges, (3) improper gauge installation, and (4) taking readings on unlevel surfaces. Because we performed gauge intercomparison measurements ourselves with focused attention on avoiding these issues, they are not reflected in our results. Future work should consider

the impacts of these potential error sources on citizen scientist measurements. Since it is likely that effective training and follow-up is the key to minimizing such errors, future work should also explore the effectiveness of different training approaches on different audiences.

## Recruiting and Motivating Citizen Scientists Discussion

Our results showed that citizen scientists recruited via random site visits (R4; alpha = 0.01), social media (R2; alpha = 0.05), and outreach programs (R3; alpha = 0.05) on average took significantly more measurements than those recruited from personal connections (R1). Since all but two citizen scientists recruited from random site visits were also paid, it is not clear if the greater number of measurements was due to the recruitment method or payment, or a combination of the two. Citizen scientists who were recruited via social media had to take several self-initiated steps to move from (1) initially seeing something about S4W-Nepal on social media to (2) collecting precipitation data during the 2018 monsoon. In contrast, the barrier to entry for other recruitment methods was lower, and was externally initiated through interpersonal interactions. Therefore, the initial investment and motivation level of citizen scientists who joined the monitoring campaign through social media is relatively higher.

A survival analysis of volunteers in CoCoRaHS, the longest running large scale citizen science-based precipitation monitoring effort, found that retirement aged participants (i.e., ages 60 and above) were most likely to continue taking measurements (Sheppard et al., 2017). This suggests that older citizen scientists are most easily motivated, at least in a western context. While we did not have any retirement aged participants, our oldest age group (>25) actually had the largest attrition rates (52%). Future citizen science projects in Asia should focus on involving older citizen scientists to test the validity of this finding in the context of Nepal or other Asian settings.

Since payment appears to be an effective motivation, future work should explore how payment can be used as an effective means of recruitment. Also, recruitment of citizen scientists should be expanded to focus on retirement age groups and on clear communication of the usefulness of generated precipitation data.

While we only observed statistically significant differences in citizen science performance due to payment, roughly half of the bachelor's students involved in the project continued their involvement in the project (attrition rate was 53% for the 5 months campaign) without monetary motivations (no bachelor's students received payments). This suggests that students can be motivated to participate in citizen science projects with incentives like (1) the opportunity to use data for their research projects (e.g., bachelor's theses), (2) lucky draws (i.e., raffles or giveaways), and (3) by receiving certificates of involvement. However, these student-focused incentives often lead to data collection in urban areas, and may not be effective at generating data in rural areas with limited student populations and relatively low scientific literacy levels. In such areas, payments may be the most effective near-term incentive.

Survey results from CoCoRaHS volunteers have shown that a significant motivational factor is the knowledge that the data they are providing is useful (Reges et al., 2016). Therefore, a key component of any citizen science project should be "closing the loop" back to citizen scientists by clearly communicating the usefulness of their data, along with easy to understand examples. Our experience has shown that the difficulty of "closing the loop" increases as the citizen scientists' scientific literacy decreases. Therefore, in places like rural Nepal with, on average, relatively low scientific literacy rates, additional efforts must be made to properly contextualize and connect abstract concepts like data collection and fact-based decision making to the daily lives of community members. Payments might also be an important intermediate solution to motivate involvement while generational improvements in scientific literacy are realized.

Finally, even though we specifically reinforced the value of measuring zeros during training, our results suggested that the magnitude of precipitation was an important motivator for citizen scientists. However, there was some noise in this relationship because for the citizen scientists who did not take measurements, it was unknown whether this occurred because (1) there was no measurable precipitation in their gauge that day, or (2) they simply did not take a measurement. Regardless, this suggests that it may be difficult to motivate people to continue taking regular measurements outside the monsoon season, so focused monsoon monitoring campaigns are a good solution.

## Performance of Citizen Scientists Discussion

Our findings reinforce the importance of including photographic records so that citizen science observations can be quality controlled and corrected if necessary. In our 5-month campaign, 9% of measurements required corrections; if not for photographic records, these errors may have been more difficult to detect, or may have gone unnoticed. It is important to note that the feedback we provided to citizen scientists about their errors during the campaign most likely led to fewer errors than there would have been without feedback. Future work should explore the opportunity to automate the quality control process by leveraging machine learning techniques to automatically retrieve correct values from photographs of measurements. Meniscus errors were more difficult to identify and correct from photographic records. Training citizen scientists to read the lower meniscus was at times a difficult task, because of the small variations in readings, often on the order of only a few millimeters.

## Cost per Observation Discussion

Median CPOs of 0.47 USD for both volunteer and paid citizen scientists were roughly equivalent to 1 h of labor at nationally averaged rates (0.48 USD per hour; see section Cost per Observation for details). The cost per observation analysis revealed well over an order of magnitude difference between

minimum and maximum average CPO for both volunteer and paid citizen scientists; this demonstrates the sensitivity of CPO to the number of observations. Our initial findings suggest that personal relationships and social media are the most cost-effective means of recruitment. A limitation of this study is that only two different groups of motivations were applied to volunteer and paid citizen scientists, respectively.

There was no increase in data accuracy with increases in CPO, thus efforts to minimize CPO do not appear to systematically lower PR. An important part of sustaining citizen science efforts is funding, and all efforts to minimize CPO while maintaining data quality will lead to lower funding requirements and greater chances of sustainability.

Since it is difficult to predict how citizen scientists will respond to recruitment and motivational efforts, returns on investments (as partially quantified by CPO) in citizen science monitoring efforts are uncertain and difficult to predict. Improved characterization of the effectiveness of different recruitment and motivational strategies will facilitate better understanding of the returns from citizen science-based precipitation monitoring investments.

## Outlook

Using gauges constructed by citizen scientists could make citizen science rainfall monitoring approaches more scalable. However, if such gauges are used in future studies, the errors related to differences in construction quality should be evaluated. Since this study did not investigate potential gauge errors arising from (1) gauge clogging, (2) incomplete draining, (3) improper gauge installation, and (4) taking readings on unlevel surfaces, future work should focus on characterizing these errors. Additionally, the effectiveness of different training approaches aimed at minimizing such errors should be evaluated. Opportunities to automate the quality control review process used in this study (i.e., manual retrieval of correct rainfall values from photographs) should also be investigated.

While leveraging personal relationships was a cost-effective means of citizen scientist recruitment, relying on this method poses challenges to scalability. Future efforts should focus on development and refinement of more scalable approaches. We see young researchers (grade 8 through graduate school) as potential catalysts toward expanding and sustaining citizen science-based monitoring efforts. Future work should explore how sustainable measurements of precipitation (and other parameters) can be achieved by linking standard measurement goals and methods developed by professional scientists with (1) young researchers, (2) citizen science at the community level, and (3) a common technology platform including low-cost sensors (not necessarily electronic). Involving young researchers in this process has the potential benefits of both improving the quality of their education and level of practical experience, while simultaneously providing valuable data to support fact-based decision making. As previously mentioned (see section Recruiting and Motivating Citizen Scientists Discussion), the potential role of retired aged participants (i.e., ages 60 and above) in Asian citizen science projects, along with the possibility of using payment as a means of recruitment should also be investigated.

Finally, future efforts should explore the potential for cross-cutting organizations to facilitate and catalyze this process by linking young water-related researchers across a range of academic institutions related to water including: natural sciences, agriculture, engineering, forestry, economics, sociology, urban planning, etc. Desired outcomes of these links would be to (1) encourage young researchers to focus their efforts on relevant and multidisciplinary research topics and (2) encourage academic institutions to integrate participatory monitoring into their curricula and academic requirements (Shah and Martinez, 2016). Ultimately, these young researchers can then become the champions of engaging citizen scientists in the communities where they grew up, live, research, and work.

# CONCLUSIONS

Our results illustrate the potential role of citizen science and low-cost precipitation sensors (e.g., repurposed soda bottles) in filling globally growing precipitation data gaps, especially in resource constrained environments like Nepal. Regardless of how simple low-cost gauges may be, it is critical to perform detailed error analyses in order to understand and correct, when possible, low-cost gauge errors. In this study, we analyzed three types of S4W gauge errors: evaporation (0.5 mm day$^{-1}$), concrete soaking (3.9 mm initial and 0.8 mm subsequent), and condensation (0.31 mm). Compared to standard DHM gauges, S4W and CoCoRaHS cumulative gauge errors were −2.9 and 0.3%, respectively, and were relatively small given the magnitude of other errors (e.g., wind induced) that affect all "catch" type gauges.

In total, 154 citizen scientists participated in the project, and on average performed 42 measurements ($n$ = 6,656 total) during the 5-month campaign from May to September 2018. Citizen scientists recruited via random site visits, social media, and outreach programs (listed in decreasing order) took significantly more measurements than those recruited via personal connections. Payment was the only categorization (i.e., not gender, education level, or age) that caused a statistically significant difference in the number of measurements per citizen scientist, and was therefore an effective motivational method. We identified three categories of citizen science observation errors ($n = 592$; 9% of total measurements): unit ($n = 50$; 8% of errors), meniscus ($n = 346$; 58% of errors), and unknown ($n = 196$; 33% of errors). Our results illustrate that simple smartphone-based metadata like GPS-generated coordinates, date and time, and photographs are essential for citizen science projects. Estimated cost per observation (CPO) was highly dependent on the number of measurements taken by each participant and ranged from 0.07 to 14.68 USD and 0.30 to 11.99 USD for volunteer and paid citizen scientists, respectively. Median CPOs were 0.47 USD for both volunteer and paid citizen scientists. There was no increase in data accuracy with increases in CPO, thus efforts to minimize CPO do not appear to systematically lower citizen scientist performance.

## DATA AVAILABILITY

The datasets generated and analyzed for this study can be found in the FigShare digital repository. All Python scripts used to analyze data and develop visualizations are included in the following GitLab repository: https://gitlab.com/jeff50/soda_bottle_science.

## AUTHOR CONTRIBUTIONS

JD had the initial idea for this investigation and designed the experiments in collaboration with MR, ND, AP, RP, and NvdG. Field work was performed by JD, AP, ND, and RP. JD prepared the manuscript with valuable contributions from all co-authors.

## FUNDING

This work was supported by the Swedish International Development Agency (SIDA) under grant number 2016-05801 and by SmartPhones4Water (S4W).

## ACKNOWLEDGMENTS

Most importantly, we want to thank each and every citizen scientist who joined the S4W-Nepal family during our monsoon monitoring campaigns. This research would not be possible without them. We appreciate the dedicated efforts of Eliyah Moktan, Anurag Gyawali, Amber Bahadur Thapa, Surabhi Upadhyay, Pratik Shrestha, Anu Grace Rai, Sanam Tamang, Kristina M. Davids, and the rest of the S4W-Nepal team of young researchers. We would also like to thank Dr. Ram Devi Tachamo Shah, Dr. Deep Narayan Shah, and Dr. Narendra Man Shakya for their supervision and support of this work. Lastly, thanks to the reviewers for their useful comments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00046/full#supplementary-material

## REFERENCES

Aaltonen, A., Elomaa, E., Tuominen, A., and Valkovuori, P. (1993). "Precipitation measurement and quality control, v2," in *Proceedings of the International Symposium on Precipitation and Evaporation* (Bratislava: Slovak Hydrometeorological Inst. and Swiss Federal Institute of Technology).

Anokwa, Y., Hartung, C., and Brunette, W. (2009). Open source data collection in the developing world. *Computer* 42, 97–99. doi: 10.1109/MC.2009.328

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., et al. (2014). Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. *Front. Earth Sci.* 2, 1–21. doi: 10.3389/feart.2014.00026

CEIC (2019). *Nepal 2018 per Capita Gross Domestic Product (GDP).* Available online at: https://www.ceicdata.com/en/indicator/nepal/gdp-per-capita (Accessed January 29, 2019).

Cifelli, R., Doesken, N., Kennedy, P., Carey, L. D., Rutledge, S. A., Gimmestad, C., et al. (2005). The community collaborative rain, hail, and snow network–informal education for scientists and citizens. *Am. Meteorol. Soc.* 86, 78–79. doi: 10.1175/BAMS-86-8-1069

Davids, J. C., Rutten, M. M., Pandey, A., Devkota, N., van Oyen, W. D., Prajapati, R., et al. (2018b). Citizen science flow – an assessment of citizen science streamflow measurement methods, Hydrol. *Earth Syst. Sci. Discuss.* 1:28. doi: 10.5194/hess-2018-425

Davids, J. C., Rutten, M. M., Shah, R. D. T., Shah, D. N., Devkota, N., Izeboud, P., et al. (2018a). Quantifying the connections—linkages between land-use and water in the Kathmandu Valley, Nepal. *Environ. Monit. Assess.* 190:304. doi: 10.1007/s10661-018-6687-2

Flohn, H. (1957). Large-scale aspects of the "summer monsoon" in South and East Asia. *J. Meteor. Soc. Japan.* 75, 80–186.

Golubev, V. S., Groisman, P. Y., and Quayle, R. G. (1992). An evaluation of the United States standard 8-in. nonrecording raingage at the valdai polygon, Russia. *J. Atmos. Ocean. Technol.* 9, 624–629. doi: 10.1175/1520-0426(1992)009<0624:AEOTUS>2.0.CO;2

Habib, E., Krajewski, W. F., and Kruger, A. (2001). Sampling errors of tipping-bucket rain gauge measurements. *J. Hydrol. Eng.* 6, 159–166. doi: 10.1061/(ASCE)1084-0699(2001)6:2(159)

Hendriks, M. R. (2010). *Introduction to Physical Hydrology*. New York, NY: Oxford University Press.

Hut, R. (2013). New Observational Tools and Datasources for Hydrology [dissertation]. [Delft (NL)]: Delft University of Technology. doi: 10.4233/uuid:48d09fb4-4aba-4161-852d-adf0be352227

Illingworth, S. M., Muller, C. L., Graves, R., and Chapman, L. (2014). UK citizen rainfall network: a pilot study. *Weather* 69, 203–207. doi: 10.1002/wea.2244

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., et al. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bull. Am. Meteorol. Soc.* 98, 69–78. doi: 10.1175/BAMS-D-14-00283.1

Koutsouris, A. J., Chen, D., and Lyon, S. W. (2016). Comparing global precipitation data sets in eastern Africa: a case study of Kilombero Valley, Tanzania. *Int. J. Climatol.* 36, 2000–2014. doi: 10.1002/joc.4476

Krajewski, W. F., Ciach, G. J., and Habib, E. (2003). An analysis of small-scale rainfall variability in different climatic regimes. *Hydrol. Sci. J.* 48, 151–162. doi: 10.1623/hysj.48.2.151.44694

Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621. doi: 10.1080/01621459.1952.10483441

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F. (2017). Inroads of remote sensing into hydrologic science during the WRR era. *Water Resour. Res.* 51, 7309–7342. doi: 10.1002/2015WR017616

Llorens, P., and Domingo, F. (2007). Rainfall partitioning by vegetation under Mediterranean conditions . a review of studies in Europe. *J. Hydrol.* 37–54. doi: 10.1016/j.jhydrol.2006.10.032

Lundberg, A., Eriksson, M., Halldin, S., Kellner, E., and Seibert, J. (1997). New approach to the measurement of interception evaporation. *J. Atmos. Ocean. Technol.* 14, 1023–1035. doi: 10.1175/1520-0426(1997)014<1023:NATTMO>2.0.CO;2

Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 50–60. doi: 10.1214/aoms/1177730491

Marin, C. T., Bouten, W., and Sevink, J. (2000). Gross rainfall and its partitioning into throughfall , stemflow and evaporation in four forest ecosystems in northwest. *J. Hydrol.* 237, 40–57. doi: 10.1016/S0022-1694(00)00301-2

Mason, C. E., and Garbarino, J. (2016). The power of engaging citizen scientists for scientific progress. *J. Microbiol. Biol. Educ.* 17, 7–12. doi: 10.1128/jmbe.v17i1.1052

McKinley, D. C., Miller-Rushing, A. J., Ballard, H. L., Bonney, R., Brown, H., Cook-Patton, S. C., et al. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol. Conserv.* 208, 15–28. doi: 10.1016/j.biocon.2016.05.015

Nayava, J. L. (1974). Heavy monsoon rainfall in Nepal. *Weather* 29, 443–450. doi: 10.1002/j.1477-8696.1974.tb03299.x

Nayava, J. L. (1980). Rainfall in Nepal. *Himal. Rev. Nepal Geogr. Soc.* 12, 1–18.

Nešpor, V., and Sevruk, B. (1999). Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Ocean. Technol.* 16, 450–464. doi: 10.1175/1520-0426(1999)016<0450:EOWIEO>2.0.CO;2

Pokharel, A. K., and Hallett, J. (2015). Distribution of rainfall intensity during the summer monsoon season over Kathmandu, Nepal. *Weather* 70, 257–261. doi: 10.1002/wea.2544

Pokharel, B., Marahatta, S., Giri, B., and Stahl, P. (2016). *CORAM-Nepal Website*. Available online at: http://coramnepal.org/about-coram/

Reges, H. W., Doesken, N., Turner, J., Newman, N., Bergantino, A., and Schwalbe, Z. (2016). CoCoRaHS: the evolution and accomplishments of a volunteer rain gauge network. *Bull. Am. Meteorol. Soc.* 97, 1831–1846. doi: 10.1175/BAMS-D-14-00213.1

Sanz, F. S., Holocher-Ertl, T., Kieslinger, B., García, F. S., and Silva, C. G. (2014). *White Paper on Citizen Science in Europe*. Available online at: http://www.socientize.eu/sites/default/files/white-paper_0.pdf

Sevruk, B., and Klemm, S. (1989). *Instruments and observing methods (Report No. 39)–catalogue of national standard precipitation gauges*. Geneva, Switzerland. Available online at: http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-39.pdf

Shah, H. R., and Martinez, L. R. (2016). Current approaches in implementing citizen science in the classroom. *J. Microbiol. Biol. Educ.* 17, 17–22. doi: 10.1128/jmbe.v17i1.1032

Sheppard, S. A., Turner, J., Thebault-Spieker, J., Zhu, H., and Terveen, L. (2017). Never too old, cold or dry to watch the sky. *Proc. ACM Hum. Comput. Interact.* 1, 1–21. doi: 10.1145/3134729

Shrestha, M. L. (2000). Interannual variation of summer monsoon rainfall over Nepal and its relation to southern oscillation index. *Meteorol. Atmos. Phys.* 75, 21–28. doi: 10.1007/s007030070012

SRTM (2000). *Shuttle Radar Topography Mission*. Available online at: https://earthexplorer.usgs.gov/ (Accessed September 14, 2016).

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L. (2018). A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Rev. Geophys.* 56, 79–107. doi: 10.1002/2017RG000574

Tauro, F., Selker, J., Van De Giesen, N., Abrate, T., Uijlenhoet, R., Porfiri, M., et al. (2018). Measurements and observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle. *Hydrol. Sci. J.* 63, 169–196. doi: 10.1080/02626667.2017.1420191

Thapa, B. R., Ishidaira, H., Pandey, V. P., and Shakya, N. M. (2017). A multi-model approach for analyzing water balance dynamics in Kathmandu Valley, Nepal. *J. Hydrol. Reg. Stud.* 9, 149–162. doi: 10.1016/j.ejrh.2016.12.080

Thimonier, A. (1998). Measurement of atmospheric deposition under forest canopies: Some recommendations for equipment and sampling design. *Environ. Monit. Assess.* 52, 353–387. doi: 10.1023/A:1005853429853

Turner, A. G., and Annamalai, H. (2012). Climate change and the South Asian summer monsoon. *Nat. Clim. Chang.* 2, 587–595. doi: 10.1038/nclimate1495

Ueno, K., Toyotsu, K., Bertolani, L., and Tartari, G. (2008). Stepwise onset of monsoon weather observed in the Nepal Himalaya. *Mon. Weather Rev.* 136, 2507–2522. doi: 10.1175/2007MWR2298.1

Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J. (2008). Rainfall and sampling uncertainties: a rain gauge perspective. *J. Geophys. Res. Atmos.* 113, 1–12. doi: 10.1029/2007JD009214

Volkmann, T. H. M., Lyon, S. W., Gupta, H. V., and Troch, P. A. (2010). Multicriteria design of rain gauge networks for flash flood prediction in semiarid catchments with complex terrain. *Water Resour. Res.* 46, 1–16. doi: 10.1029/2010WR009145

WMO (2008). *Guide to Meteorological Instruments and Methods of Observation*. WMO-No. 8. Geneva, Switzerland: World Meteorological Organization Available online at: https://library.wmo.int/pmb_ged/wmo_8_en-2012.pdf

Yang, D., Goodison, B. E., Metcalfe, J. R., Golubev, V. S., Bates, R., Pangburn, T., et al. (1998). Accuracy of NWS 8″ standard nonrecording precipitation gauge: Results and application of WMO intercomparison. *J. Atmos. Ocean. Technol.* 15, 54–68. doi: 10.1175/1520-0426(1998)015<0054:AONSNP>2.0.CO;2

**Conflict of Interest Statement:** 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.

*Copyright © 2019 Davids, Devkota, Pandey, Prajapati, Ertis, Rutten, Lyon, Bogaard and van de Giesen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 03 April 2019
doi: 10.3389/feart.2019.00067

# Low-Cost, Open-Source, and Low-Power: But What to Do With the Data?

*Jeffery S. Horsburgh[1]*, Juan Caraballo[2], Maurier Ramírez[2], Anthony K. Aufdenkampe[3], David B. Arscott[4] and Sara Geleskie Damiano[4]*

[1] Department of Civil and Environmental Engineering and Utah Water Research Laboratory, Utah State University, Logan, UT, United States, [2] Utah Water Research Laboratory, Utah State University, Logan, UT, United States, [3] LimnoTech, Oakdale, MN, United States, [4] Stroud Water Research Center, Avondale, PA, United States

OPEN ACCESS

**Edited by:**
Rolf Hut,
Delft University of Technology,
Netherlands

**Reviewed by:**
Ute Wollschläger,
Helmholtz Centre for Environmental Research (UFZ), Germany
Ioannis N. Daliakopoulos,
Technological Educational Institute of Crete, Greece
Chet Udell,
Oregon State University,
United States

***Correspondence:**
Jeffery S. Horsburgh
jeff.horsburgh@usu.edu

**Specialty section:**
This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science

**Received:** 07 December 2018
**Accepted:** 19 March 2019
**Published:** 03 April 2019

**Citation:**
Horsburgh JS, Caraballo J, Ramírez M, Aufdenkampe AK, Arscott DB and Damiano SG (2019) Low-Cost, Open-Source, and Low-Power: But What to Do With the Data? Front. Earth Sci. 7:67. doi: 10.3389/feart.2019.00067

There are now many ongoing efforts to develop low-cost, open-source, low-power sensors and datalogging solutions for environmental monitoring applications. Many of these have advanced to the point that high quality scientific measurements can be made using relatively inexpensive and increasingly off-the-shelf components. With the development of these innovative systems, however, comes the ability to generate large volumes of high-frequency monitoring data and the challenge of how to log, transmit, store, and share the resulting data. This paper describes a new web application that was designed to enable citizen scientists to stream sensor data from a network of Arduino-based dataloggers to a web-based Data Sharing Portal. This system enables registration of new sensor nodes through a Data Sharing Portal website. Once registered, any Internet connected data-logging device (e.g., connected via cellular or Wi-Fi) can then post data to the portal through a web service application programming interface (API). Data are stored in a back-end data store that implements Version 2 of the Observations Data Model (ODM2). Live data can then be viewed using multiple visualization tools, downloaded from the Data Sharing Portal in a simple text format, or accessed via WaterOneFlow web services for machine-to-machine data exchange. This system was built to support an emerging network of open-source, wireless water quality monitoring stations developed and deployed by the EnviroDIY community for do-it-yourself environmental science and monitoring, initially within the Delaware River Watershed. However, the architecture and components of the ODM2 Data Sharing Portal are generic, open-source, and could be deployed for use with any Internet connected device capable of making measurements and formulating an HTTP POST request.

**Keywords: low-cost, open-source, environmental sensors, data management, Arduino, Mayfly datalogger, EnviroDIY, Monitor My Watershed**

## INTRODUCTION

Although it is increasingly common for research groups, organizations, and agencies to collect time series data using *in situ* environmental sensors (Hart and Martinez, 2006; Rundel et al., 2009; Muste et al., 2013), the cost of environmental sensors and sensing systems is still a major limitation to their more widespread and long-term use. The base cost for the data logging components of a scientific-grade, *in situ* environmental monitoring station can be upward of $5000 USD, excluding the cost of the sensors for collecting the data, which may cost many thousands more. Additionally, while commercially available data logging and telemetry systems generally have robust and proven capabilities, they also tend to be proprietary, manufacturer specific, and closed, making it difficult in some cases to integrate dataloggers, communication peripherals, and sensors across manufacturers. These challenges associated with using existing commercial environmental sensing equipment, along with the now ubiquitous and inexpensive availability of easy to use microcontroller units such as the Arduino suite of products[1], single-board computers like the Raspberry Pi[2], and the diverse array of Internet of Things (IoT) devices have driven new innovations in low-cost, low-power, and do-it-yourself (DIY) environmental sensing and data logging (hereafter referred to as "low-cost sensing") (Baker, 2014; Ferdoush and Li, 2014; Wickert, 2014; Sadler et al., 2016; Beddows and Mallon, 2018).

Using increasingly off-the-shelf components, scientists of varying skill levels can now develop functional dataloggers for tens or hundreds of dollars rather than thousands, with capabilities for integrating high quality environmental sensors, or less expensive sensors that are now also increasingly available (Ensign et al., 2019). A variety of communication options are available for telemetering data, including cellular, spread spectrum radio, and Wi-Fi, and applications include continuous monitoring of indoor and outdoor air quality (Gualtieri et al., 2017; Karami et al., 2018), monitoring of ambient environmental conditions (Faustine et al., 2014; Adu-Manu et al., 2017), adaptive workflows and decision support using real-time data (e.g., Wong and Kerkez, 2016), among others. This ability to assemble fully functional environmental sensor stations for much lower cost is attractive to scientists, who, in many cases, wish to increase the spatial and temporal coverage of their data collection activities. Lower cost can potentially mean more stations, more sensors, and more information. Lower cost has also made these types of devices attractive to many citizen science data collection efforts.

With the development of these innovative low-cost sensing systems, however, comes the ability to generate large volumes of high-frequency data and the challenges of how to log, transmit, store, manage, and share the resulting data (Abu-Elkheir et al., 2013). Sensor data can be difficult to manage, especially as the number of sites, variables, and the time period over which observations are collected increases (Jones et al., 2015). Because Arduino microcontrollers, Raspberry Pi computers, and other systems like them are not purpose built as environmental dataloggers, one major challenge for using them in low-cost sensing applications lies in programming them to function as dataloggers (Jiang et al., 2016; Mazumdar et al., 2017). While this is becoming easier as the number of examples shared on the Internet increases, this is still left to the user. In contrast, many commercially available, purpose-built dataloggers make much of this type of programming transparent to the user through the use of datalogger program development software provided by the manufacturer.

Another major challenge that many projects and data managers face is how to consolidate data from a network of monitoring sites to a centralized location where they can be stored, archived, checked for quality, and then used for scientific analyses or shared with potential users (Rundel et al., 2009; Jones et al., 2017). Potential heterogeneity in the syntax and semantics of the data can complicate this step (Samourkasidis et al., 2018). Commercial sensing systems usually come with a proprietary software product that provides this functionality, whereas low-cost sensing systems are usually custom built and lack robust software that provides these capabilities. Finally, providing convenient methods for web-based access to visualize and download observational data for a variety of users whose technical skills may vary can also be challenging (Horsburgh et al., 2011; Demir and Krajewski, 2013; Muste et al., 2013; Mason et al., 2014) – yet these are basic capabilities needed for managing and sharing environmental sensor data, regardless of how they are collected.

In this paper we describe a web-based software application called the ODM2 Data Sharing Portal that was designed and developed to enable streaming of data from low-cost sensing systems deployed in the field to a centralized, web-based data repository. The specific driver for creating this software was to support data collection and management for a group of conservation organizations and citizen scientists in the Delaware River Watershed in the eastern United States who are deploying water quality monitoring sites using an Arduino-based, EnviroDIY Mayfly Data Logger Board™ [3] paired with low-cost hydrologic and water quality sensors (see Monitor My Watershed Data Sharing Portal Case Study). While the ODM2 Data Sharing Portal was built to support the emerging network of Arduino-based sensor nodes in the Delaware River Watershed, the architecture and components are generic, open-source, and could be deployed by other initiatives and groups needing a centralized data repository for environmental sensor data. Specific contributions of this work include an innovative, push-based architecture and simple messaging protocol that enables communications between a network of remote monitoring sites and a centralized data portal server. We describe our approach for storing and managing the sensor data and associated metadata, as well as techniques for producing high-performance, web-based visualization and access to the data. Finally, we provide an open-source implementation of the web portal and data management functionality we found necessary to support a community of citizen scientists in developing a network of low-cost environmental sensing stations.

[1] https://www.arduino.cc/

[2] https://www.raspberrypi.org/

[3] http://www.EnviroDIY.org

# MATERIALS AND METHODS

## Design and Overall Software Architecture

Our goal in developing the ODM2 Data Sharing Portal was to provide a system that could be used by citizen scientists to stream data from a variety of low-cost water quality sensing stations, such as those powered by Arduino-based EnviroDIY Mayfly dataloggers, to a centralized data repository where they could be stored, managed, and accessed by other members of the citizen science and water resources community. The following requirements motivated our implementation:

(1) Open-source software development to facilitate free deployment and software reuse;
(2) A graphical user interface (GUI) that supports users with varying levels of technical capabilities, including citizen scientists, and functions on a variety of computing platforms;
(3) Ability to register monitoring sites and configure the list of observed variables at each site;
(4) Support for a standard data/messaging protocol that enables pushing data from monitoring sites in the field using a variety of low-cost dataloggers to the centralized server;
(5) Persistent storage of the data pushed to the centralized server from monitoring sites;
(6) Ability to discover and access data from other monitoring sites within the network;
(7) Basic security to support ownership of registered sites and data and to avoid unwanted spam;
(8) Screening-level visualizations of data for monitoring the status of data collection sites;
(9) Access to download monitoring data in a comma-separated text format;
(10) Machine-to-machine data discoverability and access via web services; and
(11) Administrative functions for managing metadata for sensors, variables, and units.

The ODM2 Data Sharing Portal was designed as a web application with a web browser-based GUI. The overall Architecture of the software consists of a user interface layer, a web framework layer, a web service layer, and a data storage layer (**Figure 1**). In the following sections, we describe the high-level design of each of the architectural layers, their key components, and their basic functionality. In Section "Results" we describe a specific implementation of the ODM2 Data Sharing Portal for the Monitor My Watershed network of water quality monitoring sites, each of which uses an Arduino-based EnviroDIY Mayfly datalogger. Finally, in the "Discussion and Conclusions" section we discuss the capabilities of the system, some of the challenges we faced in our implementation, and how they were overcome.

### User Interface Layer

The user interface layer was implemented primarily using HTML5, cascading style sheets (CSS), and JavaScript, which function in all modern web browsers. This meets the requirements for operation across multiple computer operating systems as well as ensuring that functionality of the Data Sharing Portal is presented to users in a way that does not require specialized software installation. It also ensures that the functionality of the data sharing portal is available to users of varying technical capabilities. We assumed that most, if not all, potential users are familiar with using a web browser. A number of common and openly available front-end development tools were used to facilitate development of the web user interface (**Table 1**). We provide these here for completeness and to document the dependencies of the ODM2 Data Sharing Portal code.

The user interface of the ODM2 Data Sharing Portal consists of three main pages that are focused on meeting the functional requirements listed above. The "My Sites" page enables users to register new monitoring sites and manage their list of registered sites and followed sites. The "Site Details" page enables users to edit the metadata for a monitoring site and manage the list of variables measured at a site. The "Browse Sites" page is provided for discovering and accessing sites registered by other users. Additionally, pages are provided for creating a new user account, logging into the portal, and editing a user's profile. Finally, for administrative users of the system, an "Admin" page is provided for modifying lists of sensors, variables, and units presented to users when they are registering sites. Specific functionality of each of these pages is presented in the context of

**TABLE 1 |** Tools used for developing the web user interface of the Data Sharing Portal.

| Tool | Purpose |
|---|---|
| Material Design Lite URL: https://getmdl.io/ Version: 1.3.0 | Defines the theme of the website and offers a wide variety of easy-to-use components following Material Design principles. Provides a web adaption of Google's Material Design to provide users with familiar look and feel. |
| Bootstrap URL: https://getbootstrap.com Version: 4.0.0-beta.2 | Used mainly for its grid system that assists in page layout along with responsive design tools to ensure that pages work across a variety of devices including computers, tablets, and mobile phones. |
| jQuery URL: https://blog.jquery.com/2017/03/20/jquery-3-2-1-now-available/ Version: 3.2.1 | Library that simplifies development of JavaScript for document object model (DOM) manipulation and front-end logic, including event handling, animation, and Ajax calls. |
| d3.js URL: https://d3js.org Version: 4.13.0 | JavaScript library used for developing dynamic visualizations of data. Used to generate time series visualizations. |
| Font Awesome URL: http://fontawesome.io Version: 4.7.0 | Provides a set of high quality, vector icons to provide visual meaning to controls used throughout the website |
| Google Maps API URL: https://developers.google.com/maps/documentation/javascript/tutorial Version: Latest available | JavaScript API that provides customizable and dynamic maps. |

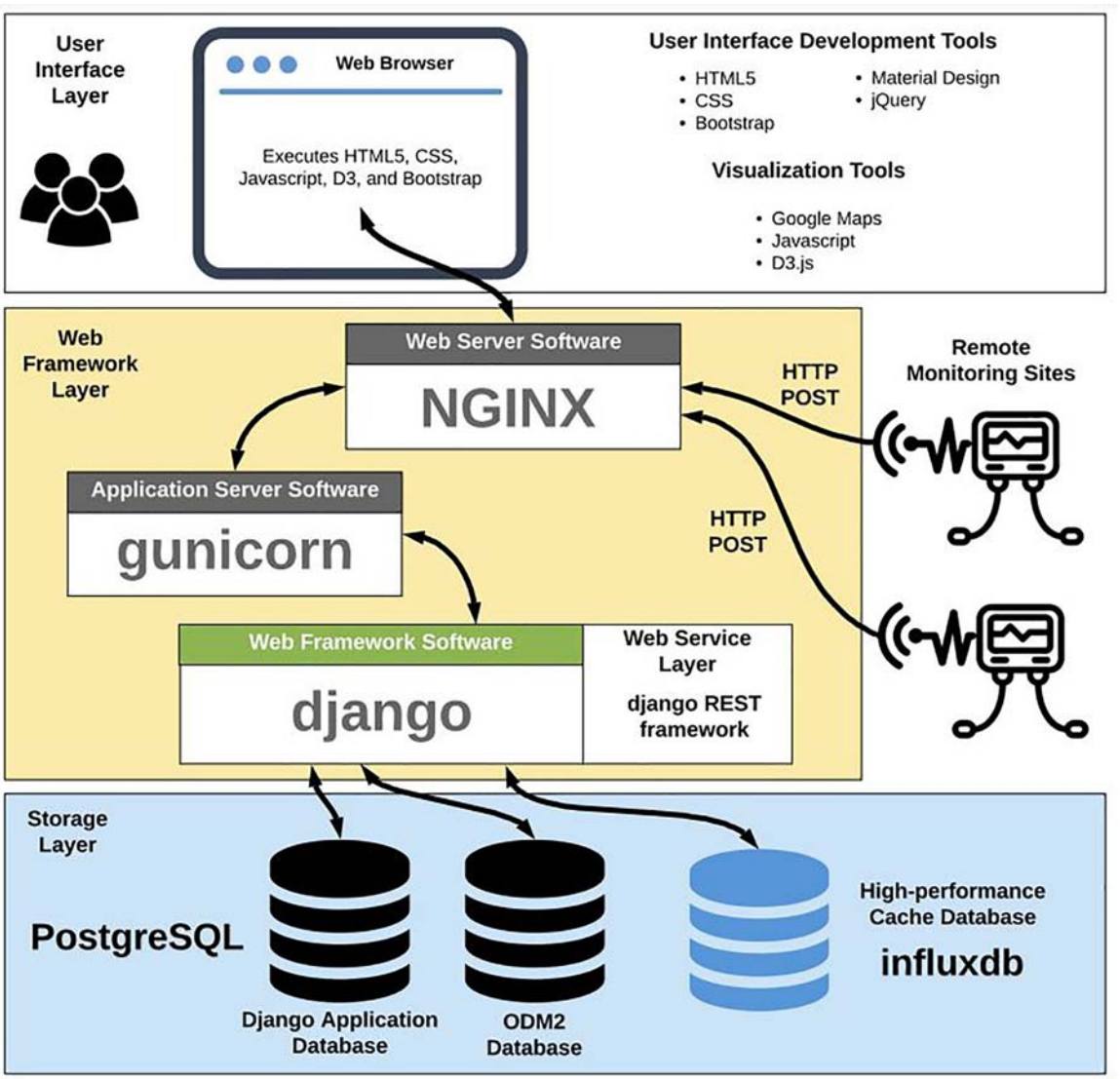



**FIGURE 1 |** Architecture of the Data Sharing Portal.

the Monitor My Watershed instance of the ODM2 Data Sharing Portal (see Results).

## Web Framework Layer

The ODM2 Data Sharing Portal was developed using the Python Django web framework[4]. We chose it over other frameworks because it is freely available, open-source, and supports rapid and straightforward development of common website functionality (e.g., user and account management, authentication, content management, etc.) using existing web components that are reliable, interchangeable, and scalable. Because it is Python based, it can be deployed on multiple server platforms (e.g., Linux or Windows) and can be used with a variety of web server software applications [e.g., Apache, NGINX, and Microsoft's Internet Information Services (IIS)]. These capabilities enable multiple options for deployment; however, for the ODM2 Data Sharing Portal, we targeted deployment of the Django Web Framework on an Ubuntu Linux server using a combination of the NGINX web server along with the Gunicorn app server. NGINX generally handles serving the static content of the ODM2 Data Sharing Portal website, whereas Gunicorn handles any web requests that must be dynamically generated. The combination of Django, NGINX, and Gunicorn is a common deployment environment for open-source web applications targeted for deployment on a Linux server.

## Data Storage Layer

The ODM2 Data Sharing Portal uses a combination of technologies in its storage layer. First, Django's Object-Relational Mapping (ORM) functionality is used along with an instance

[4]https://www.djangoproject.com

of PostgreSQL[5] to store Django's native database. Django uses its native database to store dynamic configuration data (e.g., users, sessions, permissions), along with other cached application data for faster access. In addition to Django's database, we also implemented an instance of ODM2 (Horsburgh et al., 2016) in PostgreSQL. ODM2 provides an extensive information model for storing observational data along with metadata describing monitoring sites, deployed sensors, observed variables and units, sensor depth/height, and individuals and organizations responsible for data collection, making it an obvious choice to serve as the back-end data store and archive for the Data Sharing Portal. We chose PostgreSQL for implementing the relational database components of the storage layer because it integrates well with Django's ORM functionality, provides robust and advanced relational database functionality, is Structured Query Language (SQL) compliant, and is freely available and open-source.

The final component of the data storage layer is a cache database that we implemented for providing high-performance data queries and time series data access. It is used in generating visualizations of the time series data for display on the website and for providing high-performance data download. The cache database was created in the InfluxDB time series database system[6], which is a high-performance data store written specifically for storing, managing, and real-time querying of timestamped data like those produced by environmental monitoring sites. Time series databases like InfluxDB have been used extensively with financial data, but have more recently been adapted for use in a variety of newer applications, including storing and managing high-resolution data resulting from monitoring of computational server systems and infrastructure (e.g., development operations or "DevOps") and storing and managing timestamped data from IoT applications. Time series databases are optimized for storage, summarization, and aggregation of timestamped data, along with handling time-dependent queries over large numbers of data values, making a time series database ideal for the data caching needs of the portal. InfluxDB is freely available and part of a set of open-source core components that also have commercial offerings.

### Web Service Layer

The primary function of the web service layer is to enable Internet-connected dataloggers to submit data to an instance of the Data Sharing Portal. We chose a push-based communication model where individual dataloggers push their data to the central repository for three main reasons. First, this negates the need for each individual datalogger to have a static and unique network or Internet Protocol (IP) address that can be consistently accessed by a centralized server. This is an important consideration because low-cost dataloggers may use a variety of hardware (e.g., Arduino versus Raspberry Pi) and a variety of means and service providers for connecting to the Internet (e.g., cellular, Wi-Fi, spread-spectrum radio, or a combination of these). Thus, they may not always have static IP addresses. We anticipated that it would likely be impossible for a centralized server to consistently connect to and pull data from all of the registered monitoring sites.

Second, the push model relies on the portal exposing a standard data submission interface to which remote dataloggers can push their data. With a standardized data submission interface, the portal needs only focus on receiving and acting upon requests from remote dataloggers and does not have to concern itself with making low-level device connections and mediating across communication protocols that may be inconsistent across different types of dataloggers. Indeed, reliance on a push model and a standardized data submission interface means that any Internet connected device or datalogger can push data to an instance of the Data Sharing Portal.

Last, the push model can result in significant power economy for low-power dataloggers deployed in the field because they do not have to stay awake to listen for pull requests from a centralized server. Each data collection device has full autonomy to send data to the server as often as it needs to and only when it needs to, which provides the owner of the datalogger with considerable flexibility in choosing data collection, recording, and transmission schedules that meet data collection needs while balancing power requirements.

Using the Django REST Framework[7], which is an extension of Django for building representational state transfer (REST) web services, we built a REST web service that enables any Internet-connected device to send data to an instance of the ODM2 Data Sharing Portal using standard HTTP POST requests. POST requests sent to the server are encoded using JavaScript Object Notation (JSON), and the portal returns standard HTTP responses (e.g., CREATED 201 when a POST request successfully creates new data in the portal's database) that can be interpreted by the datalogger to determine whether a request was successfully received and processed. As a simple security measure aimed at preventing unauthorized spam requests to the web service, we implemented a token-based authorization system for web service requests. Each registered data collection site is assigned a unique identifier and an authorization token visible only to the site owner. Each web service request received by the portal is first checked to make sure that a valid authorization token is provided and that it matches the identifier of the site in the request. Any requests with invalid tokens or mismatched tokens and site identifiers are automatically ignored. The JSON format for POST requests and the syntax of tokens and identifiers used in the messages are described in more detail in Section "Results."

## RESULTS

### Monitor My Watershed Data Sharing Portal Case Study

In this section, we describe a production instance of the ODM2 Data Sharing Portal software for the Monitor My Watershed®network[8]. Monitor My Watershed is an evolving program for conservation organizations, citizen scientists, and

[5]https://www.postgresql.org/

[6]https://www.influxdata.com/time-series-platform/influxdb/

[7]https://www.django-rest-framework.org/

[8]http://MonitorMyWatershed.org

students that bridges science, technology, engineering, and mathematics (STEM) by incorporating open-source hardware and software, environmental monitoring, ecosystem science, and data analysis and interpretation (Bressler et al., 2018; Ensign et al., 2019). The program is multi-faceted, with goals to (a) enhance knowledge and stewardship of fresh water and other natural resources, (b) increase citizen access, use, collection, and sharing of environmental data, (c) increase STEM literacy, and (d) develop methods, protocols, curricula, and workshop materials to support STEM educators and programs.

A core component of Monitor My Watershed is a network of monitoring sites deployed by participants using EnviroDIY Mayfly dataloggers (see Mayfly Loggers and the EnviroDIY Modular Sensors Library). The ODM2 Data Sharing Portal described in this paper was developed to capture, manage, and provide access to environmental monitoring data from these DIY devices and for aquatic macroinvertebrate data that are part of the Leaf Pack Network® stream ecology program. These online tools are part of a broader set of digital tools available at https://WikiWatershed.org, that are designed to support researchers, conservation practitioners, municipal decision-makers, educators, and students that are interested in water resources and environmental stewardship.

In 2010, a research team at the Stroud Water Research Center started developing and deploying open-source hardware and software devices to build autonomous water quality monitoring stations with real-time data telemetry. The primary motivation was to reduce costs in order to increase the spatial resolution of data for various research studies by deploying more measurement sites in streams and rivers in the Delaware River Watershed (and elsewhere). The team realized the potential for these devices to be useful for both the greater research community and also the watershed conservation community, launching the EnviroDIY website[9] in 2013 to share their approaches and encourage a community of contributors to share their DIY technology for environmental monitoring, find resources, or pose questions to other users. In 2014, the William Penn Foundation funded a training and support program as an expansion of EnviroDIY under the umbrella of the Delaware River Watershed Initiative (DRWI). The DRWI is a multi-year effort supporting conservation organizations working to protect and restore stream health in the Delaware River Watershed (Freedman et al., 2018; Johnson et al., 2018).

The DRWI effort, among others, has led to the development of the tools described herein that support the use of low-cost, open-source, and low-power devices for monitoring environmental conditions, in particular sensors collecting data on water level, water temperature, specific conductivity, turbidity, dissolved oxygen, and other water quality and meteorological sensor arrays. Most of the sensors in use are commercially available, bare-wire devices that can be programmed to communicate with Arduino compatible devices like the EnviroDIY Mayfly Data Logger Board. Today, there are hundreds of devices deployed throughout the Delaware River Watershed with the help of more than 50 non-profit organizations, and hundreds of registered members using https://EnviroDIY.org as a social networking website to share their DIY technology.

[9]https://EnviroDIY.org

## Mayfly Loggers and the EnviroDIY Modular Sensors Library

Participants in the Monitor My Watershed network are using Arduino-based EnviroDIY Mayfly dataloggers[10] to deploy their water quality monitoring sites. The EnviroDIY Mayfly is a user-programmable microcontroller board specifically designed to meet the needs of solar-powered, wireless environmental data logging. It uses an ATmega 1284p processor and is fully compatible with the Arduino interactive development environment (IDE) software. In addition to a more powerful processor, it has enhanced flash memory for storing larger datalogging programs, or sketches, along with additional RAM, additional input pins for sensors, a real time clock, an onboard MicroSD memory card socket, an XBee module socket for integration of communication peripherals, and a solar charge regulator. These hardware enhancements, which grew from the need for options to better enable low-cost and low-power environmental monitoring, make the EnviroDIY Mayfly a more capable datalogger when compared to many other Arduino boards. The EnviroDIY Mayfly is commercially available for purchase at a cost of $60 USD via Amazon, and hardware designs, code examples, and documentation are openly available in the Mayfly GitHub repository[11]. The relatively low cost and open nature of the EnviroDIY Mayfly design made it an ideal platform on which to build the citizen science monitoring efforts of the Monitor My Watershed network.

Do-it-yourself practitioners generally find rapid success at reading data from simple sensors to an EnviroDIY Mayfly or other Arduino board. However, it is much more challenging to program an Arduino to perform all of the required functions of a solar-powered monitoring station that collects data from several environmental sensors, saves observations to a MicroSD card, transmits data to a public server like the ODM2 Data Sharing Portal, and puts the sensors to sleep to conserve power between logging intervals. To make this easier for citizen scientists and other potential users, we developed the EnviroDIY Modular Sensors Arduino code library[12] to support wireless, solar-powered environmental data logging applications. The Modular Sensors library coordinates these tasks by "wrapping" native sensor code libraries and other well-developed IoT code libraries into simplified, high level functions with unified conventions for arguments and returns. These wrapper functions also serve to harmonize the process of iterating through the powering up and logging of data from a diverse set of sensors and variables, avoiding code conflicts and minimizing power consumption. In addition, the library supports saving data to a MicroSD memory card, transmitting data wirelessly to an instance of the ODM2 Data Sharing Portal, and putting the processor, sensors, and peripherals to sleep to conserve power. Example code sketches included in the library were designed to serve as a sort of menu

[10]https://www.EnviroDIY.org/mayfly/

[11]https://github.com/EnviroDIY/EnviroDIY_Mayfly_Logger

[12]https://github.com/EnviroDIY/ModularSensors

of options, where users select the options they need for their specific monitoring site along with specifying their site-specific configuration (i.e., unique registration token, site identifier, and variable identifiers) after registering their site with the portal. Last, a Wiki provides extensive documentation[13] and a tutorial guide for first-time users[14].

It is beyond the scope of this paper to describe all of the functionality of the Modular Sensors Library. However, the following high-level functions, which are called within an Arduino datalogging sketch, are the basis for enabling the communication between an Internet connected EnviroDIY Mayfly and an instance of the ODM2 Data Sharing Portal:

- *setToken( )*: Sets the unique registration token for the monitoring site to the value assigned by the ODM2 Data Sharing Portal when the site was registered in the portal.
- *setSamplingFeatureUUID( )*: Sets the universally unique identifier for the monitoring site assigned by the ODM2 Data Sharing Portal when the site was registered.
- *completeUpdate( ):* Powers all sensors and iterates through them requesting values as soon as they become available from the sensor. After all values from an individual sensor are received, that sensor is put to sleep and powered down. The iteration attempts to minimize the amount of time all sensors are active in order to conserve battery life.
- *printSensorDataJSON( )*: Prints a properly formatted JSON string containing the site's registration token, sampling feature identifier, variable identifiers, observation time stamp, and numeric values of the observed variables.
- *postDataEnviroDIY( )*: Creates the proper HTTP headers and sends a JSON string to an instance of the ODM2 Data Sharing Portal as an HTTP POST request. Returns an HTTP response code from the ODM2 Data Sharing Portal indicating whether the POST request was successful.

## POSTing Data to the Portal

HTTP POST requests containing observation data values can be sent to an instance of the ODM2 Data Sharing Portal with any desired temporal frequency. Requests include HTTP headers and a JSON-encoded body (**Figure 2**). The registration token in the header serves as the authentication for the ODM2 Data Sharing Portal to ensure that the POST request is from a valid, registered site. Within the body, the sampling feature and variable identifiers are used by the ODM2 Data Sharing Portal to match the data values in the POST request with the correct monitoring site and variable in the database. Any data values within a POST request that have a valid registration token and sampling feature identifier but invalid variable identifiers are ignored. Any number of observed variables can be sent with an individual POST request, but each is associated with a single sampling feature and single timestamp that identifies the data and time at which the values were recorded by the datalogger. Timestamps are encoded using the International Standards Organization (ISO) 8601 standard for encoding date and time strings (International Standards Organization [ISO], 2004).

[13]https://github.com/EnviroDIY/ModularSensors/wiki

[14]https://envirodiy.github.io/LearnEnviroDIY/

```
POST /api/data-stream/ HTTP/1.1
Host: www.monitormywatershed.org
TOKEN: 0cd0616f-cf03-4789-aa28-82bca1b847f1
Content-Type: application/json

{
  "sampling_feature": "f319af6a-3091-4070-b3ad-a606a7fdbed4",
  "timestamp": "2016-12-08T14:45:01-07:00",
  "f8fbf90e-f59d-4736-af66-91fbee455433": 8.3,
  "52e6d5ce-eca1-4545-9b01-607a487cbfc0": 10.54
}
```

**FIGURE 2 |** Format of the HTTP POST requests sent to an instance of the Data Sharing Portal to post data. In this example, numeric values are specified for two measured variables.

## Portal Deployment on Server Hardware

Although all of the ODM2 Data Sharing Portal components could be installed on the same Linux server, for performance and security reasons, the Monitor My Watershed instance of the ODM2 Data Sharing Portal was deployed on two separate virtual machines running within a VMWare ESXi virtualization environment. The first machine serves as the web server for the portal website and web services. The second machine is a dedicated database server. This separation of concerns ensures that processor intensive tasks on the database server do not slow the web server down and affect the user experience. It also allowed us to keep the database server behind institutional firewalls to limit the surface area for potential security issues.

Both machines were created using Ubuntu Linux Version 16.04[15], which was the latest version available at the time the machines were built and is freely available for download (the latest version available for download is 18.04). The web server was allocated four processor cores and eight GB of RAM, while the database server was allocated six cores and 16 GB of RAM. In monitoring these machines, the allocated resources have been more than adequate to serve the needs of the Monitor My Watershed network, with processor and memory usage of each machine generally being well below 25%.

## Graphical User Interface

### *My sites: registering and managing monitoring sites*

The My Sites page (**Figure 3**) consists of a map-based display of all of the monitoring sites that a user has registered within the portal along with access to view the details of each individual registered site via the Site Details page (described below). Users can register new sites on this page by filling in a form with the new site's descriptive metadata, including the site's geographic location. The descriptions of existing sites can be edited using this same form. To enhance the sharing aspects of the portal, we also added to the My Sites page a list of sites that the user is following. Followed sites are those registered by other users of the portal that the current user finds interesting or useful. Following

[15]https://www.ubuntu.com/

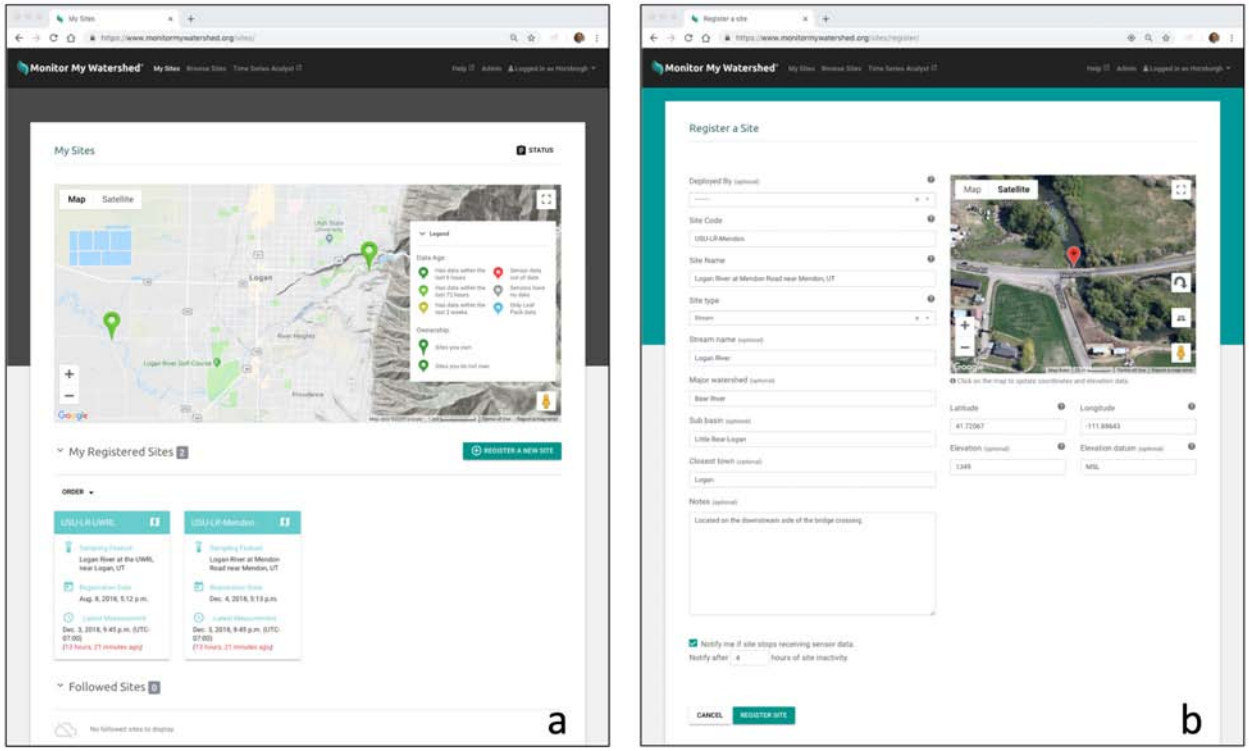



**FIGURE 3 |** Detailed views of the My Sites page of the Data Sharing Portal. **(a)** Map view of sites the user has registered and the cards for registered sites. **(b)** Form for creating a new site or editing an existing site.

a monitoring site is initiated by clicking on a check box on the Site Details page for any site registered within the system.

Once a user has created a new site, the list of sensors deployed at that site and the list of measured variables and their units can be configured on the Site Details page. Users can also opt to be notified by the portal if it stops receiving sensor data for that site. When this option is selected, the user will be alerted via email when the portal does not receive any new data for the site for more than a configurable number of hours. The data alerts were implemented as a Django script that runs on the web server and is scheduled as a cron job to run every 15 min.

### *Site details: adding and managing sensors and observed variables*

The Site Details page (**Figure 4**) provides a public view of the descriptive metadata for a monitoring site. For the owner of the site, it provides options for editing the site description, managing sensors and observed variables for the site, viewing and downloading data for the site, configuring the site to share its data to HydroShare (see Integration With the CUAHSI HIS and HydroShare), and deleting the site. Editing the site's description and deleting the site can be done by the site's owner by clicking buttons at the top of the page. When a user chooses to delete a site from the portal, that site and all of its associated sensor data are removed from the portal and its databases. Given that users create the data uploaded to the portal, we opted to enable them to delete the data. However, we also provided users with a mechanism for permanently preserving their data in an open data repository (see Integration With the CUAHSI HIS and HydroShare).

The unique identifiers associated with a site, including its registration token and its sampling feature ID, along with the unique identifiers for each of the measured variables are displayed on the page as well as via a pop-up window that makes it convenient for the user to copy the identifiers and paste them into their Arduino (or other) datalogger program for that site. To protect the security of a registered site, these codes are only displayed to the site's owner. Users that do not own the site can view the site's metadata, access and download the data, and choose an option to follow the site, which adds that site to a section in their My Sites page.

Toward the bottom of the Site Details page, users are presented with metadata about each variable measured at that site and screening-level visualizations of the data. Each measured variable is displayed on a card with the most recent data value shown and a sparkline plot showing the latest 72 h of data. The background of the sparkline plot is colored to indicate the age of the most recently received data value. Plots shaded green have reported data within the last 72 h and plots shaded red have not. This is a simple and quick indication of both data quality and age for users that can give at-a-glance information about whether a site is reporting data (based on the shading of the sparkline plot) and whether a sensor may be malfunctioning (based on the last reported value and the values shown in the sparkling plot). Each of the variable cards also includes a link to display the last 72 h

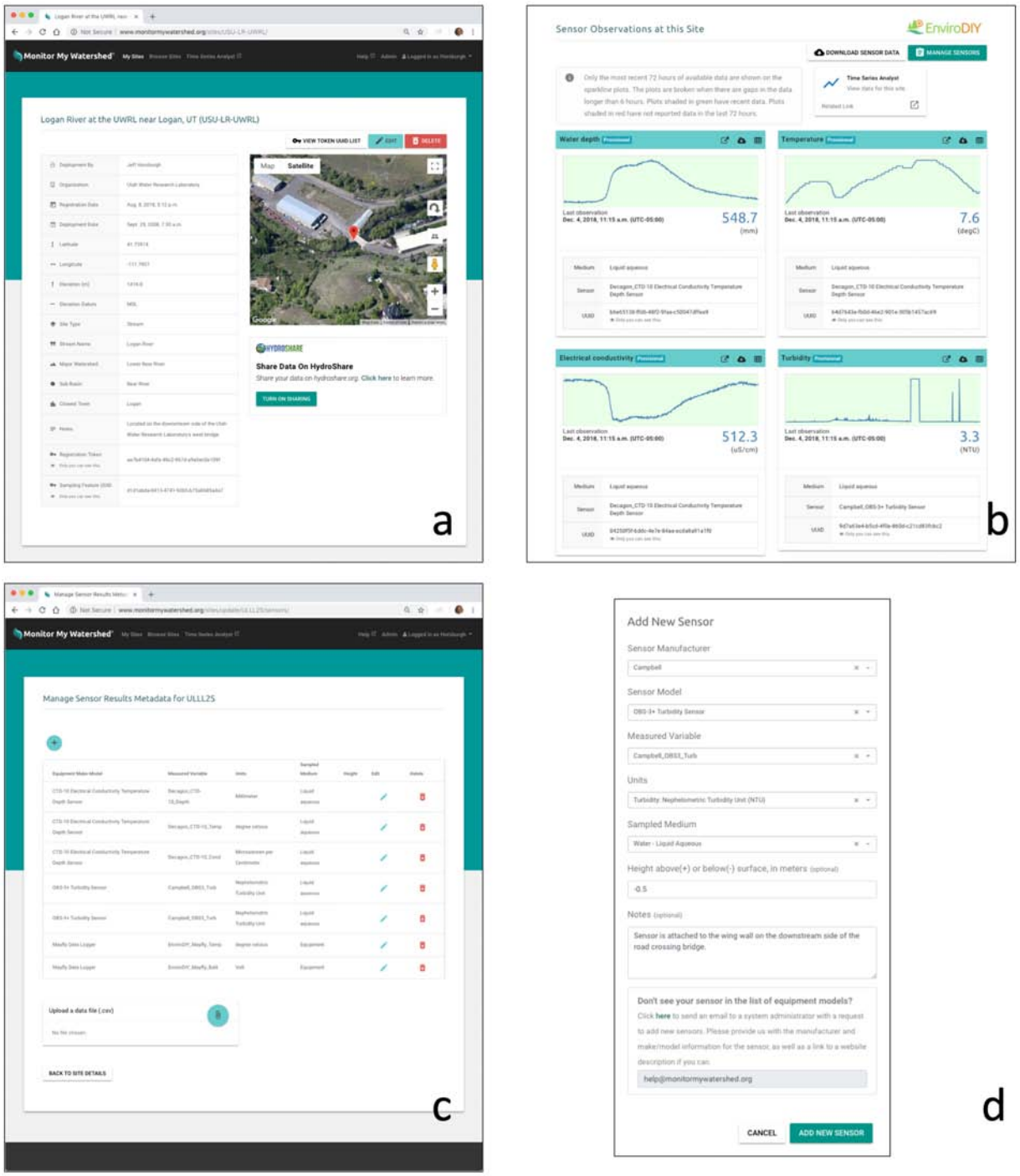



**FIGURE 4 |** Detailed views of the Site Details page of the Data Sharing Portal. **(a)** Detailed metadata for a registered site. **(b)** Sparkline visualizations of data for each of the variables measured at a site. **(c)** Interface for managing sensors and measured variables at a site. **(d)** Window for adding or editing a sensor for a site.

of data in a tabular view so individual values can be inspected as well as a link to download a comma-separated text file for all of the recorded data for that variable. An additional link is provided to download a single comma-separated values text file containing the data for all measured variables at that site.

Users can manage the list of measured variables at a site by clicking the Manage Sensors button. A new measured variable can be added by selecting options from pre-populated lists of sensor manufacturers, sensor models, measured variables, and units. Additionally, the user can select the environmental

medium in which the sensor is installed (e.g., air, water, sediment) and can optionally specify a height above or below the surface to enable installation of multiple sensors making simultaneous measurements at a single site, but at different heights or depths (e.g., multiple temperature sensors installed at different depths in a water column). We chose to have citizen scientists choose from pre-populated lists of sensors, variables, and units because our experience has shown that this significantly simplifies the entry of metadata describing the observed variable and ensures that metadata for all sensors and measured variables are complete and consistent. The tradeoff is that administrators of the portal must add the lists of sensor manufacturers, sensors, measured variables, and units to their instance of the ODM2 Data Sharing Portal before they can be used (see Administrative Functions). Where users wish to add a sensor, measured variable, or use units that do not already exist in the drop-down lists, the "Add New Sensor" form provides an email address for contacting an administrator of the system to get them added. Users can edit existing measured variables at a site and delete them, which removes that measured variable and any associated data from the portal's underlying databases.

As a final option under managing sensors and measured variables, users can upload a comma separated values text file containing sensor data to be parsed into the portal's databases. This option is important because it enables users to upload data to the portal under circumstances where communications are lost at a monitoring site making it impossible to send data via HTTP POST requests, or where sites are simply operated without a telemetry connection but with periodic data downloads (e.g., for remote sites with no nearby cellular data network). A Django script parses uploaded data files, compares the data from the file to data within the portal's database, and adds any new data from the file to the portal. Any data in the file that already exists in the portal's database is ignored. We modeled the format of the upload data file (**Figure 5**) after the file format captured on the MicroSD card by datalogger programs built using the Modular Sensors Library to ensure that users could easily download data files from their datalogger's MicroSD card and then upload them directly to the portal. However, these files can also be constructed using code, in a text editor, or via Microsoft Excel (e.g., in the case a user wants to upload historical data for a site).

The first column of the data file contains the timestamp in ISO 8601 format. Each subsequent column in the file contains the numeric data values for one measured variable at the site. The first line of the file contains the universally unique sampling feature identifier in the first column, and then each subsequent column contains the unique identifier for the measured variable whose numeric values appear in that column. Files can contain any number of measured variable columns and any number of rows of data. Additional header rows are allowed at the top of the file, but are ignored by the data loading script.

### *Browse sites: discovering and accessing data*

To enhance the sharing aspect of the portal, public access to the Site Details page for each site registered within the portal is provided via the Browse Sites page (**Figure 6**). A Google Maps-based interface is provided that shows the location of all monitoring sites registered with the portal. Sites are indicated on the map with markers that display site ownership (i.e., sites the user owns are shown with a different symbol than sites owned by other users) and the age of the data available at the site (i.e., sites having data within the past 6 h are colored green, whereas sites with data older than 2 weeks are colored red). Users can search sites using the search box at the top of the map, which performs a keyword search on the Site Code and Site Name metadata fields across all sites. Users can also browse sites by entering filter criteria in the faceted browsing panel on the left of the window to search sites by data type, organization, and site type. When search criteria are entered, the map view is automatically zoomed to the extent of sites that meet the specified criteria. Clicking on

| Sampling Feature: bffbd411-27d7-48a6-bc8c-fb4edd0c628512ba4d-9516-414f-93dd-db7281d8d76 5825859-fd69-4c63-9db2-5759641ab3e1 | | |
|---|---|---|
| Data Logger: CMP01_Mayfly_170292 | Campbell_OBS3_Turb | YosemiTech_Y502-A_DOmgL |
| Data Logger: CMP01_Mayfly_170292 | turbidity | oxygenDisssolved |
| Data Logger: CMP01_Mayfly_170292 | nephelometricTurbidityUnit | milligramPerLiter |
| Date and Time in UTC-6 | Campbell_OBS3_Turb | YosemiTech_Y502-A_DOmgL |
| 2008-09-30 00:00:00 | 31.65 | 7.751076985 |
| 2008-09-30 00:30:00 | 34.51 | 7.696071275 |
| 2008-09-30 01:00:00 | 30.51 | 7.649400566 |
| 2008-09-30 01:30:00 | 29.76 | 7.596061856 |
| 2008-09-30 02:00:00 | 32.77 | 7.556057147 |
| 2008-09-30 02:30:00 | 32.42 | 7.521051437 |
| 2008-09-30 03:00:00 | 33.64 | 7.471045728 |
| 2008-09-30 03:30:00 | 33 | 7.436041018 |
| 2008-09-30 04:00:00 | 31.58 | 7.427703309 |
| 2008-09-30 04:30:00 | 32.87 | 7.421031599 |
| 2008-09-30 05:00:00 | 33.17 | 7.41935889 |
| 2008-09-30 05:30:00 | 31.08 | 7.42435418 |
| 2008-09-30 06:00:00 | 30.23 | 7.44601547 |
| 2008-09-30 06:30:00 | 30.16 | 7.474344761 |

**FIGURE 5 |** Format of the data upload file showing a brief snippet of data for a site measuring two variables. Files are stored in comma separated value (CSV) format. Comma separators have been omitted from this view of the file for clarity.

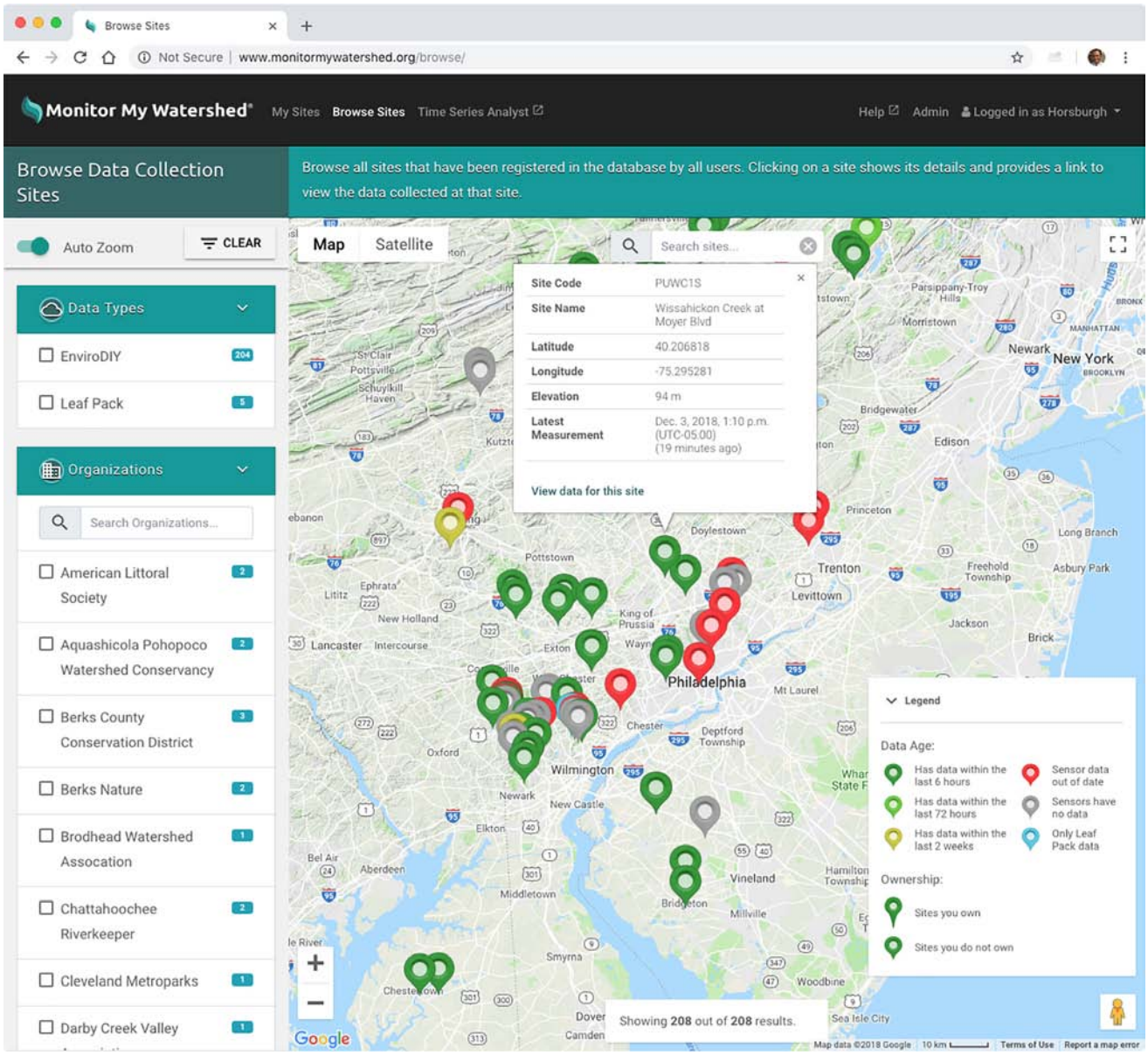



**FIGURE 6 |** Detailed view of the Browse Sites page of the Data Sharing Portal.

a site marker on the map shows a pop-up window with basic metadata about that site. Included is a link to "View data for this site," which opens the public view of the Site Details page for the selected site.

## Administrative Functions

Because the primary focus of the ODM2 Data Sharing Portal was citizen science and DIY users, we chose to simplify the input of metadata about sensors, measured variables, and units so that users could select from predefined lists that were already populated within the system. This proved effective at ensuring that the metadata descriptions created by users were complete. However, doing so required that we keep the list of sensors, measured variables, and units up to date. To avoid modifying the code of the portal or requiring low-level database edits every time a new sensor or variable needed to be added, we used Django's automatic admin interface to create this functionality for a small number of system administrator users. When users with admin rights log into the portal, they can access the admin functionality using a link in the main title bar. This exposes a simple set of Django admin pages for creating new sensors, measured variables, and units. These pages add newly created items to Django's native ORM database, which means that once they are created by an administrator, they are automatically available for use within the portal.

### Integration With the CUAHSI HIS and HydroShare

To best serve the needs of the conservation and environmental science communities for data discoverability, accessibility, and archiving, we enabled automated data exchange with the Water Data Services managed by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI). This ensures that the portal is not a stove pipe for contributed data (i.e., we wanted users to be able to get their data into and out of the portal). To enable machine-to-machine communication of data, we deployed the WaterOneFlow for Python (WOFPy)[16] web services on the portal's web server and registered them with the CUAHSI Hydrologic Information System (HIS) (Horsburgh et al., 2009, 2010). The WOFPy services connect directly to the portal's ODM2 PostgreSQL database, serve site and time series level metadata to the central HIS metadata catalog, and serve time series of data values using the WaterOneFlow web service methods standardized by the CUAHSI HIS. Data values are delivered over the web in a standardized extensible markup language (XML) encoding called Water Markup Language (WaterML) (Zaslavsky et al., 2007). By doing so, we made all Monitor My Watershed data searchable and accessible via CUAHSI's data client application[17] and all other WaterOneFlow/WaterML client applications.

We also connected the Monitor My Watershed Data Sharing Portal to HydroShare[18], which is a file-based data sharing and publication system operated by CUAHSI (Horsburgh et al., 2015). This allows users to connect their Data Sharing Portal account profile to their HydroShare account and then sync their data from the portal to HydroShare either on demand or on a scheduled basis with a user-configurable frequency. When a user chooses to connect a monitoring site in the portal with HydroShare by turning on sharing via the Site Details page, all of the time series measured at that site are converted to a comma-separated text file (one per variable) with a detailed metadata header and uploaded to a HydroShare resource using HydroShare's web service application programming interface (API). This enables users to easily move all of their sensor data to an open data repository that offers broader data sharing and formal data publication [i.e., HydroShare issues a citable digital object identifier (DOI) for published datasets and makes them immutable]. These automated data exchanges, with federally supported data cyberinfrastructure and using established environmental data standards for interoperability, distinguish the ODM2 Data Sharing Portal over other IoT data systems.

## DISCUSSION AND CONCLUSION

The combination of functionality provided by the ODM2 Data Sharing Portal meets many of the most common needs for streaming environmental sensor data to the web and all of the requirements we identified for a citizen science and DIY environmental data portal aimed at low-cost sensing. Users' ability to register new data collection sites, describe which data are being collected using the robust metadata model provided by ODM2, and manage their list of registered sites using a web-based GUI enables them to begin logging data from a monitoring site after some basic training. Map-based browsing and display of registered monitoring locations, the faceted browsing interface, and visualization of sites on a map by the age of collected data provide a dashboard for users to monitor the health of their sites and to discover sites and data collected by others. No specialized software or expertise are required to use these tools, which was important for our use case and significantly lowers the bar for getting started with data collection and for accessing the resulting data. More technical users can export selected datasets in a CSV text file format for more sophisticated analyses or visualization in separate data analysis software.

Because the ODM2 Data Sharing Portal uses standard HTTP POST requests for streaming data from the field to the web, any Internet connected device capable of making measurements and formulating an HTTP POST request can send those observations to an instance of the ODM2 Data Sharing Portal. This met our needs in supporting the network of Arduino-based dataloggers in the Delaware River Watershed, each of which sends an HTTP POST request to insert its data into the portal as new data are collected. It also enabled us to insert data from data collection sites that existed before the ODM2 Data Sharing Portal came online via Python scripting to ensure that historical data for existing sites were not lost. Additionally, since the capabilities of the Monitor My Watershed instance of the ODM2 Data Sharing Portal are not specific to the network of sites within the Delaware River Watershed, the network of monitoring sites registered with the Monitor My Watershed website has now grown well beyond the boundaries of the Delaware River Watershed, with more than 190 registered monitoring sites from nearly 70 contributors affiliated with more than 50 organizations, totaling more than 78 million data values at the time of this writing.

While the ODM2 data model proved to be capable of storing the needed metadata for describing monitoring sites, sensors, measured variables, etc., we were unable to obtain acceptable performance for all of the data management, visualization, and download capabilities of the portal website using only an ODM2 database implemented in PostgreSQL. Performance of functionality for generating the screening-level sparkline visualizations and CSV download files on demand for users proved to be unacceptably slow when the number of measured variables at a site grew beyond three to four and when the number of observations for each variable grew beyond a few thousand records. These performance limitations drove our implementation of the high-performance data cache using InfluxDB. When data POST requests are received by an instance of the ODM2 Data Sharing Portal, the new data values are written to both the ODM2 database in PostgreSQL and to the data cache in InfluxDB. Any functionality that needs high-performance access to data values gets them from InfluxDB. Any functionality that requires access to detailed metadata about a site, observed variables, sensors, etc. queries that information from the ODM2 PostgreSQL database. The ODM2 PostgreSQL database also serves as the definitive, archival version of the

[16]https://github.com/ODM2/WOFpy

[17]http://data.cuahsi.org

[18]http://www.hydroshare.org

data from which the InfluxDB cache can be reconstructed at any time if needed. By keeping the PostgreSQL database, we preserved the ability to perform expressive queries using the full syntax of SQL (as opposed to the "SQL-like" query language provided by InfluxDB) on the metadata stored in the ODM2 database. We also maintained much simpler support for enforcing metadata constraints and business rules (e.g., enforcing required versus optional metadata elements) that would have been harder to implement using the unstructured metadata approach of InfluxDB. Other approaches for high performance access to data values could have been investigated, including using materialized views in PostgreSQL or the TimescaleDB extension for PostgreSQL. However, our use of InfluxDB provided the performance and scalability that we needed.

HTTP and REST web services are ubiquitous on the web, integrated well with our chosen development architecture (Python, Django, and the Django REST Framework), and met our communication needs for the first releases of the ODM2 Data Sharing Portal. However, there are disadvantages to this approach – mainly the "overhead" size of HTTP POST requests relative to the volume of data contained within them. This overhead increases the volume of cellular data consumed by a datalogger, which can increase operating costs for monitoring sites using cellular modems. It can also increase the daily electrical power requirements for the monitoring site devices (i.e., a shorter radio pulse requires less power to transmit). We are now investigating potential enhancements to the ODM2 Data Sharing Portal, including enabling the use of Message Queue Telemetry Transport (MQTT) as a communication protocol. MQTT is increasingly used by IoT applications due to its smaller footprint and lower bandwidth consumption. Other potential enhancements under consideration for the Data Sharing Portal include automating and streamlining the entry of site and sensor metadata to avoid redundancy and streamline the process, more advanced tools to support quality assurance for submitted data (e.g., automated value range checks), additional tools for data visualization, and the addition of capabilities for post processing and quality control of submitted data.

The ODM2 Data Sharing Portal was developed over a period of multiple years and has had eight major releases to date. We have received input and feedback about functionality from researchers working on the project and participating DIY users and citizen scientists that we have used to refine the design and functionality of the site. Although the ODM2 Data Sharing Portal was conceptualized and initially implemented for the Monitor My Watershed network of monitoring sites, it was designed for and can be adapted for potential reuse. The components we used in developing the portal are all freely available, and the source code for the portal is shared on GitHub[19] under the liberal BSD-3 open source license. To deploy a new instance of the portal to support a different project or data collection network, users would need to procure the necessary server infrastructure (either physical or virtual), modify the styling of the site to suit their needs by replacing logos and modifying the CSS, and then deploy the software. Directions for deploying the data sharing portal software are provided in the GitHub repository. We anticipate that the ODM2 Data Sharing Portal software and/or the methods we used in its design and development may be useful for other organizations that need to provide capabilities for streaming environmental sensor data along with public visualization and data access capabilities for conservation, citizen science, or research efforts.

[19]https://github.com/ODM2/ODM2DataSharingPortal

## SOFTWARE AVAILABILITY

The software described in this paper includes the ODM2 Data Sharing Portal and associated web services for enabling upload of sensor data from Internet connected devices and the Modular Sensors Arduino library. All of the source code for the ODM2 Data Sharing Portal and related web services is available for download via the GitHub repository at https://github.com/ODM2/ODM2DataSharingPortal. The most recent release for the portal software at the time of this writing was Version 0.9.5 and is available via Zenodo (Caraballo et al., 2019). The production instance of the Monitor My Watershed Data Sharing Portal is available at http://MonitorMyWatershed.org. Code for the Modular Sensors Arduino library is available at https://github.com/EnviroDIY/ModularSensors, with the latest release for the library at the time of this writing being Version 0.17.2.

## AUTHOR CONTRIBUTIONS

JH was the main author of the manuscript. JH and AA co-architected the ODM2 Data Sharing Portal software, including design and specifications, with assistance from DA. JC developed the back-end code, databases, and server infrastructure for the ODM2 Data Sharing Portal software and the Monitor My Watershed implementation. MR developed the front-end code and graphical user interface for the Data Sharing Portal software and Monitor My Watershed implementation. SD is the primary author of the Modular Sensors Arduino library. AA, SD, and DA performed extensive testing of the portal. All authors contributed to writing text and editing the manuscript.

## FUNDING

Funding for this work was provided by the William Penn Foundation under grant 158-15. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the William Penn Foundation.

## ACKNOWLEDGMENTS

We are gratefully acknowledged the work and contributions of the EnviroDIY community and those who participated in testing and advancing the Monitor My Watershed Data Sharing Portal software.

## REFERENCES


Abu-Elkheir, M., Hayajneh, M., and Abu Ali, N. (2013). Data management for the internet of things: design primitives and solution. *Sensors* 13, 15582–15612. doi: 10.3390/s131115582

Adu-Manu, K. S., Tapparello, C., Heinzelman, W., Apietu Katsriku, F., and Abdulai, J. (2017). Water quality monitoring using wireless sensor networks: current trends and future research directions. *ACM Trans. Sens. Netw.* 13, 1–41. doi: 10.1145/3005719

Baker, E. (2014). Open source data logger for low-cost environmental monitoring. *Biodivers. Data J.* 2:e1059. doi: 10.3897/BDJ.2.e1059

Beddows, P. A., and Mallon, E. K. (2018). Cave Pearl data logger: a flexible Arduino-based logging platform for long-term monitoring in harsh environments. *Sensors* 18:530. doi: 10.3390/s18020530

Bressler, D., Jackson, J. K., Ehrhart, M., and Arscott, D. B. (2018). Building a more scientifically informed community in the Delaware River Basin. *Narrat. Inq. Bioeth.* (in press).

Caraballo, J., Ramirez, M., Horsburgh, J. S., Blackburn, C., Fryar-Ludwig, K., Aufdenkampe, A. K., et al. (2019). *ODM2 Data Sharing Portal: Release 0.9.6 (Version v0.9.6), Zenodo*. Available at: http://doi.org/10.5281/zenodo.2563182 (accessed February 12, 2019).

Demir, I., and Krajewski, W. F. (2013). Towards an integrated flood information system: centralized data access, analysis, and visualization. *Environ. Model. Softw.* 50, 77–84. doi: 10.1016/j.envsoft.2013.08.009

Ensign, S., Arscott, D., Hicks, S., Aufdenkampe, A., Muenz, T., Jackson, J., et al. (2019). A digital mayfly swarm is emerging. *EOS* 100. doi: 10.1029/2019EO116611

Faustine, A., Mvuma, A. N., Mongi, H. J., Gabriel, M. C., Tenge, A. J., and Kucel, S. B. (2014). Wireless sensor networks for water quality monitoring and control within lake victoria basin: prototype development. *Wireless Sens. Netw.* 6, 281–290. doi: 10.4236/wsn.2014.612027

Ferdoush, S., and Li, X. (2014). Wireless sensor network system design using Raspberry Pi and Arduino for environmental monitoring applications. *Procedia Comput. Sci.* 34, 103–110. doi: 10.1016/j.procs.2014.07.059

Freedman, P., Arscott, D. B., Haag, S., and Hall, K. (2018). Data and models help define effective strategies in the Delaware River Watershed Initiative. *Water Resour. Impact* 20, 19–21.

Gualtieri, G., Camilli, F., Cavaliere, A., De Filippis, T., Di Gennaro, F., Di Lonardo, S., et al. (2017). An integrated low-cost road traffic and air pollution monitoring platform to assess vehicles' air quality impact in urban areas. *Trans. Res. Procedia* 27, 609–616. doi: 10.1016/j.trpro.2017.12.043

Hart, J. K., and Martinez, K. (2006). Environmental sensor networks: a revolution in the earth system science? *Earth Sci. Rev.* 78, 177–191. doi: 10.3390/s18072248

Horsburgh, J. S., Aufdenkampe, A. K., Mayorga, E., Lehnert, K. A., Hsu, L., Song, L., et al. (2016). Observations data model 2: a community information model for spatially discrete Earth observations. *Environ. Model. Softw.* 79, 55–74. doi: 10.1016/j.envsoft.2016.01.010

Horsburgh, J. S., Morsy, M. M., Castronova, A., Goodall, J. L., Gan, T., Yi, H., et al. (2015). HydroShare: sharing diverse hydrologic data types and models as social objects within a hydrologic information system. *J. Am. Water Resour. Assoc.* 52, 873–889. doi: 10.1111/1752-1688.12363

Horsburgh, J. S., Tarboton, D. G., Maidment, D. R., and Zaslavsky, I. (2011). Components of an environmental observatory information system. *Comput. Geosci.* 37, 207–218. doi: 10.1016/j.cageo.2010.07.003

Horsburgh, J. S., Tarboton, D. G., Piasecki, M., Maidment, D. R., Zaslavsky, I., Valentine, D., et al. (2009). An integrated system for publishing environmental observations data. *Environ. Model. Softw.* 24, 879–888. doi: 10.1016/j.envsoft.2009.01.002

Horsburgh, J. S., Tarboton, D. G., Schreuders, K. A. T., Maidment, D. R., Zaslavsky, I., and Valentine, D. (2010). "Hydroserver: a platform for publishing space-time hydrologic datasets," in *Proceedings of the 2010 American Water Resources Association Spring Specialty Conference Geographic Information Systems (GIS) and Water Resources VI*, (Orlando, FL: American Water Resources Association), 10–11.

International Standards Organization [ISO] (2004). *Data Elements and Interchange Formats – Information Interchange – Representation of Dates and Times*, 3rd Edn. 33. Available at: https://www.iso.org/standard/40874.html

Jiang, Q., Kresin, F., Bregt, A. K., Kooistra, L., Pareschi, E., van Putten, E., et al. (2016). Citizen sensing for improved urban environmental monitoring. *J. Sens.* 2016:5656245. doi: 10.1155/2016/5656245

Johnson, A., Billett, C., and Boon, N. (2018). The delaware river watershed initiative. *Water Resour. Impact* 20, 10–11.

Jones, A., Aanderud, Z. T., Horsburgh, J. S., Eiriksson, D., Dastrup, D., Cox, C., et al. (2017). Designing and implementing a network for sensing water quality and hydrology across mountain to urban transitions. *J. Am. Water Resour. Assoc.* 53, 1095–1120. doi: 10.1111/1752-1688.12557

Jones, A. S., Horsburgh, J. S., Reeder, S. L., Ramirez, M., and Caraballo, J. (2015). A data management and publication workflow for a large-scale, heterogeneous sensor network. *Environ. Monit. Assess.* 187:348. doi: 10.1007/s10661-015-4594-3

Karami, M., McMorrow, G., and Wang, L. (2018). Continuous monitoring of indoor environmental quality using an Arduino-based data acquisition system. *J. Build. Eng.* 19, 412–419. doi: 10.1016/j.jobe.2018.05.014

Mason, S. J. K., Cleveland, S. B., Llovet, P., Izurieta, C., and Poole, G. C. (2014). A centralized tools for managing, archiving, and serving point-in-time data in ecological research laboratories. *Environ. Model. Softw.* 51, 59–69. doi: 10.1016/j.envsoft.2013.09.008

Mazumdar, S., Wrigley, S., and Ciravegna, F. (2017). Citizen science and crowdsourcing for Earth observations: an analysis of stakeholder opinions on the present and future. *Remote Sens.* 9:87. doi: 10.3390/rs9010087

Muste, M. V., Bennerr, D. A., Secchi, S., Schnoor, J. L., Kusiak, A. K., Arnold, N. J., et al. (2013). End-to-end cyberinfrastructure for decision-making support in watershed management. *J. Water Resour. Plan. Manage.* 139, 565–573. doi: 10.1061/(ASCE)WR.1943-5452.0000289

Rundel, P. W., Graham, E. A., Allen, M. F., Fisher, J. C., and Harmon, T. C. (2009). Environmental sensor networks in ecological research. *New Phytol.* 182, 589–607. doi: 10.1111/j.1469-8137.2009.02811.x

Sadler, J. M., Ames, D. P., and Khattar, R. (2016). A recipe for standards-based data sharing using open source software and low-cost electronics. *J. Hydroinform.* 18, 185–197. doi: 10.2166/hydro.2015.092

Samourkasidis, A., Papoutsoglou, E., and Athanasiadis, I. N. (2018). A template framework for environmental timeseries data acquisition. *Environ. Model. Softw.* (in press). doi: 10.1016/j.envsoft.2018.10.009

Wickert, A. D. (2014). The ALog: inexpensive, open-source, automated data collection in the field. *Bull. Ecol. Soc. Am.* 95, 166–176. doi: 10.1890/0012-9623-95.2.68

Wong, B. P., and Kerkez, B. (2016). Real-time environmental sensor data: an application to water quality using web services. *Environ. Model. Softw.* 84, 505–517. doi: 10.1016/j.envsoft.2016.07.020

Zaslavsky, I., Valentine, D., and Whiteaker, T. (eds) (2007). *CUAHSI WaterML. OGC Discussion Paper OGC 07-041r1. Version 0.3.0.*


**Conflict of Interest Statement:** 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.

*Copyright © 2019 Horsburgh, Caraballo, Ramírez, Aufdenkampe, Arscott and Damiano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 29 May 2019
doi: 10.3389/feart.2019.00138

# Talking SMAAC: A New Tool to Measure Soil Respiration and Microbial Activity

*Ayush Joshi Gyawali*, Brandon J. Lester and Ryan D. Stewart*

*School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA, United States*

Soil respiration measurements are widely used to quantify carbon fluxes and ascertain soil biological properties related to soil microbial ecology and soil health, yet current methods to measure soil respiration either require expensive equipment or use discrete spot measurements that may have limited accuracy, and neglect underlying response dynamics. To overcome these drawbacks, we developed an inexpensive setup for measuring $CO_2$ called the soil microbial activity assessment contraption (SMAAC). We then compared the SMAAC with a commercial infrared gas analyzer (IRGA) unit by analyzing a soil that had been subjected to two different management practices: grass buffer vs. row crop cultivation with tillage. These comparisons were done using three configurations that detected (1) *in situ* soil respiration, (2) $CO_2$ burst tests, and (3) substrate induced respiration (SIR), a measure of active microbial biomass. The SMAAC provided consistent readings with the commercial IRGA unit for all three configurations tested, showing that the SMAAC can perform well as an inexpensive yet accurate tool for measuring soil respiration and microbial activity.

**Keywords: substrate induced respiration, soil microbial activity, soil health, environmental sensing, soil $CO_2$**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology,*
*Netherlands*

***Reviewed by:***
*Ademir Araujo,*
*Federal University of Piauí, Brazil*
*Claudio Mondini,*
*Council for Agricultural*
*and Economics Research, Italy*

****Correspondence:***
*Ayush Joshi Gyawali*
*ayushg7@vt.edu*

***Specialty section:***
*This article was submitted to*
*Soil Processes,*
*a section of the journal*
*Frontiers in Earth Science*

***Received:*** *01 March 2019*
***Accepted:*** *15 May 2019*
***Published:*** *29 May 2019*

***Citation:***
*Joshi Gyawali A, Lester BJ and Stewart RD (2019) Talking SMAAC: A New Tool to Measure Soil Respiration and Microbial Activity. Front. Earth Sci. 7:138. doi: 10.3389/feart.2019.00138*

## INTRODUCTION

Increased soil respiration due to warmer temperatures may exacerbate global climate change (Rustad et al., 2000; Davidson and Janssens, 2006; Bond-Lamberty et al., 2018), as soils currently have an gross efflux of ∼60 Gt C $yr^{-1}$ and represent one of the two largest terrestrial sources of carbon fluxes. Sequestering more carbon in soils has become a goal of climate mitigation efforts, such as the *four per mille* initiative (Minasny et al., 2017), and with particular emphasis on soils that have been degraded by human activities (Lal, 2004). Soil respiration measurements can help to inform such sequestration efforts, while also providing a means to monitor the health, and function of agricultural soils (Mondini et al., 2010; Allen et al., 2011). In the laboratory, soil respiration measurements are used to interpret soil microbial characteristics, for example using assays like SIR (Bradford et al., 2010), carbon mineralization (Song et al., 2014), and catabolic response profile (Casas et al., 2011).

Soil respiration is often assessed by measuring changes in carbon dioxide ($CO_2$) concentration within a controlled volume over some period of time, and rely on either spot samples or integrated measurements. Spot samples are often analyzed using gas chromatography (GC) techniques (McGowen et al., 2018). Multiple GC measurements can also be combined for integrated measurements. However, these GC measurements can be costly, particularly

**Abbreviations:** IRGA, infrared gas analyzer; SIR, substrate induced respiration; SMAAC, soil microbial activity and assessment contraption.

when many samples are required. IRGA devices provide integrated flux measurements, and have been widely used to quantify soil respiration in forest (Gaudinski et al., 2000; Ladegaard-Pedersen et al., 2005; Don et al., 2009) and agricultural ecosystems (Smukler et al., 2012). IRGA-based measurements have also been used to study microbial community composition (Fierer et al., 2003), which represents one of the important properties related to soil function (Mukhopadhyay et al., 2014). While IRGA-based devices provide the most accurate flux data (Rowell, 1995), such sensors are often expensive, putting them beyond the means of many practitioners, and power-intensive, limiting their usefulness in the field.

Integrated measurements can also be collected using chemical titration with potassium hydroxide, KOH, or sodium hydroxide, NaOH (Haney R.L. et al., 2008). While titration methods are straightforward and can be done without expensive devices, there are concerns over the accuracy of the titration process (Haney R. et al., 2008). These methods often under-estimate soil respiration when compared to IRGA measurements (Ferreira et al., 2018). To add to this, titration methods often require substantial labor and laboratory space to conduct.

Finally, both spot and integrated samples can be analyzed using colorimetric techniques. For spot samples, colorimetric tubes can be used (Patil et al., 2010), while colorimetric paddles can provide integrated flux measurements (Sciarappa et al., 2016; Norris et al., 2018). Micro-respiration measurements, which quantify soil respiration and microbial community physiological profiles using indicator dyes in agar gel, also use colorimetric techniques (Campbell et al., 2003; Renault et al., 2013). Even though individual sampling units are relatively inexpensive, the materials are not re-usable and quickly become cost-prohibitive as the numbers of samples rise.

To address the above-mentioned shortcomings, we present an inexpensive Arduino-powered and IRGA-based $CO_2$ measurement device, called the soil microbial activity assessment contraption (SMAAC). The SMAAC has considerable flexibility, as we demonstrate using three different configurations: (1) SMAAC-Field, where the device was used to quantify soil respiration in a field setting; (2) SMAAC-Burst, where the device was used to analyze $CO_2$ evolution upon rapid re-wetting of air-dried soil; and (3) SMAAC-Biomass, where the device was used to quantify SIR. To validate these configurations, we compared the measurements provided by the SMAAC with those from a commercial field-portable IRGA system. These examples reveal that the SMAAC can perform well as an inexpensive yet accurate tool to measure soil respiration.

# MATERIALS AND METHODS

## Soil Microbial Activity Assessment Contraption (SMAAC) Description and Calibration

The sensor platform consists of four main components (**Figure 1**).

(1) Arduino Uno (Arduino LLC, Ivrea, Italy).
(2) Adafruit Data Logger Shield (Adafruit Industries, New York, NY, United States).
(3) Sandbox Electronics 10,000 ppm CO2 sensor (Sandbox Electronics, China).
(4) 5 V DC power source.

The Arduino Uno is an open source/open hardware microcontroller based on the ATMEGA 328P. It has no storage space or accurate time-keeping abilities on its own, so the data logger shield contains a real time clock (RTC) and additional circuitry to store data on a removable SD card. The SMAAC was powered using four 1.5 V AA batteries. This configuration provided up to 21 h of readings at the rate of 20 readings per minute.

The $CO_2$ sensor requires only 4 wires to communicate with the Arduino (+V, RX, TX, and Ground). The sensor uses I2C (Intra Integrated Circuit) serial protocol and determines $CO_2$ concentration using non-dispersive infrared absorbance (NDIR). Example code for integrating this sensor with the Arduino is available at https://github.com/SandboxElectronics/NDIRZ16.

The $CO_2$ sensor has an option to calibrate itself to 400 ppm $CO_2$ based on ambient readings. To verify that this first-order calibration is accurate enough for scientific use, we checked the sensor accuracy using known $CO_2$ standards ($n$ = 2). Here, the sensor was installed via a rubber stopper into a 1 L jar (**Figure 2a**). The jar was filled with $CO_2$-free air, and then 0.1 L of 1000 ppm $CO_2$ gas was replaced within the jar (providing a 100 ppm concentration within the jar). This process was repeated a second time with 1000 ppm $CO_2$ air, and also two times each with 2000 and 5000 ppm $CO_2$ air (providing concentrations of 200 and 500 ppm within the jar). The results obtained from SMAAC for these standards were repeatable within ±20 ppm and accurate within the ±50 ppm sensor limit.

## Soil Description

We tested the SMAAC with a Weaver series silt loam soil (*Fine-loamy, mixed, active,* and *mesic Fluvaquentic Eutrudepts*), located at Kentland Farm at Virginia Tech (37.198, -80.575). To include different soil microbial activity levels, we sampled two locations in adjacent fields that were managed using (1) perennial grass cover and (2) row crop cultivation with moldboard tillage. The pH of the grass-covered soil was 6.4 and of the tilled soil was 6.6, putting the soil at the upper pH limit for performing static chamber measurements [e.g., West and Sparling (1986) recommend pH $\leq$ 6.5]. We performed three tests in which the SMAAC measurements were compared to a commercially available self-contained IRGA unit (LI-COR 8100 with 20 cm diameter 8100–8103 survey chamber, LI-COR, Lincoln, NE, United States): SMAAC-Field, SMAAC-Burst, and SMAAC-Biomass.

## Field and Laboratory Measurements

### SMAAC-Field Soil Respiration Test

We used 200 mm (diameter) by 150 mm (height) PVC columns for the field measurements. We collected a 2-min $CO_2$ respiration measurement first using the SMAAC located within the LI-COR 8100–8103 sampling chamber (i.e., SMAAC-simultaneous;

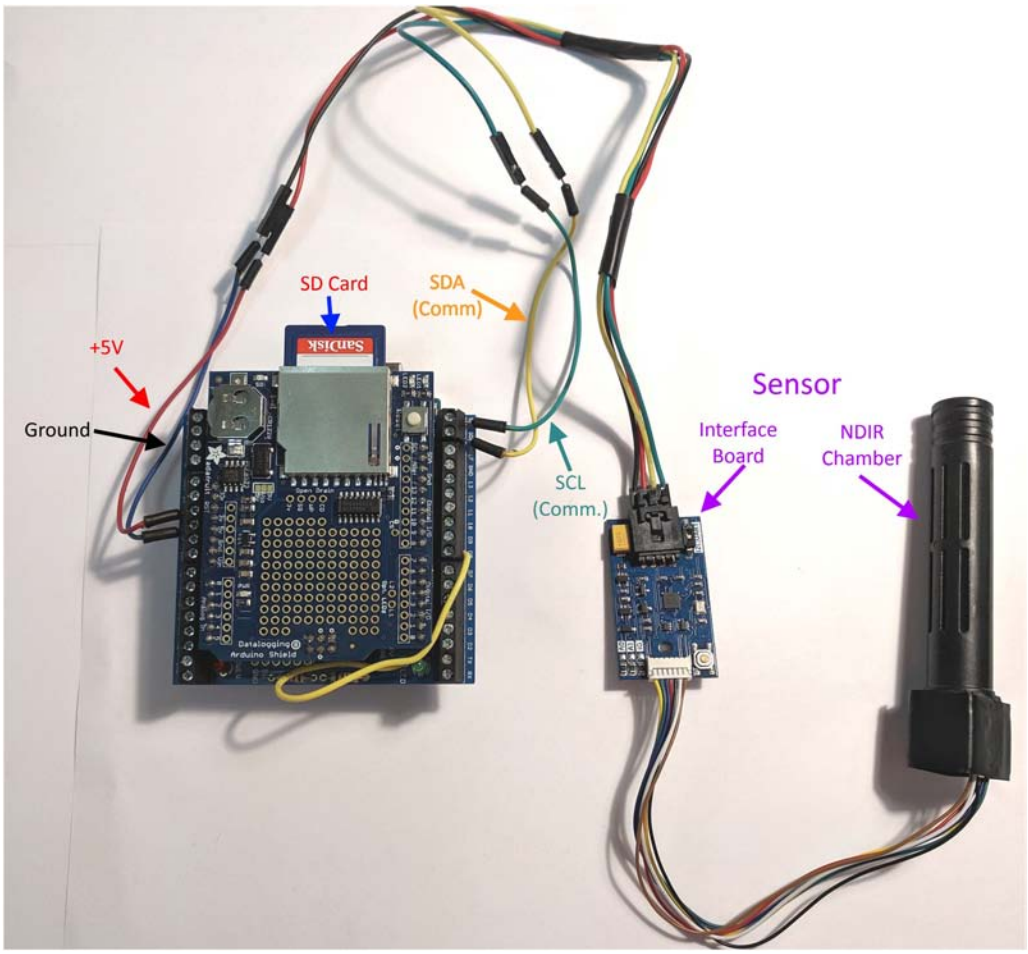



**FIGURE 1 |** Schematic of the soil microbial activity assessment contraption (SMAAC). NDIR, non-dispersive infrared sensor used to detect $CO_2$; SCL, serial clock line, used to synchronize data and commands between the Arduino and the interface board; SDA, serial data line, used to send and receive serial data and commands to the interface board.

**Figure 2b**). Note that the sampling chamber provided an air-tight seal around the PVC column during measurements. Immediately after this first measurement the LI-COR unit was removed and the ring was capped with an airtight rubber cap (i.e., SMAAC-independent; **Figure 2c**). The SMAAC then collected a second 2-min measurement. The $CO_2$ flux [$f_{CO2}$; (N $L^{-2}$ $t^{-1}$)] was estimated as:

$$f_{CO_2} = \frac{P_0 V_c}{R T_0 A} \frac{\Delta C}{\Delta t} \quad (1)$$

where $P_0$ is the pressure in the chamber [M $L^{-1}$ $t^{-2}$], assumed to be equal to atmospheric pressure, $V_c$ is the volume of the sampling chamber plus any tubing and pumps [$L^3$], $R$ is the ideal gas law constant [M $L^2$ $N^{-1}$ $T^{-1}$ $t^{-2}$], $T_0$ is the temperature of the air [T], $A$ is the area of exposed soil [$L^2$], and $\Delta C$ is the change in $CO_2$ concentration on a molar basis [N $N^{-1}$] per change in time $\Delta t$ [t].

Four rings were sampled for each of the grass-covered and tilled soils ($n = 4$).

## SMAAC-Burst $CO_2$ Test

For the $CO_2$ burst test, we placed 200 g of 4-mm sieved and air-dried soil from the two sites into a 200 mm diameter by 150 mm tall column. The water holding capacity for each soil sample was measured using the funnel method (Fierer et al., 2006). Water was added dropwise to each soil sample using a syringe until the sample reached 50% water holding capacity. Once the soil samples were wetted, the SMAAC was placed on the soil surface (**Figure 2d**). The LI-COR 8100 sampling hood was then placed on top. Both instruments collected readings several times a minute for at least 2 h. For each instrument, the readings collected were averaged per minute for graphing purposes ($n = 4$ per soil).

## SMAAC-Biomass Substrate Induced Respiration (SIR)

We also compared LI-COR 8100 and SMAAC measurements during a test designed to mimic SIR measurements (Fierer et al., 2003; Strickland et al., 2010). Refrigerated soil samples from the fields were brought to room temperature overnight. We placed 80 g (equivalent dry mass) of 4-mm sieved soil samples into a

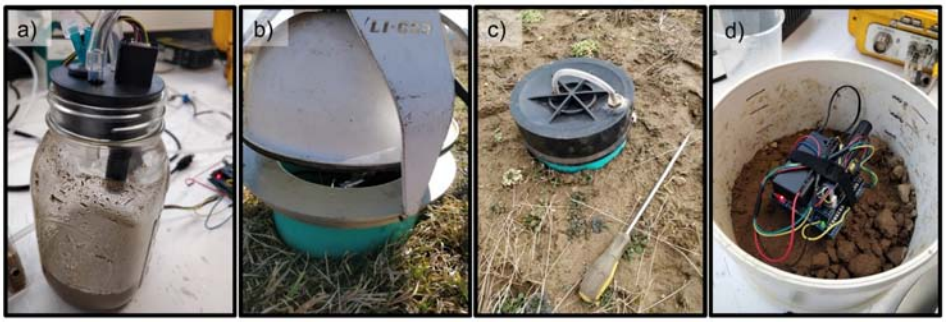



**FIGURE 2 |** Measurement setups for: **(a)** SMAAC-Biomass substrate induced respiration (SIR) measurement; **(b)** SMAAC-Field flux measurement with SMAAC simultaneously located within the LI-COR 8100 sampling chamber; **(c)** SMAAC-Field flux measurement with SMAAC independent of the LI-COR unit; and **(d)** SMAAC-Burst laboratory CO2 burst measurement.

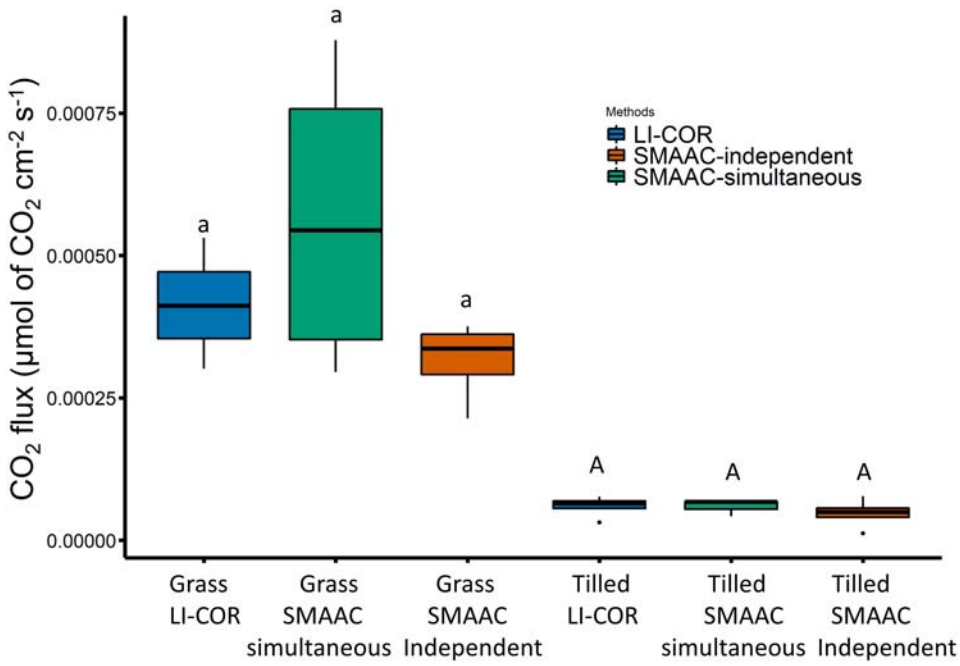



**FIGURE 3 |** $CO_2$ fluxes measured in the field by the LI-COR (blue), SMAAC-simultaneous (green), and SMAAC-independent (orange). Different small letters indicate grass-covered soil fluxes are statistically different; different capital letters indicate tilled soil fluxes are statistically different (ANOVA with Tukey's HSD; $P < 0.05$).

1 L glass jar (**Figure 2a**). We then added 0.16 L of autolyzed yeast solution made from 12 g of yeast extract (BD Biosciences, San Jose, CA, United States) in 1 L of DI water as a substrate. The mixture of soil and substrate was shaken with no cover for 10 min. We then sealed the jar using a rubber stopper that had the SMAAC sensor and a septum mounted through it. Using the septum, we flushed the headspace of the jar using $CO_2$ free air for 7 min. Then the jar was maintained at 20°C for 4 h. After 4 h, we collected a gas sample through the septum using a syringe. This sample was injected into the LI-COR 8100 unit to quantify the $CO_2$ concentration in the jar headspace. The 4-h $CO_2$ reading from the SMAAC was also analyzed. Both measurements of headspace $CO_2$ were converted to SIR units (µg C $g^{-1}$ dry soil $h^{-1}$) based on the dry mass of soil. Three replicates were analyzed for the grass buffer and moldboard plowed soils ($n = 3$).

## Statistical Analyses

All statistical analysis and figures were done in R Version 3.5.0 (R Development Core Team., 2018). Analysis of variance (ANOVA) was used to compare the three types of measurements performed in the SMAAC-Field configuration (i.e., LI-COR, SMAAC-simultaneous, and SMAAC-independent). During the SMAAC-Burst and SMAAC-Biomass tests, the Student's *t*-test was used to compare results from the LI-COR vs. the SMAAC. Measurements were analyzed separately for the grass-covered and tilled soils. $\alpha = 0.05$ was used to test for significance throughout this study.

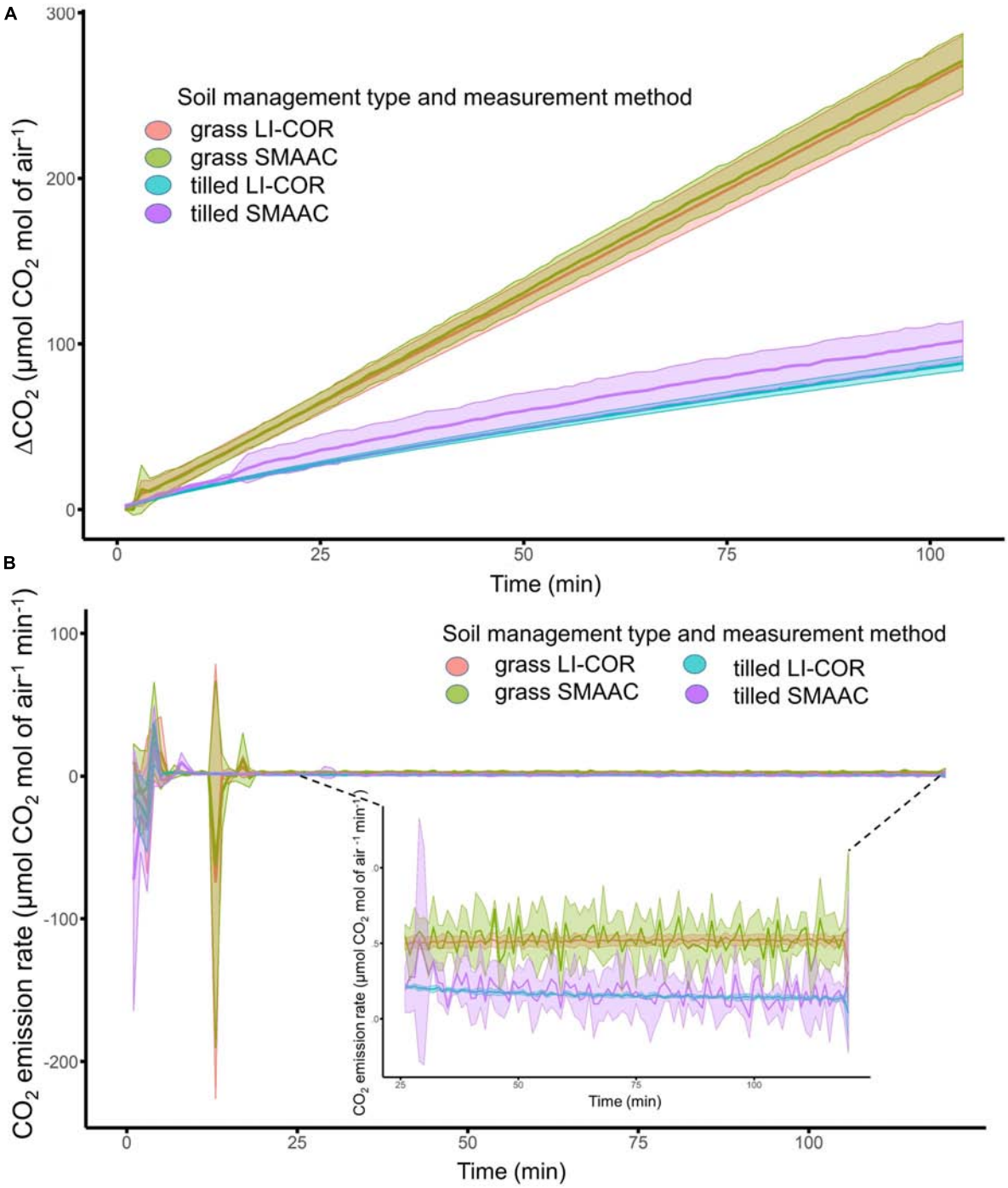



**FIGURE 4 | (A)** Laboratory measurement of change in $CO_2$ through time (min) for grass-covered and tilled soils, measured using a LI-COR 8100, and the SMAAC-Burst configuration. Solid lines represent mean values, and shaded areas represent standard deviations from the means. **(B)** $CO_2$ emission rates for grass-covered and tilled soils, measured using the LI-COR, and the SMAAC-Burst configuration.

# RESULTS

## SMAAC-Field Soil Respiration Test

For the SMAAC-Field respiration test, the LI-COR 8100 and SMAAC were used to quantify $CO_2$ flux over a 2-min period, with the SMAAC both placed within (SMAAC-simultaneous) and without (SMAAC-independent) the LI-COR sampling chamber. Both instruments showed that the grass-covered soil had a higher $CO_2$ flux than the tilled soil (**Figure 3**). The flux measured for the grass buffer soil by the LI-COR ($4.1 \times 10^{-4}$ µmol

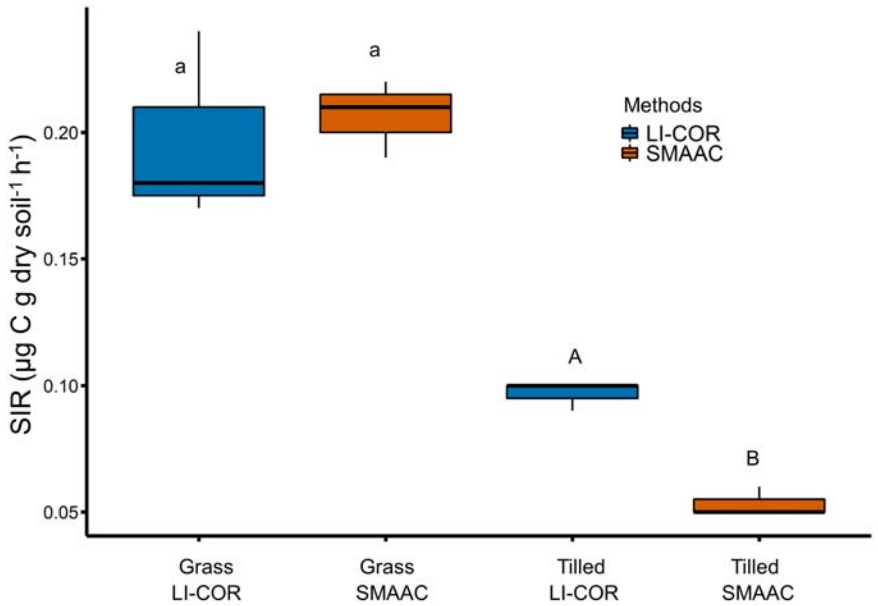



**FIGURE 5 |** Substrate induced respiration measured by a LI-COR 8100 (blue) vs. the SMAAC-Biomass configuration (orange). Different small letters indicate grass-covered soil fluxes are statistically different; different capital letters indicate tilled soil fluxes are statistically different (Students *t*-test; $P < 0.05$).

$CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $9.9 \times 10^{-5}$ standard deviation, SD) was not significantly different than fluxes determined via the SMAAC-simultaneous ($5.6 \times 10^{-4}$ µmol $CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $2.7 \times 10^{-4}$ SD) or SMAAC-independent ($3.1 \times 10^{-4}$ µmol $CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $7.2 \times 10^{-5}$ SD) tests. For the tilled soil, the LI-COR flux ($5.9 \times 10^{-5}$ µmol $CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $1.9 \times 10^{-5}$ SD) was again not significantly different from the fluxes measured during the SMAAC-simultaneous ($5.9 \times 10^{-5}$ µmol $CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $1.5 \times 10^{-5}$ SD) and SMAAC-independent ($4.7 \times 10^{-5}$ µmol $CO_2$ $cm^{-2}$ $s^{-1}$ $\pm$ $2.6 \times 10^{-5}$ SD) tests.

## SMAAC-Burst $CO_2$ Burst Test

The SMAAC-Burst configuration produced consistent results compared to the LI-COR 8100 unit for both the grass-covered and tilled soils (**Figure 4**), with similar mean values and standard deviations calculated from the four physical replicates for each soil (**Figure 4A**). We observed relatively large fluctuations in $CO_2$ emission rates, especially during the first 20 min of the experiment (**Figure 4B**). After this initial period, $CO_2$ emission rates fluctuated more for SMAAC compared to LICOR, though the mean rates were generally consistent between methods (**Figure 4B**). Both instruments showed that the $CO_2$ burst was larger in the grass-covered soil compared to the tilled soil (**Figure 4**).

## SMAAC-Biomass Substrate Induced Respiration Test

Results generated using both the LI-COR 8100 and the SMAAC-Biomass consistently showed that the grass-covered soil had higher SIR values than the tilled soil (**Figure 5**). The LI-COR (0.19 µg C g dry $soil^{-1}$ $h^{-1}$ $\pm$ 0.03 SD) and SMAAC (0.21 µg C g dry $soil^{-1}$ $h^{-1}$ $\pm$ 0.01 SD) measurements were not statistically different for the grass-covered soil ($P \geq 0.05$). However, the LI-COR SIR value (0.09 µg C g dry $soil^{-1}$ $h^{-1}$ $\pm$ 0.005 SD) for the tilled soil was significantly higher than the SMAAC SIR value (0.05 µg C g dry $soil^{-1}$ $h^{-1}$ $\pm$ 0.005 SD; $P = 0.0009$).

# DISCUSSION

In this study we developed three configurations of an Arduino-based $CO_2$ sensor that allowed us to assess soil microbial activity. Our instrument, deemed the SMAAC, was then compared against a commercial IRGA unit (LI-COR 8100). Overall, the SMAAC generated similar results to the commercial IRGA, with signficant differences only observed when SIR was quantified for the tilled soil (**Figure 5**). In this example, the SIR value from the SMAAC-Biomass configuration was approximately half of the value estimated by the LI-COR. The reason for the discrepancy may relate to the accuracy of the SMAAC IRGA sensor (50 ppm per the manufacturer). Even though our calibration analysis determined that the instrument provided consistent readings for $CO_2$ concentrations between 100 and 500 ppm, the sensor accuracy implies that the error can exceed 10% for $CO_2$ concentrations < 500 ppm. Using the sensor to measure low $CO_2$ concentrations may therefore require extra precautions such as using longer run times, greater number of replicates, and more frequent calibration. We also note that we did not test the sensor beyond 1,000 ppm, so the calibration should also be assessed when using SMAAC to measure higher $CO_2$ concentrations.

The SMAAC tended to show more measurement noise than the LI-COR when assessing $CO_2$ fluxes, e.g., the field flux

measurements from SMAAC-simultaneous vs. LI-COR setups in the grass-covered soil (**Figure 3**), or the emissions rates calculated for both soils with the SMAAC-Burst (**Figure 4B**). However, during the field flux measurements, the SMAAC-independent test had a slightly lower median flux and a smaller standard deviation than either the LI-COR or SMAAC-simultaneous. This result may reflect the influence of the LI-COR pump unit, which provided continuous circulation of air in the chamber. At the same time, our flux calculations (Eq. 1) assumed that the volume of the air, $V_c$, for the LI-COR and SMAAC-simultaneous setups was equal to the LI-COR sampling chamber plus the internal pump volume of the LI-COR. We did not account for the volume or the exposed surface area of the soil occupied by the SMAAC itself, thus potentially introducing minor error into the flux calculations for those tests. We also note here that the SMAAC and LI-COR both showed high variability in emissions during the initial 20 min of the $CO_2$ burst test experiment. This result may reflect an equilibration period within the glass jar, particularly in response to the initial soil disturbance during wetting the soil and sealing the system.

The total cost of the SMAAC was ∼$150, making it at least two orders of magnitude less expensive than commercial IRGA units. Despite the low cost, the SMAAC still maintained reasonable accuracy in all three configurations tested, and performed repeatable measurements when compared with $CO_2$ standards. The SMAAC is lighter weight and requires less power than commercial IRGA units, increasing its usefulness when performing extended measurements or working in remote locations. An additional benefit of the SMAAC comes from its small form factor: it can be placed directly inside the headspace of samples, thus eliminating the need to pull discrete gas samples using a syringe. Removing this step eliminates a potential source of error, particularly since many commercial IRGA pump units are not fully sealed.

The SMAAC may open new avenues of inquiry related to soil respiration measurements, both in terms of the configurations shown here as well as other possible configurations yet to be developed. For example, we focused our tests on closed chamber measurements, since those are commonly used to evaluate soil $CO_2$ fluxes, and perform measurements such as SIR. The closed chamber measurements also lended themselves to direct comparison with the commercial IRGA unit. However, $CO_2$ can also be measured using open systems (Norman et al., 1997; Alterio et al., 2006) or in continually flushed chambers (Chow et al., 2006). Using the SMAAC in open/purged systems thus represents an area of possible future development.

Similarly, since the SMAAC system is inexpensive and easy to assemble, multiple sensors could be used concurrently to better quantify spatial, and temporal variability in soil biological measurements, for example by analyzing multiple chambers simultaneously and thereby providing similar functionality as multiplexer units often offered with commercial IRGAs. Finally, direct continuous logging of $CO_2$ evoluation during measurements may help generate new insights. For example, the grass-covered vs. tilled soil showed different temporal trends in the SMAAC-Burst test (**Figure 4B**), where the grass-covered soil produced a constant $CO_2$ efflux rate over the 2-h test period vs. a decreasing $CO_2$ efflux rate for the tilled soil. While the underlying mechanisms controlling these different responses remain beyond the scope of this current paper, it is nonetheless worth noting that it would not be possible to observe such trends without the high measurement frequency offered by IRGA-based instruments such as SMAAC.

## CONCLUSION

The SMAAC developed in this study represents an low cost yet reliable way to measure $CO_2$ fluxes from soils. The results obtained from the SMAAC were consistent with those from a commercial IRGA unit for both field and laboratory measurements. In this study we highlighted three SMAAC configurations that were designed to assess different aspects of soil microbial activity and function, yet the SMAAC also has the potential to generate additional applications and insights. As an example, by having the SMAAC-Burst and SMAAC-Biomass units placed inside the closed headspace above samples, we generated near-continuous measurements of $CO_2$ evolution through time. Such $CO_2$ trends may provide new understanding of soil microbial processes that is not possible via traditional discrete measurements. In conclusion, the SMAAC is a promising tool for measuring soil respiration and microbial activity that warrants usage by the broader scientific community.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

AJG contributed to designing and conducting the experiments, performing the analyses, and writing the first draft of the manuscript. BL contributed to making the sensor, conducting the experiments, and editing the manuscript. RS contributed to generating the main idea, providing guidance throughout the experiments and analysis, and writing and editing the manuscript.

## FUNDING

This work was provided by the U.S. Department of Agriculture NRCS Conservation Innovation Grant #69-3A75-14-260 and also provided in part by the U.S. Department of Agriculture National Resources Conservation Service Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture. We also would like to thank Virginia Tech Open Access Subvention fund for support regarding publication fees.

## REFERENCES


Allen, D. E., Singh, B. P., and Dalal, R. C. (2011). Soil health indicators under climate change: a review of current knowledge. *Soil Health Clim. Change.* 29, 25–45. doi: 10.1007/978-3-642-20256-8_2

Alterio, G., Giorio, P., and Sorrentino, G. (2006). Open-system chamber for measurements of gas exchanges at plant level. *Environ. Sci. Technol.* 40, 1950–1955. doi: 10.1021/es052094o

Bond-Lamberty, B., Bailey, V. L., Chen, M., Gough, C. M., and Vargas, R. (2018). Globally rising soil heterotrophic respiration over recent decades. *Nature* 560, 80–83. doi: 10.1038/s41586-018-0358-x

Bradford, M. A., Watts, B. W., and Davies, C. A. (2010). Thermal adaptation of heterotrophic soil respiration in laboratory microcosms. *Glob. Change Biol.* 16, 1576–1588. doi: 10.1111/j.1365-2486.2009.02040.x

Campbell, C. D., Chapman, S. J., Cameron, C. M., Davidson, M. S., and Potts, J. M. (2003). A rapid microtiter plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities by using whole soil. *Appl. Environ. Microbiol.* 69, 3593–3599. doi: 10.1128/aem.69.6.3593-3599.2003

Casas, C., Omacini, M., Montecchia, M. S., and Correa, O. S. (2011). Soil microbial community responses to the fungal endophyte Neotyphodium in Italian ryegrass. *Plant Soil* 340, 347–355. doi: 10.1007/s11104-010-0607-8

Chow, A. T., Tanji, K. K., Gao, S., and Dahlgren, R. A. (2006). Temperature, water content and wet–dry cycle effects on DOC production and carbon mineralization in agricultural peat soils. *Soil Biol. Biochem.* 38, 477–488. doi: 10.1016/j.soilbio.2005.06.005

Davidson, E. A., and Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* 440:165. doi: 10.1038/nature04514

Don, A., Rebmann, C., Kolle, O., Scherer-Lorenzen, M., and Schulze, E. D. (2009). Impact of afforestation-associated management changes on the carbon balance of grassland. *Glob. Change Biol.* 15, 1990–2002. doi: 10.1111/j.1365-2486.2009.01873.x

Ferreira, C. R. P. C., Antonino, A. C. D., Sampaio, E. V. D. S. B., Correia, K. G., Lima, J. R. D. S., Soares, W., et al. (2018). Soil CO2 efflux measurements by alkali absorption and infrared gas analyzer in the brazilian semiarid region. *Revista Brasileira de Ciência do Solo* 42:e0160563.

Fierer, N., Colman, B. P., Schimel, J. P., and Jackson, R. B. (2006). Predicting the temperature dependence of microbial respiration in soil: a continental-scale analysis. *Glob. Biogeochem. Cycles* 20:GB3026.

Fierer, N., Schimel, J. P., and Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biol. Biochem.* 35, 167–176. doi: 10.1016/s0038-0717(02)00251-1

Gaudinski, J. B., Trumbore, S. E., Davidson, E. A., and Zheng, S. (2000). Soil carbon cycling in a temperate forest: radiocarbon-based estimates of residence times, sequestration rates and partitioning of fluxes. *Biogeochemistry* 51, 33–69.

Haney, R., Brinton, W., and Evans, E. (2008). Soil CO 2 respiration: comparison of chemical titration. CO 2 IRGA analysis and the Solvita gel system. *Renew. Agri. Food Sys.* 23, 171–176. doi: 10.1017/s174217050800224x

Haney, R. L., Brinton, W., and Evans, E. (2008). Estimating soil carbon, nitrogen, and phosphorus mineralization from short-term carbon dioxide respiration. *Commun. Soil Sci. and Plant Anal.* 39, 2706–2720. doi: 10.1080/00103620802358862

Ladegaard-Pedersen, P., Elberling, B., and Vesterdal, L. (2005). Soil carbon stocks, mineralization rates, and CO2 effluxes under 10 tree species on contrasting soil types. *Can. J. Forest Res.* 35, 1277–1284. doi: 10.1139/x05-045

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science* 304, 1623–1627. doi: 10.1126/science.1097396

McGowen, E. B., Sharma, S., Deng, S., Zhang, H., and Warren, J. G. (2018). An automated laboratory method for measuring CO2 emissions from soils. *Agri. Environ. Lett.* 3, 2–3. doi: 10.2134/ael2018.02.0008

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., et al. (2017). Soil carbon 4 per mille. *Geoderma* 292, 59–86. doi: 10.1016/j.scitotenv.2017.12.263

Mondini, C., Sinicco, T., Cayuela, M. L., and Sanchez-Monedero, M. A. (2010). A simple automated system for measuring soil respiration by gas chromatography. *Talanta* 81, 849–855. doi: 10.1016/j.talanta.2010.01.026

Mukhopadhyay, S., Maiti, S., and Masto, R. (2014). Development of mine soil quality index (MSQI) for evaluation of reclamation success: a chronosequence study. *Ecol. Eng.* 71, 10–20. doi: 10.1016/j.ecoleng.2014.07.001

Norman, J., Kucharik, C., Gower, S., Baldocchi, D., Crill, P., Rayment, M., et al. (1997). A comparison of six methods for measuring soil-surface carbon dioxide fluxes. *J. Geophys. Res. Atmos.* 102, 28771–28777. doi: 10.1029/97jd01440

Norris, R., Chim, B. K., Evanylo, G., Reiter, M., and Thomason, W. (2018). assessment of in-season soil nitrogen tests for corn planted into winter annual cover crops. *Soil Sci. Soc. Am. J.* 82, 1428–1436.

Patil, R. H., Colls, J. J., and Steven, M. D. (2010). Effects of CO2 gas as leaks from geological storage sites on agro-ecosystems. *Energy* 35, 4587–4591. doi: 10.1016/j.energy.2010.01.023

R Development Core Team. (2018). *A language and Environment for Statistical Computing*. Austria: R Foundation for Statistical Computing.

Renault, P., Ben-Sassi, M., and Berard, A. (2013). Improving the MicroResp substrate-induced respiration method by a more complete description of CO2 behavior in closed incubation wells. *Geoderma* 207, 82–91. doi: 10.1016/j.geoderma.2013.05.010

Rowell, M. J. (1995). Colorimetric method for CO2 measurement in soils. *Soil Biol. Biochem.* 27, 373–375. doi: 10.1016/0038-0717(94)00218-p

Rustad, L. E., Huntington, T. G., and Boone, R. D. (2000). Controls on soil respiration: implications for climate change. *Biogeochemistry* 48, 1–6. doi: 10.1111/gcb.13308

Sciarappa, W., Murphy, S., Quinn, V., Barresi, R. and Ward, D. (2016). Assessing soil health in highbush blueberry with the Solvita CO2 respiration test. *Acta Hortic.* 1180, 327-334

Smukler, S., O'Geen, A., and Jackson, L. (2012). Assessment of best management practices for nutrient cycling: A case study on an organic farm in a Mediterranean-type climate. *J. Soil Water Conser.* 67, 16–31. doi: 10.2489/jswc.67.1.16

Song, Y., Song, C., Tao, B., Wang, J., Zhu, X., and Wang, X. (2014). Short-term responses of soil enzyme activities and carbon mineralization to added nitrogen and litter in a freshwater marsh of Northeast China. *Eur. J. Soil Biol.* 61, 72–79. doi: 10.1016/j.ejsobi.2014.02.001

Strickland, M. S., Devore, J. L., Maerz, J. C., and Bradford, M. A. (2010). Grass invasion of a hardwood forest is associated with declines in belowground carbon pools. *Glob. Change Biol.* 16, 1338–1350. doi: 10.1111/j.1365-2486.2009.02042.x

West, A., and Sparling, G. (1986). Modifications to the substrate-induced respiration method to permit measurement of microbial biomass in soils of differing water contents. *J. Microbiol. Methods* 5, 177–189. doi: 10.1016/0167-7012(86)90012-6


**Conflict of Interest Statement:** 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.

*Copyright © 2019 Joshi Gyawali, Lester and Stewart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

BRIEF RESEARCH REPORT
published: 31 May 2019
doi: 10.3389/feart.2019.00137

# OPEnS Hub: Real-Time Data Logging, Connecting Field Sensors to Google Sheets

*Thomas DeBell[1]*, Luke Goertzen[2], Lars Larson[1], William Selbie[2], John Selker[1] and Chet Udell[1]*

[1] Openly Published Environmental Sensing Lab, Biological and Ecological Engineering, Oregon State University, Corvallis, OR, United States, [2] Openly Published Environmental Sensing Lab, Computer Science, Oregon State University, Corvallis, OR, United States

In Earth science, we must often collect data from sensors installed in remote locations. Retrieving these data and storing them can be challenging. Present options include proprietary commercial dataloggers, communication devices, and protocols with rigid software and data structures that may require ongoing expenses. While there are open-source solutions that include telemetry, such as EnviroDIY's Mayfly, none presently generate real-time, remotely accessible workbooks (Aufdenkampe et al., 2017; EnviroDIY, 2018). The Openly Published Environmental Sensing (OPEnS) Lab developed the OPEnS Hub, a new approach to using low-power, open-source hardware and software to achieve real-time data logging from the field to the web. The Hub is an order of magnitude less expensive than commercial products, inherently modular and flexible, and aims to reduce technical barriers for users with little programming experience (DeBell, 2019). Data can be collected remotely using a host of transmission protocols to relay data from distributed *in situ* monitoring devices. The Hub mesh-networks with several nodes and backs up to an onboard microSD card. Telemetry options include 900 MHz Long Range Radio (LoRa) with up to 25 km range and Nordic Radio Frequency (nRF) for higher data rates (Feather, 2018). Ongoing transmissions from the Hub to the internet currently employ Ethernet with potential support for Wi-Fi and the cell network. The Hub engages a dynamic, low-latency portal to Google Sheets via the free Application Programming Interface (API), *PushingBox*, and an adaptable Google Apps Script. This framework was tested on 12 individual sensors nodes at remote sites in Oregon. This manuscript details our methods and evaluates *PushingBox*, Google Apps Script, Adafruit Industries' open-hardware Feather development boards, the Hypertext Transfer Protocol (HTTP), and the aforementioned modes of data transfer.

**Keywords: open-source, *in situ*-sensing, arduino, lora, google-sheet, data-logging, IOT, low-cost**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology, Netherlands*

***Reviewed by:***
*Andrew Wickert,*
*University of Minnesota Twin Cities, United States*
*Jeffery S. Horsburgh,*
*Utah State University, United States*

****Correspondence:***
*Thomas DeBell*
*debellt@oregonstate.edu;*
*tcdebell@ncsu.edu*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *01 November 2018*
***Accepted:*** *15 May 2019*
***Published:*** *31 May 2019*

***Citation:***
*DeBell T, Goertzen L, Larson L, Selbie W, Selker J and Udell C (2019) OPEnS Hub: Real-Time Data Logging, Connecting Field Sensors to Google Sheets. Front. Earth Sci. 7:137. doi: 10.3389/feart.2019.00137*

## INTRODUCTION

Advancements in sensing technology have sparked a new age of data acquisition that continues to change how we understand the world around us. However, proprietary data loggers can be prohibitively expensive for distributed *in situ* sensing. These systems often store data onboard, demanding intermittent retrieval from the field or requiring ongoing fees for remote access with

satellite telemetry (Communications, 2018). Proprietary systems often require separate data loggers at each sensor location, making spatially distributed sensing costly (CR1000, 2018). Although commercial loggers come preassembled and tested, by using open source alternatives, users gain an ever-growing community of collaborators and a robust, inexpensive platform. The OPEnS Hub is less than one-tenth the price of common commercial options with a hardware cost of $128, neglecting the cost of assembly time and testing which varies depending on the user's technical background (DeBell, 2019).

One solution to the problem of logging data from remote locations leverages the "Internet of Things" (IoT) movement: everything can be connected to the internet. Specifically, the OPEnS Lab has established an "Internet of Agriculture" (IoA) initiative using open-source IoT-enabled devices to collect scientific data on environmental conditions. A significant challenge to the IoA is that systems are deployed in remote areas where Wi-Fi is not accessible. Existing open-source dataloggers such as the Northern Widget LLC ALog are proven as reliable tools for automated field data acquisition, but still lack telemetry (Wickert, 2014). However, a more recent open-source development by a University of North Texas research group supports the Zigbee telemetry protocol with an Xbee module and hosts data to the web using a Raspberry Pi and custom web interface (Ferdoush and Li, 2014; Raspberry and Pi Foundation, 2019). Additionally, EnviroDIY's Mayfly offers an Arduino-based system with onboard telemetry options such as Xbee (900 MHz and 2.4 GHz), LoRa, and WiFi that interfaces with cloud-based data platforms including the Model My Watershed Web app (Hicks et al., 2019).

To build on the existing open source systems above, the goal of the OPEnS Hub was to create an inherently modular, cost-effective platform with a continuous, real-time link to Google Sheets. This process allows for data to be shared, viewed, and analyzed by anyone of the two billion active Google account users in the familiar Google ecosystem (Popper, 2017). We sought to develop a device that accommodates a variety of long-range wireless telemetry options and to provide open-source documentation (see GitHub) at a technical level such that a farmer, scientist, or student would be able to replicate our work (DeBell, 2019). Tutorials, computer-aided design files (CAD), code and other supporting documentation for the Hub are located at the project GitHub repository, https://github.com/OPEnSLab-OSU/OPEnS-Hub_Frontiers. A release of the GitHub repository was deposited in Zenodo for archival purposes (DeBell et al., 2019). The OPEnS Hub stands to simultaneously lower the cost of experimentation and data collection while breaking down traditional technical barriers.

# MATERIALS AND METHODS

## Hardware

The physical components of the Hub rely on an open-hardware suite of development boards produced by Adafruit Industries and driven by the ATMEGA32u4 microcontroller (Feather, 2018). We chose the Adafruit Feather line of development boards for their low power requirements (∼0.7 mA standby), smaller form factor, and embedded telemetry options, when compared to the ubiquitous Arduino Uno (∼15 mA standby) (SparkFun, 2015; DeBell, 2019). Variants of the Feather include onboard modules enabling 900 MHz Long Range Radio (LoRa) transmissions or Wi-Fi/Ethernet connectivity. Stackable "FeatherWing" extensions for the development boards include the Global System for Mobile Communication (GSM), 2.4 GHz Nordic Radio Frequency (nRF), and Bluetooth modules. Feathers are programmed using C++ (International Organization for Standardization, 2013) in the Arduino platform (Arduino, 2019). The boards selected for field implementation were the Ethernet FeatherWing (Adafruit Industries, 2019) to connect the Hub to the web, the real-time clock FeatherWing (Adafruit Industries, 2018b) to make accurate timestamps of transmissions, and the LoRa-enabled development board (Adafruit Industries, 2018a) which accesses a non-licensed 900 MHz radio band to transmit data from the sensors to the logger. A 3-ft-long, 8-dB, 50-Ohm impedance, omnidirectional radio antenna was used to improve transmission strength. Custom, 3D-printed enclosures were designed in Autodesk's Fusion 360 (A360, 2019) to protect the Hub from field conditions. This produced a housing that could be rapidly modified to meet varying configurations with a production cost of $12 (DeBell, 2019). A comprehensive list of hardware can be found in the bill of materials included in the Section **Supplementary Materials**.

## Software

A cloud service was utilized to process, store, and provide users with remote access to the collected data. Google's App Script was chosen because it is free and can be easily modified in a language similar to JavaScript. This application also makes data available in a simple, familiar environment and displays near real-time updates using Google's reliable spreadsheet interface. The Google ecosystem lends itself well to open data and readily pairs with open-hardware.

The process of getting field data to a Google spreadsheet requires several steps. Data must first be packaged into a format that can be sent and parsed, the device must connect to the internet, and a Hypertext Transfer Protocol (HTTP) request containing the data triggers an Application Programming Interface (API), PushingBox (PushingBox, 2018). This API was primarily chosen because it is free to use, compatible with open-hardware, and it does not require a secure connection to move data into its "scenarios" before offloading this information into Google Sheets (see Pushingbox folder on GitHub).

Each sensor node sends the spreadsheet ID, tab ID, and column names alongside the data so that the App Script can create any number of Google Sheets from a single Hub. To achieve this, each node sends data in key-value pairs (KVP). For every data point sent, the Hub specifies the origin of the data (i.e., the column in the spreadsheet) to be correctly organized, coupled with the data value itself. As a result, each data point requires two HTTP *GET* arguments. Although sending these KVPs adds to the total packet size, this protocol enables dynamic addition or removal of sensors without needing to change the App Script.

The next steps no longer involve the development board; the API can extract and forward data from the Hub to a Google Script. When the Google Script receives a *GET* argument, it creates a JavaScript dictionary, relating the keys to the values which will identify the correct spreadsheet and tab and finally write these data into the corresponding columns. Next, it accesses the specified spreadsheet and tab and checks the most recent column headers. The data is then sorted into the correct columns, or a new header is created if the data keys have changed since the last upload. A full visual representation of this process is in **Figure 1**.

Much of the complexity of this routine stems from the limited processing capacity of Arduino-like devices for supporting the Secure Sockets Layer (SSL) or Transport Layer Security (TLS) encryption protocol required for HTTPS (HTTP Secure). This barrier is nontrivial because Google Scripts/Apps can only be accessed via secure connections. As such, the device needs to offload the direct communication with the script to another platform such as the PushingBox API. While PushingBox can trigger a variety of services upon receiving a HTTP request, the OPEnS Hub sends data to the script URL which effectively converts the original HTTP request from the Hub to a HTTPS request to reach the Google script.

## Lab Testing

Since each sensor deployment configuration is unique, it was necessary to be able to test each device individually and in concert over the internet gateway to know that data was transcribed correctly to the spreadsheet. First, testing was done to confirm that the sensors were transmitting the correct data at specified intervals to the Hub. This also tested the system's scalability by proving that multiple devices could transmit to the Hub simultaneously without losing or corrupting data. The use of a free API presented one of the significant constraints of the project because each account is limited to 1,000 HTTP requests per day. For initial testing, the sampling frequency was 5 min or 288 readings per day. The system was then scaled to support any number of devices as long as the sampling frequency did not exceed 1,000 requests per day. Prototype testing simulated field conditions by sending transmissions over a kilometer, subjecting the enclosure to precipitation, and exposing the system to high UV intensity. Although no field testing was done beyond 1 km in the field, Adafruit Industries states that the upper limit for their LoRa transmitter and receiver can be upward of 25 km line of sight (Adafruit Industries, 2019).

## Field Testing

Although there are a variety of telemetry options supported by the OPEnS Hub, LoRa radio proved to be the most applicable for field testing at ranges exceeding half a kilometer. Field testing consisted of three deployments among two different sites. The first two field experiments were conducted at the H.J. Andrews Experimental Forest (**Figure 2A**) near Blue River, Oregon in July 2017 and July 2018, and the third was at Lewis Brown Farms (**Figure 2B**) near Corvallis, Oregon in April 2018 (DeBell, 2019).

The first experiment consisted of a Hub equipped with LoRa radio and a wired Ethernet connection and one LoRa-enabled weather station located approximately half a kilometer away through the densely wooded forest. The following test at Lewis Brown Farms consisted of a variety of sensor types all equipped with LoRa radios transmitting at intervals of 10 min for two weather stations and 15 min for three soil moisture sensors. These

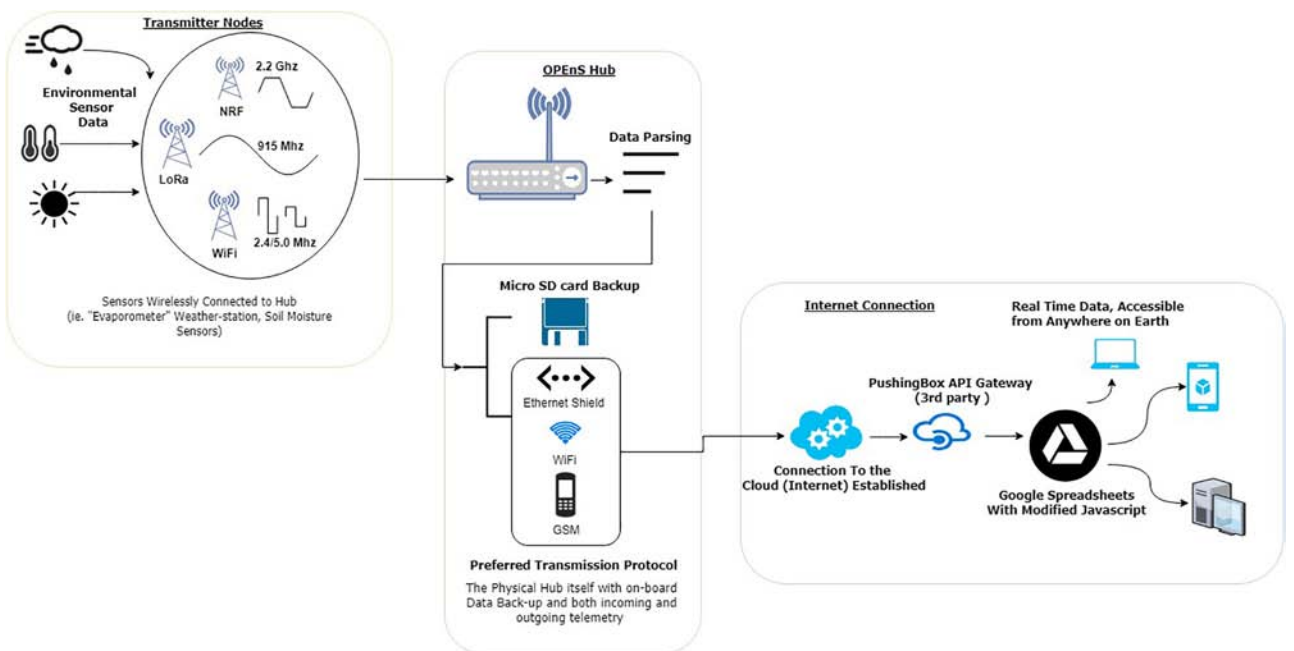



**FIGURE 1 |** The depiction above represents the data pipeline from the point of acquisition in the field to observation on personal devices.

data were broadcast at a maximum distance of 0.45 kilometers to the Hub which was connected via Ethernet. The final field deployment was conducted, again at the H.J. Andrews Experimental Forest, with five weather stations transmitting a variety of environmental conditions at varying distances from the Hub. The longest transmission reached 0.58 kilometers. The format and metadata of the generated Google spreadsheet are outlined in the GitHub repository under "field data," and a map of the field sites showing the Hub in relation to the nodes can be found in **Figure 2** (DeBell, 2019).

## RESULTS

The system was validated in the field at two locations with a total hardware cost of $128 (DeBell, 2019). The first deployment (represented by the purple pin in **Figure 2**) yielded almost 2 months of consistent data transmissions approximately half a kilometer through dense forest. Weather data was reported at 5-min intervals to Google Sheets with less than 10 s of latency. The second deployment demonstrated the capability to receive sensor data from multiple nodes over a period of 4 months. The App Script proved sufficiently dynamic to generate separate tabs for each device and place their respective dataset into the correct columns, producing a spreadsheet populated with over 300,000 data points. The third and final deployment of this study resulted in weather station data received from 5 devices dispersed across the H.J. Andrews Experimental Forest with transmission distances up to half a kilometer. Cumulative data transmission from these three experiments exceeded 400,000 individual points. See the GitHub repository to access field data spreadsheets. The third experiment was cut short due to battery damage at the transmitter nodes caused by a preliminary enclosure design that was permeable to rainwater.

## DISCUSSION

An initial challenge was that the data transmission and the spreadsheet were inherently coupled, which resulted in an end product that lacked flexibility. The spreadsheet assumed the incoming data's order and placed it accordingly, which meant that if the nodes ever changed the data transmitted or the way the Hub started processing data, then the spreadsheet would organize it incorrectly. This problem was resolved by altering the functionality of the nodes to send KVP so that the data could be order-agnostic. This strategy resulted in a spreadsheet that accurately displays data in the correct columns, regardless of the

**FIGURE 2 |** Map of field site at H.J. Andrews Experimental Forest **(A)** and Lewis Brown Farms **(B)**, showing relative distances from nodes to Hub.

order of data received, making the system truly dynamic in the event of dropped radio data packets. However, the transmissions were restricted to only 13 different sensor variables as a result.

Stackable telemetry modules are available for nRF, WiFi, and GSM which plug directly into the header pins of the Adafruit Feather. This requires only minor changes to the transmission code which is under further development on our associated GitHub repository, "Internet of Ag" (Goertzen et al., 2018). The Hub's potential for interchangeable incoming (LoRa, nRF, and Wi-Fi) and outgoing (Ethernet and GSM) transmissions allows for future customization depending on the application of use. This modularity enables transmission over several kilometers at low bandwidths (LoRa and GSM) or shorter distance at much higher bandwidths (Wi-Fi, Ethernet). It is also notable that LoRa technology is still developing and has been expanded to transmit to an ever-growing constellation of satellites, making this technology truly global in its applicability (Telkamp, 2018; Semtech and Lacuna, 2019).

## CONCLUDING REMARKS

The scope of field research using distributed sensors is often restricted by the need to manually retrieve data from remote locations. Moreover, proprietary data logging systems can be prohibitively expensive when scaled to support multiple sensor nodes. To address this challenge, we developed a modular Hub with open-source software, open-hardware and a myriad of telemetry options to push data from the field to Google Sheets in real time, making use of a platform that over two billion people currently use. The OPEnS Hub costs $128, and current ongoing telemetry is free. The Hub has relayed over 400,000 data points through dense forest, proving robust operation under field conditions.

The OPEnS Hub leverages the IoT movement and applies its low-cost and flexible framework to environmental sensing networks. The comprehensive library of code, supporting files, and tutorials on our GitHub helps to break down technical barriers by allowing citizen scientists, farmers, and students to increase the extent and precision of their monitoring efforts without undergoing the complex development process. By expanding access to open-source environmental sensing, the OPEnS Hub broadens the potential for cost-effective precision agriculture, larger field experiments, and new applications for mass data analytics that are yet to be discovered.

## AUTHOR CONTRIBUTIONS

TD designed the framework for the project, wrote the majority of the manuscript, and constructed the physical device. LG and WS contributed to the software library development, and acted as a reference for all software considerations. LL served as chief editor of the manuscript and provided imperative guidance in the construction of the article. CU was the primary mentor on the project. JS served as the head principal investigator on the project.

## FUNDING

This work was supported in part by the USDA National Institute of Food and Agriculture, Hatch project NI18HFPXXXXXG055, the Agricultural Science Foundation at Oregon State University, and Oregon State University Scaled Learning Innovation Grant. All work was done using the resources of the Openly Published Environmental Sensing (OPEnS) Lab at Oregon State University. Student funding for the project was provided in part by the College of Agricultural Science through the beginning and continuing research support programs.

## ACKNOWLEDGMENTS

We are especially indebted to all those who have taken time to provide feedback on this project including Gordon Godshalk, Katherine Darr, and Carolyn Gombert. Additionally, we would like to thank the members of the OPEnS lab, especially Cara Walter, who has provided aid to this project since the very beginning. We also thank our thoughtful Frontiers reviewers and editor for their constructive comments and suggestions. The ability to approach technical writing critically and iteratively is largely a result of the mentorship provided by Dr. Chad Higgins. Lastly, we would like to thank all the faculty, staff and researchers of the Biological and Ecological Engineering department at Oregon State University, for creating a professional and delightful place to work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00137/full#supplementary-material

## REFERENCES

A360 (2019). *Project Collaboration in the Cloud*. Available at: https://a360.autodesk.com/ (accessed October 29, 2018).

Adafruit Industries (2018a). *LoRa Radio FeatherWing - RFM95W 900 MHz*. Available at: https://www.adafruit.com/product/3231 (accessed October 27, 2018).

Adafruit Industries (2018b). *DS3231 Precision RTC FeatherWing - RTC Add-on For Feather Boards*. Available at: https://www.adafruit.com/product/3028 (accessed October 27, 2018).

Adafruit Industries (2019). *Ethernet FeatherWing ID: 3201 - $19.95: Adafruit Industries*. Available at: https://www.adafruit.com/product/3201 (accessed October 27, 2018).

Arduino (2019). *Software*. Available at: https://www.arduino.cc/en/Main/Software (accessed October 29, 2018).

Aufdenkampe, A. K., Damiano, S. G., Hicks, S., and Horsburgh, J. S. (2017). "EnviroDIY Modular Sensors: A Library to give Environmental Sensors a Common Interface of Functions for use with Arduino-Compatible Dataloggers," in *American Geophysical Union, Fall Meeting*, (New Orleans, LA: American Geophysical Union).

Communications (2018). *Wireless, Remote, Hard-Wired, Direct, Or Two-Way Communication*. Available at: http://www.campbellsci.com/communications (accessed October 27, 2018).

CR1000 (2018). *Measurement and Control Datalogger*. Available at: https://www.campbellsci.com/cr1000 (accessed October 27, 2018).

DeBell, T. (2019). *Design and Deployment of a General Purpose, Open Source LoRa to Wi-Fi Hub and Data Logger. 2017 AGU Fall Meeting*. Available at: https://agu.confex.com/agu/fm17/meetingapp.cgi/Paper/221030 (accessed October 29, 2018).

DeBell, T., Goertzen, L., Larson, L., Selbie, W., Selker, J., and Udell, C. (2019). *OPEnSLab-OSU/OPEnS-Hub_Frontiers: The OPEnS Hub*. San Francisco, CA: GitHub, Inc.

EnviroDIY (2018). *GitHub*. Available at: https://github.com/EnviroDIY (accessed February 17, 2019).

Feather (2018). *A complete line of development boards from Adafruit that are both standalone and stackable*. Available at: https://www.adafruit.com/feather?gclid=EAIaIQobChMI79-8uJCc4AIVyiCtBh21dwE5EAAYASAAEgLUa_D_BwE (accessed February 2, 2019).

Ferdoush, S., and Li, X. (2014). Wireless sensor network system design using raspberry Pi and arduino for environmental monitoring applications. *Procedia Comput. Sci.* 34, 103–110. doi: 10.1016/j.procs.2014.07.059

Goertzen, L., Selbie, W., Swopet, Noblek, Leedongj, Debellt, et al. (2018). *OPEnSLab-OSU/InternetOfAg: Loom 1.1 Release for Zenodo*. doi: 10.5281/zenodo.2483255

Hicks, S., Aufdenkampe, A., Damiano, S., and Arscott, D. (2019). *EnviroDIY Mayfly Logger: v0.5b*. doi: 10.5281/zenodo.2572006

International Organization for Standardization C++ (2013). Available at: https://isocpp.org/std/status (accessed October, 2018).

Popper, B. (2017). *Google Announces over 2 Billion Monthly Active Devices On Android. The Verge*. Available at: https://www.theverge.com/2017/5/17/15654454/android-reaches-2-billion-monthly-active-users (accessed April 1, 2019).

PushingBox (2018). *Notifications for Your Internet Of Things Devices*. Available at: https://www.pushingbox.com/ (accessed October 29, 2018).

Raspberry and Pi Foundation (2019). *Teach, Learn, and Make with Raspberry Pi. Raspberry Pi*. Available at: https://www.raspberrypi.org (accessed April 1, 2019).

Semtech and Lacuna (2019). *Receiving Messages from Space*. Available at: https://www.semtech.com/company/press/semtech-and-lacuna-receiving-messages-from-space (accessed February 10, 2019).

SparkFun, E. (2015). *Reducing Arduino Power Consumption - learn.sparkfun.com*. Available at: https://learn.sparkfun.com/tutorials/reducing-arduino-power-consumption/all (accessed February 10, 2019).

Telkamp, T. (2018). *LoRa Transmission From Low Orbit Satellite. The Things Network*. Available at: https://www.thethingsnetwork.org/article/lora-transmission-from-low-orbit-satellite (accessed October 27, 2018).

Wickert, A. D. (2014). The ALog: inexpensive, open-source, automated data collection in the field. *Bull. Ecol. Soc. Am.* 95, 166–176. doi: 10.1890/0012-9623-95.2.68

**Conflict of Interest Statement:** 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.

*Copyright © 2019 DeBell, Goertzen, Larson, Selbie, Selker and Udell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 20 June 2019
doi: 10.3389/feart.2019.00150

# Mojito, Anyone? An Exploration of Low-Tech Plant Water Extraction Methods for Isotopic Analysis Using Locally-Sourced Materials

*Benjamin M. C. Fischer[1,2]*, Jay Frentress[3], Stefano Manzoni[1,2], Sara A. O. Cousins[1,2], Gustaf Hugelius[1,2], Maria Greger[4], Rienk H. Smittenberg[2,5] and Steve W. Lyon[1,2,6]*

[1] Department of Physical Geography, Stockholm University, Stockholm, Sweden, [2] Bolin Centre for Climate Research, Stockholm, Sweden, [3] Free University of Bozen-Bolzano, Bolzano, Italy, [4] Department of Ecology, Environment and Plant Sciences, Stockholm University, Stockholm, Sweden, [5] Department of Geological Sciences, Stockholm University, Stockholm, Sweden, [6] The Nature Conservancy, Delmont, NJ, United States

The stable isotope composition of water ($\delta^{18}O$ and $\delta^{2}H$) is an increasingly utilized tool to distinguish between different pools of water along the soil-plant-atmosphere continuum (SPAC) and thus provides information on how plants use water. Clear bottlenecks for the ubiquitous application of isotopic analysis across the SPAC are the relatively high-energy and specialized materials required to extract water from plant materials. Could simple and cost-effective do-it-yourself "MacGyver" methods be sufficient for extracting plant water for isotopic analysis? This study develops a suite of novel techniques for plant water extraction and compares them to a standard research-grade water extraction method. Our results show that low-tech methods using locally-sourced materials can indeed extract plant water consistently and comparably to what is done with other state-of-the-art methods. Further, our findings show that other factors play a larger role than water extraction methods in achieving the desired accuracy and precision of stable isotope composition: (1) appropriate transport, (2) fast sample processing and (3) efficient workflows. These results are methodologically promising for the rapid expansion of isotopic investigations, especially for citizen science and/or school projects or in remote areas, where improved SPAC understanding could help manage water resources to fulfill agricultural and other competing water needs.

**Keywords: plant water extraction, cryogenic vacuum extraction, stable water isotopes, method comparison, plant sample transport, plant sample storage, low-tech and low-cost**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology, Netherlands*

***Reviewed by:***
*Tim van Emmerik,*
*Delft University of Technology, Netherlands*
*Natalie Ceperley,*
*École Polytechnique Fédérale de Lausanne, Switzerland*

****Correspondence:***
*Benjamin M. C. Fischer*
*benjamin.fischer@natgeo.su.se*

***Received:*** *27 March 2019*
***Accepted:*** *27 May 2019*
***Published:*** *20 June 2019*

***Citation:***
*Fischer BMC, Frentress J, Manzoni S, Cousins SAO, Hugelius G, Greger M, Smittenberg RH and Lyon SW (2019) Mojito, Anyone? An Exploration of Low-Tech Plant Water Extraction Methods for Isotopic Analysis Using Locally-Sourced Materials. Front. Earth Sci. 7:150. doi: 10.3389/feart.2019.00150*

## INTRODUCTION

Stable isotope ratios of water ($\delta^{18}O$ and $\delta^{2}H$), have been successfully used to study atmospheric and hydrological processes around the world for decades (Dansgaard, 1953; Craig, 1961; Sklash et al., 1976). When quantifying catchment water storage and release, water samples of rainfall, soil moisture, groundwater and stream flow are collected (Klaus and McDonnell, 2013), and subsequently analyzed for their isotopic composition and related to various catchment compartments in space and time. Technological innovations such as laser spectroscopy (Kerstel et al., 1999) have drastically reduced the cost of isotope analysis (Lis et al., 2008; Lyon et al., 2009).

This development encouraged hydrologists to collect an ever-increasing number of water samples across space (Fischer et al., 2015, 2017) and time (Berman et al., 2009; von Freyberg et al., 2016). This development also stimulated the use of stable isotopes to explore how vegetation interacts with the atmosphere and the surrounding catchment (Brooks et al., 2010; McDonnell, 2014). To determine which pools of water are used by vegetation and returned to the atmosphere as transpiration, a common approach is to analyze the isotopic composition of plant water, e.g., water found in the root, xylem and/or leaf tissues (Dawson and Ehleringer, 1991; Brooks et al., 2010; Beyer et al., 2016; Goldsmith et al., 2018). Collecting rain or stream water samples for stable isotope analysis is relatively easy with a laser spectroscope where precisions of <0.1‰ for $\delta^{18}O$ and <1‰ for $\delta^{2}H$ are achieved. However, collecting plant water is more challenging because the desired water is part of the living plant tissues and must first be extracted.

Water extraction through squeezing or cooking plant tissue to obtain chemical components and essential oils has been conducted for thousands of years (Kockmann, 2014). More recently, water extraction approaches for stable isotope analysis based on high-tech versions of squeezing or cooking plant material were developed, such as cryogenic vacuum extraction (Dalton, 1989; West et al., 2006; Koeniger et al., 2011), distillation (Vendramini et al., 2007), cryogenic freezing and crushing (Peters and Yakir, 2008), microwave (Munksgaard et al., 2014), or monitored *in situ* using the direct vapor equilibration of water (Wassenaar et al., 2008; Sprenger et al., 2015; Volkmann et al., 2016). The different high-tech methods require a controlled environment to achieve desired accuracy and precision. In addition, each of the aforementioned plant water extraction method is associated with challenges concerning accuracy, precision and repeatability (Orlowski et al., 2016a, 2018; Millar et al., 2018). Extraction time during cryogenic vacuum distillation, for example, affects the apparent stable isotope composition (West et al., 2006). In addition, different "common" methods have the tendency to co-extract various chemical compounds, which can affect the accuracy of laser spectroscopes (West et al., 2010; Millar et al., 2018). As such, there is no general agreement upon optimal or best practice for plant water extraction methods. However, the choice of extraction method may affect study results and represents a subjective and potentially influential factor.

All current plant water extraction methods tend to be resource-intensive, costly, and demand specialized materials and supporting infrastructure. These requirements limit leveraging of citizen science projects which have been beneficial for other isotopic-centered hydrological efforts, such as spatial rainfall sampling during storm events (Good et al., 2014). The relatively high resource demands of plant water extraction is especially problematic when working in remote areas that lack infrastructure where plant water isotopic information could be most useful e.g., in central Tanzania (Koutsouris and Lyon, 2018) or in northern Sweden (Dahlke et al., 2014). Therefore, methodological innovations are necessary for fast, easy, reliable and cost-efficient plant water extraction.

With this perspective in mind, this study develops do-it-yourself "MacGyver" plant water extraction methods using materials found in common kitchens or laboratories around the world and techniques that can be implemented without specialized training. As a proof of concept, we used herbaceous plants species, such as grasses and melon plants, to evaluate the effectiveness of the various techniques and compared the isotopic composition of the extracted water with a "standard" extraction technique, i.e., the cryogenic vacuum distillation. In addition, we simulated the effect of plant sample transport and storage on the plant water isotopic composition. Since all plant water extraction methods have sources of error and uncertainty, which we can control through adequate methodological characterization and clearly defined protocols, our study proposes and tests the hypothesis that simple plant water extraction methods can be used to generate isotopic data with precisions that are comparable to that of more demanding methods.

## MATERIALS AND METHODS

### Plant Material Growth and Initial Processing

Plant water extraction methods were tested on four plant material groups: (A) grass grown indoors (ryegrass; *Lolium perenne*); (B) melon plants grown indoor (water melon; *Citrullus lanatus*); (C) grass grown outdoors on a mown lawn and (D) grass grown outdoors on a grazed pasture (both the pasture and the lawn C and D are a combination of mainly *Poa annua* and *Festuca rubra*) (**Supplementary Figure S1**).

Indoor plant groups grew in trays on an office windowsill. Each tray (22 × 36 × 6 cm) contained 40, free-draining seedling pots (4 × 4 × 5 cm). Each pot contained one turf briquette, which was soaked for 30 min in water to reach field capacity before sowing five grass seeds or two melon seeds. To control isotopic composition, we used two 25-L closed-top barrels filled with tap water at the beginning of the experiment giving a constant and known isotope composition ($\delta^{18}O = 7.97 \pm 0.3$‰ and $\delta^{2}H = -62.02 \pm 0.5$‰) for the initial soaking and subsequent irrigation. One of the trays rested on a kitchen balance connected to an Arduino™ UNO micro-controller with SD-shield (AMC) to measure changes in weight due to evaporation and transpiration at 5 min intervals. In addition, AMC-connected, low-budget soil moisture sensors (HL-69) were installed into one seedling pot to monitor volumetric soil moisture content (%). The AMC information was used to monitor water content and adjust the irrigation scheme, which consisted of irrigation every 2–3 days with 10 ml of water to maintain a moisture content of approximately 60–80% across both trays. Two growing lamps (Plantagen, 6 W, 180 lumen, 265 μmol at 200 mm) were used to supplement light since the experiment was ran in the winter in Sweden (low natural radiation and short days). The lamps were positioned 40 cm above each tray and provided 20 h of light per 24 h cycle. To maintain homogeneous growing conditions, the growing pots in each tray were randomly turned around daily. After 40 days, when grass leaves reached a length of >20 cm and 2–5 mm

width and to melon plants had three leaves >15 cm long, the plants were harvested.

The plant groups grown outdoors consisted of grasses collected from a lawn and a pasture at Stockholm University's Frescati campus. Plant material samples were collected after a rain event during 1 day in October 2018 (autumn). At the moment of sampling, the qualitative soil moisture content was assessed as class 5, i.e., where squelchy noise can be heard when stepping on the ground but no water is visible (Rinderer et al., 2012). At the time of sample collection, both grasses had an average height between 10 and 20 cm, a leaf width larger than 5 mm, and were fibrous. To have a consistent sample size for the various extraction techniques (next section) and isolate potential variability in isotopic composition in the outdoor grass, the grass collected at each site was taken from three 20 × 30 cm plots located within 1 m of each other. The lawn grass and pasture grass samples were composited separately and then cut into 2 cm pieces for water extraction.

Directly after harvest of both the indoor and outdoor plant material, three replicates were prepared for each of the extraction techniques by weighing plant samples (Precisa XT4200C, ± 0.01 g). Due to the low plant weight and to be able to extract sufficient water for stable isotope analysis, a sample consisted of a leaf and stem.

## Plant Water Extraction Methods

### Reference Method (REF) – Cryogenic Vacuum Extraction

The cryogenic vacuum extraction technique described by Koeniger et al. (2011) was used as the reference method (REF method) for the evaluation of the MacGyver methods. This method was chosen because it is considered relatively inexpensive, fast, and reliable when working in well-controlled environments without material procurement limitations. The REF method (**Figure 1a**) uses a heated vial (EXE-I, Exetainer® vial with standard cap and rubber septum, Labco Ltd, Lampeter, United Kingdom) and a cold trap vial (EXE-II, Exetainer® vial with standard cap and rubber septum, Labco Ltd, Lampeter, United Kingdom). We transferred 3 g of plant material into EXE-I immediately after harvest and stored for 1 h at −20°C to avoid decomposition and fractionation. Before extraction, EXE-I and EXE-II were connected through steel capillary tubing (bended syringe 150 × 2 mm, washed and oven dried at 200°C before use) and the entire system evacuated with a hand vacuum pump (Mityvac) to a threshold of 85 kPa. EXE-I was heated for 1 h in a 100°C water bath while EXE-II rested in a Dewar flask containing liquid nitrogen (∼ −196°C). Every 15 min the Dewar flask was refilled with liquid nitrogen. After 1 h the extraction was stopped and EXE-II was sealed with Parafilm. After thawing, the extracted liquid water was pipetted into 2 ml vial for stable isotope analysis.

### Method 1 (MO) - Pestle and Mortar Extraction (Mojito Method)

A mojito is a cocktail where mint leaves are gently mashed with a muddler to extract essential oils. With this in mind, the idea of the mojito methods was born. Interested in water instead of essential oils, we transferred 5 g of plant material to a mortar immediately after harvest and slightly crushed it with a pestle until a mushy, watery puree developed (**Figure 1b**). The puree was squeezed with the pestle to separate fibrous material from the green liquid. The green liquid was transferred into a centrifuge vial and laboratory centrifuged for 30 min at 5000 rpm to separate the water from the grounded plant particles. As an alternative, in remote areas a hand-made centrifuge can be used [e.g., Bhamla et al. (2017)]. After centrifuging, the liquid water was pipetted into 2 ml vial for stable isotope analysis.

### Method 2 (MW) – Household Microwave and Re-sealable Zipper Storage Bags

This method used a standard kitchen microwave and re-sealable zipper storage bags (**Figure 1c**). We transferred 3 g of plant material into a double re-sealable zipper storage bags immediately after harvest and then microwaved at 300 W for 1 min (longer times were not used to prevent the plant material from burning). The extracted water pooled in the bottom of each plastic bag. The extracted water was transferred to a 2 ml vial for stable isotope analysis.

### Method 3 (JJ) – Jam Jar Extraction

An expandable container was constructed by affixing a latex balloon secured with a zip-tie to the top of a clean 200-ml glass jar (**Figure 1d**). We transferred 3 g of plant material immediately after harvest into the jar before sealing the container. The jar was then placed in a 100°C water bath for 1 h (same extraction time as in REF method). During the cooking process, the balloon expands and water condensates against the inner surface. After cooking, the jar was removed from the water bath and allowed to return to room temperature. Once at room temperature, the jar was unsealed and the water in the balloon was pipetted into 2 ml vials for isotopic analysis.

### Method 4 (ICE) – Ice Vacuum Extraction Using Ice Cubes and Cooking Salt

A mix of ice cubes and table salt [weight ratio 3:1 (Arbouw, 2018)] was used for cooling (−20°C) in place of the liquid nitrogen used in the aforementioned REF method (**Figure 1e**). Using the same setup outlined for the REF method, we transferred 3 g of plant material into EXE-I and stored frozen (−20°C) until extraction. As in REF method was conducted by placing EXE-I into a 100°C water bath and EXE-II into the ice-salt mixture. The ice-salt mixture was mixed every 15 min and the temperature was continuously monitored with a laboratory thermometer. After 1 h, EXE-II was removed and sealed with Parafilm®. After thawing, the extracted liquid water was pipetted into 2 ml vial for stable isotope analysis.

### Simulated Transport (REFT) and Storage Impacts (REFS)

We explored the potential impact of transport (i.e., changes introduced after sampling and moving the samples from field to the laboratory) and storage (i.e., changes introduced by delayed analysis) on the stable isotope compositions. Since our goal was to assess the magnitude of errors introduced by transport and storage, we only consider REF method for this experiment.

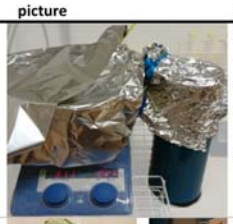
| method | picture | materials | advantage | disadvantage | usability | rank |
|---|---|---|---|---|---|---|
| **a** REF - Cryogenic vacuum extraction |  | •2x Exetainer® vial<br>•Capillary tubing<br>•Hand vacuum pump<br>•100°C water bath<br>•Dewar Flask & liquid nitrogen<br>•Parafilm®<br>•3 g plant material | •Effective water extraction for all plant groups | •High-tech and high-cost methods<br>•Blocked flows near EXE-II due to freezing of capillary affect isotope composition | 0 | |
| **b** MO - Mojito | | •Mortar and pestle<br>•Centrifuge vial<br>•Centrifuge<br>•>5 g plant material | •Easy to use<br>•Fast<br>•Use in field or in laboratory | •Not suitable for fibrous plant material<br>•Centrifuge needed<br>•No extraction efficiency | ++ | 3 |
| **c** MW - Microwave extraction | | •2x re-sealable zipper storage bags<br>•Microwave<br>•3 g plant material | •Easy to use<br>•Fast<br>•Effective water extraction for all plant groups | •Water droplets remained inside bags and likely effect on $E_{eff}$ and isotope composition | ++ | 1 |
| **d** JJ - Jam jar extraction | | •Glass jar<br>•Zip-tie<br>•Balloon<br>•100°C water bath<br>•3 g plant material | •Easy to use<br>•Application nearly everywhere | •Not suitable for fibrous plant material<br>•Some water droplets adhere to the grass and jar | + | 2 |
| **e** ICE - Ice vacuum extraction | | •2x Exetainer® vial<br>•Capillary tubing<br>•Hand vacuum pump<br>•100°C water bath<br>•Ice cubes and salt<br>•Parafilm®<br>•3 g plant material | •Effective water extraction for all plant groups<br>•No additional safety/training aspects needed | •Relative more materials needed including ice cubes<br>•Blocked flows near EXE-II due to freezing of capillary affect isotope composition | + | 1 |
| **f** REFT - Simulated transport | | •Re-sealable zipper storage bags<br>•laboratory oven at a constant 50°C for 1 hour.<br>•5 g plant material<br>•REF method | | •Water loss affect isotope composition | | |
| **g** REFS - Simulated storage | | •Re-sealable zipper storage bags<br>•Freezer -20°C for 1 hour.<br>•5 g plant material<br>•REF method | | •Water loss affect isotope composition | | |

**FIGURE 1 |** The different plant water extraction techniques used: **(a)** REF (cryogenic vacuum), **(b)** MO (mojito, left) and mojito after centrifuge (right), **(c)** MW (microwave), **(d)** JJ (jam jar), **(e)** ICE (ice cube), **(f)** REFT (REF with simulated transport, grass samples after 1 h in the oven to simulate transport of the material) and **(g)** REFS (REF with simulated storage, the grass samples thawing after 1 h in the freezer with exfiltration, i.e., loss of plant water). For each method the different materials needed, advantages, disadvantages, usability (easy, neutral, challenging indicated as ++, +, or 0) and overall rank (best to reasonable indicated as 1–3; based on *Z*-scores) are listed in the respective columns.

To simulate the transport error, 5 g of plant material were transferred immediately after harvest into a re-sealable zipper storage bag and excess air was removed by hand. This bag was placed in a second re-sealable zipper storage bag. After being sealed, the bags were stored in a laboratory oven at a constant 50°C to simulate warm transport conditions (e.g., transport in a car without refrigeration). After 1 h, 3 g of plant material were transferred into EXE-I for plant water extraction REFT (simulated transport using the REF method to extract the plant water).

To simulate the impact of storage on the stable isotope composition, 5 g of plant material were transferred into double-bagged re-sealable zipper storage bags immediately after harvest and stored in a standard freezer at −20°C for 1 h. This test allows assessing how freezing and subsequent thawing affects isotopic composition, which would be a typical concern around storage of plant samples waiting to be processed. After removal form the freezer, 3 g of the plant material were transferred into EXE-I for plant water extraction REFS (storage impact using the REF method to extract the plant water).

The extraction efficiencies ($E_{eff}$) were assessed as the ratio of the weight of the extracted plant water over the weight of the total plant water expressed as:

$$E_{eff} = \frac{W_I - W_E}{W_I - W_D} \quad (1)$$

where the weight of the pre-extraction plant sample is $W_I$, weight of the post-extraction plant sample ($W_E$) and the weight of the plant sample after oven dried at 105°C ($W_D$, weighed repeatedly until there was no change in weight).

## Isotopic Measurement and Extraction Method Comparison

Each extracted plant water sample was pipetted into a 2 mL vial (32 × 11.6 mm screw neck vials with cap and PTFE/silicone/PTFE septa). All water samples were analyzed using a Thermo Scientific isotope-ratio mass spectrometer (IRMS, Delta V Advantage Conflo IV) coupled with a Thermo Scientific Gas Bench II to determine $\delta^{18}O$. Water samples (0.2 ml) were placed in Exetainer® vials and the headspace flushed by a 0.3% $CO_2$-He gas mixture of known isotopic composition. After an equilibration phase of 24 h, the headspace vapor phase was injected 8 times which allowed for a precision of 0.08‰ for $\delta^{18}O$. Deuterium composition was determined by direct injection on the same IRMS, coupled with a Thermo Scientific High Temperature Conversion Elemental Analyzer (TC/EA), equipped with an autosampler (Thermo Scientific AI/AS 3000). Each sample was injected and analyzed 5 times. This allowed for a final precision of 0.7‰ for $\delta^2H$. Vienna Standard Mean Ocean Water (VSMOW) and Standard Light Antarctic Precipitation (SLAP) were used as internal lab standards for both water isotopes. Isotopic composition is reported normalized to the composition of VSMOW, which is defined as 0‰ $\delta^{18}O$ and 0‰ $\delta^2H$. Further, deuterium excess (D-excess) is defined as D-excess = $\delta^2H$-8·$\delta^{18}O$ (Craig, 1961). All extracted plant water samples were analyzed with an IRMS (high-tech, high-cost) contradicting the low-cost character of this study. However, using an IRMS we could avoid issues with solutes released during the extraction that impact the accuracy of laser spectroscopes (West et al., 2010; Millar et al., 2018). In this way, our analysis could focus on the water extraction method, assuming the isotope analysis was reliable.

The isotopic compositions of water from the different extraction methods were compared in dual isotope space (i.e., plotting $\delta^{18}O$ against $\delta^2H$). For each extraction method the average, standard deviation ($S_D$), range (max-min) and the difference of the average composition of a method to the average of the reference plant water extraction were determined. The plant water extraction methods were also compared by *Z-scores* (Wassenaar et al., 2012):

$$Z\text{-}scores = \frac{M_n - M_{REF}}{S_D} \quad (2)$$

where $M_n$ is the isotopic composition of water extracted with the trial method (namely MO, MW, JJ, ICE, REFT, or REFS), $M_{REF}$ is the isotopic composition of the reference method, and $S_D$ is the analysis standard deviation. Instead of using the machine precision as $S_D$ (Wassenaar et al., 2012; Orlowski et al., 2016b), 1.44‰ for $\delta^{18}O$ and 2.2‰ for $\delta^2H$ was used as $S_D$. These $S_D$ are based on the by Millar et al. (2018) reported average $S_D$ [leaf and stem obtained using the cryogenic vacuum extraction method Koeniger et al. (2011)] and were used in this study to better represent the natural variability of stable isotope composition in plant material. An adapted comparison criterion as proposed by Orlowski et al. (2016b) was used to reclassify *Z-scores* such that a *Z-score* <2 were comparable, scores from | 2–5| were considered acceptable and a score >5 was considered unacceptable.

# RESULTS

## Plant Water Extraction

All four MacGyver methods (**Figures 1b–e**) were able to extract 0.5–2 ml water from most plant groups with extraction efficiencies ranging from 0.5 to 1.0 (**Supplementary Figure S2**). In addition, the extraction efficiency varied across methods and there were qualitative differences among methods that are noted as part of the assessment of these MacGyver methods.

The MO method (**Figure 1b**) could extract water from the indoor plant groups. Despite centrifugation, it was not possible to separate all water from the puree and therefore no extraction efficiency was determined (**Figure 1b**). In contrast, outdoor grass plant material were largely fibrous given the autumn such that it was not possible to obtain enough water for stable isotope analysis.

The MW method (**Figure 1c**) was effective for all plant groups but some water droplets remained inside the re-sealable zipper storage bags due to adhesion to the inner side of the bag.

The JJ method (**Figure 1d**) was able to extract water for the indoor grown plant samples but was not able to extract sufficient water (<0.5 ml) from lawn grass samples. Some water droplets could not be piped due to adhesion to the grass and jar.

The ICE method (**Figure 1e**) and the REF method (**Figure 1a**) extracted plant water from all plant groups. As such, there was not a marked difference in the extraction efficiency comparing the MW, JJ, ICE methods and "standard" research-grade extraction technique REF.

Simulating a 1 h car ride at 50°C, the plant weight after transport decreased by 0.1–0.15 g, with the different grass samples (indoor and outdoor) losing 3% and the melon samples losing 10% of total water content, respectively (**Figure 1f**). Simulating the effect of storage (freezing and thawing), the plant weight after thawing decreased by 0.1–0.15 g with the different grass samples (indoor and outdoor) losing 1–10% of the total water content while melon plants decreased by 0.5 g, which is 20% of total water content (**Figure 1g**). From the plant materials used in REFT and REFS, we could extract 1–3 ml of water (extraction efficiency 0.3–0.98, **Supplementary Figure S2**) using the REF method.

## Isotopic Composition of Plant Water Extracted

Considering the REF method, the outdoor grass samples were more depleted in $\delta^{18}O$ and $\delta^2H$ relative to the indoor plant samples (**Table 1**). Moreover, the indoor plant samples showed evaporative enrichment, falling below the global meteoric water line (GMWL, **Figure 2**). In contrast, grass grown outdoors on the lawn or pasture clustered along the GMWL (**Figure 2**). The water used for irrigation of the indoor plants had a constant isotope composition ($\delta^{18}O = -7.97 \pm 0.3$‰ and $\delta^2H = -62.02 \pm 0.5$‰) throughout the experiment and was on the GMWL (**Figure 2**).

The isotopic range and $S_D$ obtained from a given method were 3‰ and 1.5‰ for $\delta^{18}O$, and 17.7‰ and 9‰ for $\delta^2H$ respectively (**Figure 2** and **Table 1**). The average isotope composition of water extracted with any single MacGyver method differed from the average of the REF extracted plant water (**Table 1**).

**TABLE 1 |** Comparison of isotopic composition across plant groups (columns).

| | | $\delta^{18}O$ [‰] | | | | $\delta^{2}H$ [‰] | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Grass indoor** | **Melon indoor** | **Grass lawn** | **Grass pasture** | **Grass indoor** | **Melon indoor** | **Grass lawn** | **Grass pasture** |
| Average($S_D$) | REF | 11.64 (1.7) | 4.89 (1.5) | −8.78 (0.2) | −7.95 (0.3) | −0.3 (9.6) | −23.7(1.8) | −60.9 (2.5) | −55.4 (2.1) |
| | MO | 9.99 (1.2) | 7.56 (1.9) | | | 0.4 (5) | 10.3 (8) | | |
| | MW | 12.42(0.8) | 6.98 (1.1) | −6.75 (1) | −9.61 (0.95) | 10.2 (11.1) | −0.8 (10.7) | −43.4 (4.5) | −58.5 (10.6) |
| | JJ | 11.12 (1.9) | 10.22 (0.6) | | −11.08 (0.9) | 5.9 (8.5) | 3.5 (3.3) | | −52.9 (0.2) |
| | ICE | 10.91 (3.7) | 6.75 (1) | −10.76 (4.1) | −10.09 (1.5) | −8.1 (20.7) | −5.4 (15) | −69.6 (20) | −65.7 (8.2) |
| | REFS | 9.43 (2.2) | 7.58 (0.85) | −10.71 (2.3) | −11.18 (5.1) | −12.9 (15.5) | −6.4 (5.6) | −73.9 (16.2) | −76.3 (33.3) |
| | REFT | 10.34 (3.9) | 7.64 (0.3) | −12.02 (1.1) | −10.6 (3.95) | −12.8 (13.8) | −11.9 (1.4) | −85.5 (7.1) | −73.5 (23.8) |
| Range | REF | 3.38 | 2.78 | 0.44 | 0.59 | 17.3 | 3.4 | 4.7 | 3.8 |
| | MO | 2.36 | 3.4 | | | 8.8 | 14.6 | | |
| | MW | 1.57 | 2.14 | 1.84 | 1.74 | 22.1 | 20.2 | 8.9 | 19.1 |
| | JJ | 3.29 | 1.09 | | 1.29 | 16.9 | 5.9 | | 0.6 |
| | ICE | 6.88 | 1.92 | 5.74 | 2.75 | 36.9 | 27.6 | 28.3 | 16 |
| | REFS | 4.28 | 1.71 | 3.22 | 7.19 | 30.3 | 10.6 | 22.9 | 47 |
| | REFT | 5.56 | 0.69 | 1.53 | 7.77 | 19.5 | 2.7 | 10.1 | 46.7 |

*Rows contain average, standard deviation ($S_D$) and the range (max-min) of $\delta^{18}O$ or $\delta^{2}H$, obtained by extracting plant water using the cryogenic vacuum (REF), mojito (MO), microwave (MW), jam jar (JJ), and ice cube (ICE) extraction methods, and for the simulated effect of transport (REFT) and storage (REFTS). The stable isotope dataset generated and analyzed can be found in* ***Supplementary Table S1****.*

For the different grass samples (indoor, lawn and pasture), most extraction methods provided $\delta^{18}O$ that were comparable (28 out of 32 samples, 88%) to those obtained from the REF method, such that *Z-scores* were less than 2 (**Figure 3**). The remaining extraction methods (4 out of 32 samples, 12%) provided $\delta^{18}O$ values regarded as acceptable (*Z-scores* between 2 and 5, **Figure 3**). Many extraction methods yielded $\delta^{2}H$ values comparable (11 out of 32 samples, 34%) to those from the REF method, such that *Z-scores* were less than 2 (**Figure 3**). Eight out of 32 samples (25%) provided acceptable $\delta^{2}H$ values (*Z-scores* between 2 and 5; **Figure 3**), while the remaining (13 out of 32 samples, 45%) were different from those obtained with the reference method.

The melon plants had fewer isotope values since not all methods were able to extract water (**Figure 3b**). For the methods that were able to extract water, most were comparable to the REF method for $\delta^{18}O$ (4 out of 5 samples) such that *Z-scores* were less than 2 (**Figure 3**). The remaining extraction method (1 out of 5 samples) was acceptable for $\delta^{18}O$ such that the *Z-scores* was between 2 and 5 (**Figure 3**). In contrast, most methods were different from the REF method for $\delta^{2}H$ (4 out of 5 samples), such that *Z-scores* were larger than 5 (**Figure 3**). This lower reliability for $\delta^{2}H$ with melon plants could not be explained by lower extraction efficiencies (**Supplementary Figure S3**).

Most of the samples affected by simulated transport or storage (REFT and REFS) were comparable (7 out of 14, 50%) or acceptable (7 out of 14 samples, 50%) for $\delta^{18}O$ when compared to values obtained from the REF method, such that *Z-scores* were less than 5 (**Figure 3**). In contrast, most of the samples affected by transport or storage (REFT and REFS) were unacceptable for $\delta^{2}H$ (9 out of 14, 65%), such that *Z-scores* were larger than 5. This result indicates a significant influence of transport and storage on the isotopic composition of plant water.

# DISCUSSION

## Usability of the Different Extraction Methods

Each of the investigated MacGyver methods has advantages and disadvantages concerning usability and efficiency to extract water (**Figure 1**). The MO method was easy to use in the field or in laboratory but could not extract water from fibrous plants. The MW extraction was quick and able to extract water from all different plant materials considered, but some water droplets remained in the bag, which likely had an effect on the calculated extraction efficiency and isotope composition of the extracted water. The JJ method can be applied nearly everywhere, including in remote areas with only access to an outdoor stove or fire, but had difficulty to extract water in fibrous plants and water droplets adhering to the leaves and jar, possibly effecting the calculated extraction efficiency and isotope composition of the extracted water. The disadvantage of the ICE method, which is a low-technology variant of the REF method, was that more materials including ice cubes were needed compared to other MacGyver methods. Still, a benefit of the ICE method was that it could extract water from all different plants considered and no additional safety or training aspects were needed (e.g., handling liquid nitrogen -196 °C). Hence, the ICE method could be used safely in citizen science and/or school projects.

Even though the low-technology and low-cost plant water extraction methods were able to effectively and economically extract water from plants, different aspects need to be examined in more detail. In the methods based on heating plant material to release plant water (MW, JJ, and ICE), a fixed extraction time of 1 h was selected from literature values (West et al., 2006; Koeniger et al., 2011). However, as observed by West et al. (2006), the extraction time of cryogenic vacuum extraction affects the stable isotope composition. Therefore, as a next step as

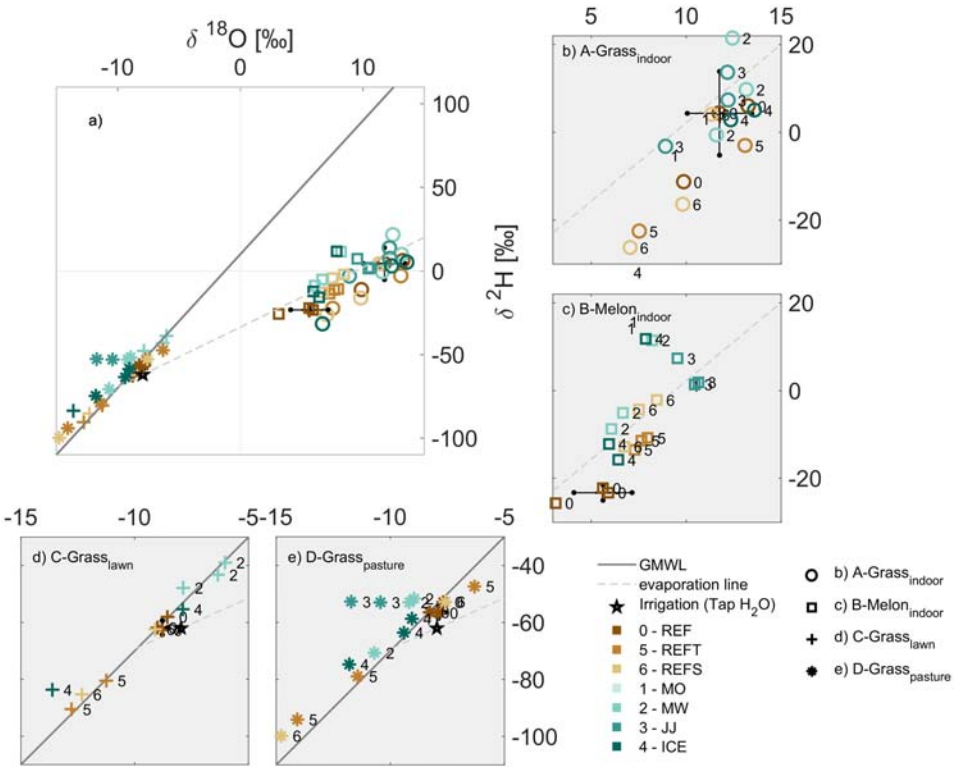



**FIGURE 2 |** Dual isotope space ($\delta^{18}O$ and $\delta^{2}H$) including all extracted plant water samples, the global meteoric water line (solid gray line), the evaporation line calculated as in Benettin et al. (2018) (dashed line) using the median irrigation water (star). The median and error of the REF are indicated as cross hairs with black closed circles. **(a)** Symbols indicate the indoor grass (circles), melon (square), campus lawn grass (cross), grass from the pasture (asterisk) and irrigation water (star). Different colors indicate cryogenic vacuum extraction (REF), simulated transport using REF (REFT), simulated storage using REF (REFS) and the different extraction methods MO, MW, JJ, and ICE (mojito, microwave, jam jar and ice cube, respectively). The side panels show data for each plant group (A-D) separately **(b–e)**.

we seek to develop and differentiate these MacGyver methods, the extraction time should be optimized for each method and investigate the effect of co-extracted chemical compounds on laser spectroscopes. In addition, it is necessary to test all methods (MO, MW, JJ, and ICE) on other types of plant (e.g., trees).

## Method Precision and Plant Water Isotopic Composition

Besides the effectiveness and applicability of each method considered in this study, it is important to assess how the isotopic signatures of the extracted water compare across the different methods.

The different MacGyver plant water extraction methods were able to extract water across a range of plant species and growing conditions (**Figure 2**). The methods seemed to correctly capture the observed evaporative enrichment in the indoor-grown plants and that outdoor grass had an isotopic composition similar to that of the GMWL (**Figure 2**). Single outliers in isotopic composition can be explained by the freezing of the outer and inner part of the syringe (for both ICE and REF) in proximity to the ice or liquid nitrogen, which blocked flows near EXE-II and impacted extraction and eventually the isotopic composition.

Comparing the isotope composition of the different plant water extraction methods across the different plant groups, the ICE method provided results that were closest to the REF method (**Figures 1**, **3** and **Supplementary Figure S4**). However, also a range of plant water stable isotope compositions larger than the precision of the stable isotope analyzer could be noticed (**Figure 2**). The average range of 3‰ for $\delta^{18}O$ for the different MacGyver methods is large but also similar to that reported by Millar et al. (2018) and West et al. (2006) using standard research-grade extraction methods. As such, the MacGyver methods can be used with some confidence knowing that the relative performance regarding final stable isotope composition is equivalent to high-tech and high-cost methods.

In general, the performance of the different methods for $\delta^{18}O$ were acceptable for all plants and for $\delta^{2}H$ for the grasses (**Figure 3** and **Supplementary Figure S4**). The higher deviation of *Z-scores* for $\delta^{2}H$ of melon plant water suggests either that fractionation occurred during sample processing or that the melon plants used in this study experience greater transpiration from the leaves. Unfortunately, due to the limited amount of plant material considered, we could not further investigate this effect through separating the different plant components (e.g., roots, stems and leaves) as was done by e.g., Millar et al. (2018). In addition, it is also uncertain whether each method extracts the same water pool from each plant type or if different pools of water are extracted

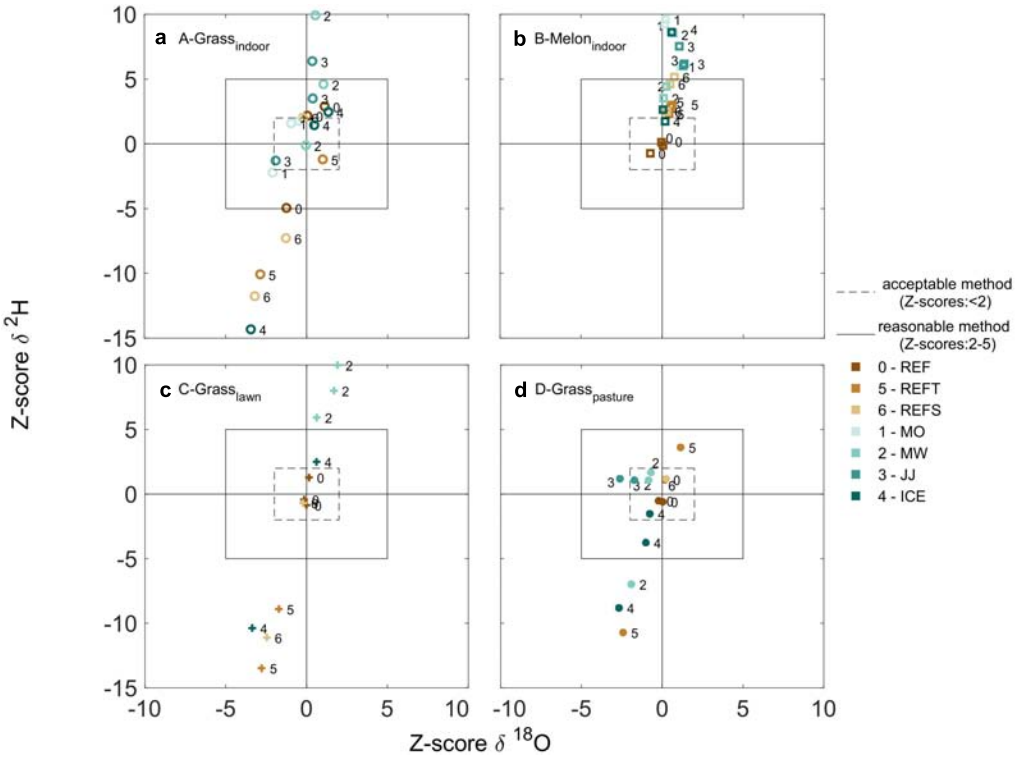



**FIGURE 3 |** For each plant group A-D **(a–d)**, performance of the different plant water extraction methods using the *Z*-scores. Symbols indicate the indoor grass (circles), melon (square), campus grass (cross), grass from the pasture (asterisk), and irrigation water (star). Different colors indicate cryogenic vacuum extraction (REF), simulated transport using REF (REFT), simulated storage using REF (REFS) and the different extraction methods MO, MW, JJ, and ICE (mojito, microwave, jam jar and ice cube, respectively).

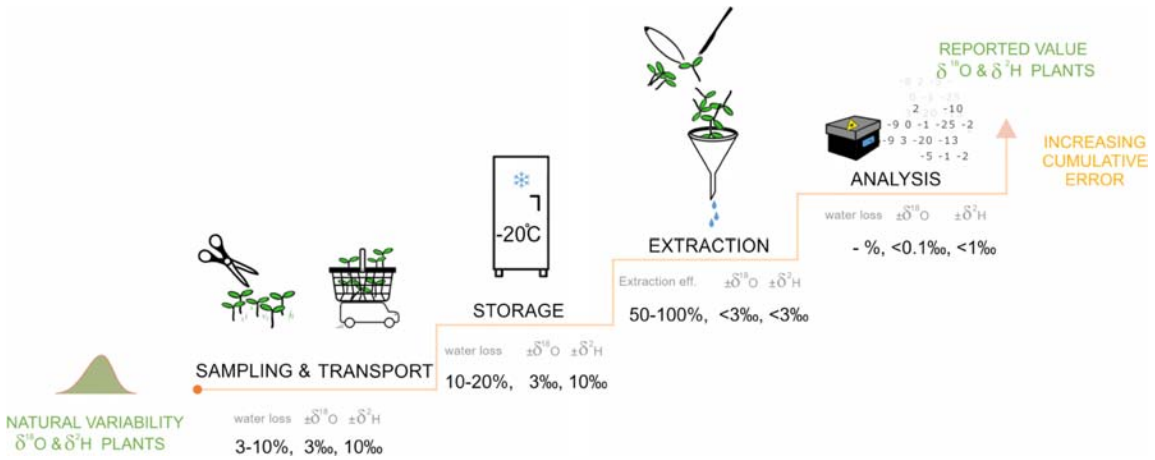



**FIGURE 4 |** The process chain from plant sample collection and transport, storage, plant water extraction and isotope analysis to the final isotope composition. At each step, the potential water loss and associated error in isotopic composition for $\delta^{18}O$ and $\delta^{2}H$ are indicated as % and ‰, respectively. With each step the cumulative error in the isotope composition increases, highlighting the importance to focus not only on the extraction technique but on the full process chain.

according to method (e.g., water from stems vs. leaves, or from xylem vs. intercellular water). This issue begs the question of how representative any bulk extraction method (i.e., cryogenic vacuum distillation) would be when it removes all water from plant tissues.

## Potential Effect of Transport and Storage on the Stable Isotope Composition

Water samples for stable isotope analysis are typically collected in the field using bottles (preferably glass or HDPE-bottles) and hermetically sealed with a cap. Under such conditions, samples

collected today could be analyzed years later (Spangenberg, 2012). Our study, however, highlights that during transport from the field to the laboratory the plant water stable isotope compositions can change considerable (**Figures 1**, **3**, **4**) and to a greater extent than the accuracy of extraction and stable isotope analysis. When plants are collected and not immediately cooled, plant material continues to transpire or lose water via evaporation from the cut surfaces, resulting in water loss of up to 10% compared to the initial sample weight. Hence, it is advisable to cool the plant material directly in the field. In addition, a common practice is to store the collected plant material in a freezer until processing to prevent decomposition and fractionation until the plant water is extracted. When freezing plant material, the cell walls burst. Upon thawing, there can be 10–20% loss of the total plant water impacting the remaining plant water isotopic composition.

Clearly, focusing only on equipment and laboratory techniques while neglecting how consistency in transport and storage can impact isotopic composition can bring about significant misinterpretations. This is where MacGyver methods could provide a remedy or supplement to standard methods such as the cryogenic vacuum extraction (Koeniger et al., 2011) or direct vapor equilibration method as proposed by Millar et al. (2018), by virtue of their speed and ease of use to help bring about consistency.

## CONCLUDING REMARKS

Our results show that simple MacGyver methods can generate isotopic data with a precision that is generally comparable to that of higher-demand research grade methods. In addition, we demonstrated that it is necessary to consider the full process chain from plant sample collection to isotope analysis, as there are several possible sources of errors along this chain (**Figure 4**). All plant water extraction methods have sources of error and uncertainty, which can be controlled through adequate methodological characterization and clearly defined protocols. Therefore, the MacGyver plant water extraction methods presented here are methodologically promising for the rapid expansion of isotopic investigation especially in remote areas with technological limitations or in citizen science and/or school projects that require high safety standards.

## DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

## AUTHOR CONTRIBUTIONS

BF designed and performed all the lab work and wrote the first draft of the manuscript. JF contributed in analyzing the isotopic composition of the water samples. All authors contributed to the revision of the manuscript, read, and approved the submitted version.

## FUNDING

This research was supported by the Bolin RA7 seed funding 2018 and by the Agricultural Water Innovations in the Tropics (AgWIT) project funded by the joint call of the Water Joint Programming Initiative (Water JPI) and the Joint Programming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI) of the European Union and partner countries. SM and SL acknowledge partial support from the Swedish Research Agencies Vetenskapsrådet, Formas, and Sida through the joint call on Sustainability and Resilience – Tackling Climate and Environmental Changes (grant VR 2016-06313). SM also acknowledges Formas (grant 2016-00998).

## ACKNOWLEDGMENTS

We thank all the people who helped and contributed to this study. Martina Hättestrand supporting the laboratory extraction, Anna Scaini for sample logistics, and Dr. Christian Ceccon for the isotope analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00150/full#supplementary-material

**FIGURE S1 |** Growing tray with indoor grass-A in windowsill **(a)**, growing tray with melon-B in windowsill **(b)**, outdoor grass-C lawn with sampling locations indicated by re-sealable zipper storage bags **(c)**, outdoor grass-D pasture (source: Google Street view) **(d)**, pod with indoor grass-A just before harvest **(e)** and pods with melon just before harvest **(f)**.

**FIGURE S2 |** For each plant group (column **A–D**), the boxplots show $\delta^2$H, $\delta^{18}$O, deuterium excesses (D-exe), and water extraction efficiency. The letters indicate the cryogenic vacuum (REF), simulated transport using REF (REFT), simulated storage using REF (REFS), mojito (MO), microwave (MW), jam jar (JJ) and ice cube (ICE) extraction methods. The red line indicates the median of the REF, and the gray lines indicate the analytical standard deviation.

**FIGURE S3 |** The extraction efficiency as a function of $\delta^{18}$O *Z*-scores (top row) and deuterium excess *Z-scores* (bottom row) for the different extraction methods: cryogenic vacuum (REF), simulated transport using REF (REFT), simulated storage using REF (REFS), microwave (MW), jam jar (JJ) and ice cube (ICE) extraction methods. Symbols indicate the indoor grass (circles), melon (square), grass from the lawn (cross), grass from the pasture (asterisk) and irrigation water (star). Different colors indicate cryogenic vacuum extraction (REF), simulated transport using REF (REFT) and simulated storage using REF (REFS).

**FIGURE S4 |** Data of **Figure 3** represented as boxplot to compare the different plant water extraction methods (mojito (MO), jam jar (JJ) and ice cube (ICE) extraction methods) for plant groups A, C, and D (plant group B excluded due to few data points) using the *Z*-score. Individual data points are indicated with circles and boxes indicate the 25 and 75th percentiles.

**TABLE S1 |** The complete stable isotope dataset that was generated and analyzed for this study.

## REFERENCES


Arbouw, E. (2018). *Cold Wine or Delicious Nachos on the Beach (Koud wijntje of heerlijke nacho's op het strand)*. Amsterdam: Volkskrant.

Benettin, P., Volkmann, T. H. M., von Freyberg, J., Frentress, J., Penna, D., Dawson, T. E., et al. (2018). Effects of climatic seasonality on the isotopic composition of evaporating soil waters. *Hydrol. Earth Syst. Sci.* 22, 2881–2890. doi: 10.5194/hess-22-2881-2018

Berman, E. S. F., Gupta, M., Gabrielli, C., Garland, T., and McDonnell, J. J. (2009). High-frequency field-deployable isotope analyzer for hydrological applications. *Water Resour. Res.* 45:W10201. doi: 10.1029/2009WR008265

Beyer, M., Koeniger, P., Gaj, M., Hamutoko, J. T., Wanke, H., and Himmelsbach, T. (2016). A deuterium-based labeling technique for the investigation of rooting depths, water uptake dynamics and unsaturated zone water transport in semiarid environments. *J. Hydrol.* 533, 627–643. doi: 10.1016/j.jhydrol.2015.12.037

Bhamla, M. S., Benson, B., Chai, C., Katsikis, G., Johri, A., and Prakash, M. (2017). Hand-powered ultralow-cost paper centrifuge. *Nat. Biomed. Eng.* 1:0009. doi: 10.1038/s41551-016-0009

Brooks, R. J., Barnard, H. R., Coulombe, R., and McDonnell, J. J. (2010). Ecohydrologic separation of water between trees and streams in a Mediterranean climate. *Nat. Geosci.* 3, 100–104. doi: 10.1038/ngeo722

Craig, H. (1961). Isotopic variations in meteoric waters. *Science* 133, 1702–1703. doi: 10.1126/science.133.3465.1702

Dahlke, H. E., Lyon, S. W., Jansson, P., Karlin, T., and Rosqvist, G. (2014). Isotopic investigation of runoff generation in a glacierized catchment in northern Sweden. *Hydrol. Process.* 28, 1383–1398. doi: 10.1002/hyp.9668

Dalton, F. N. (1989). "Plant root water extraction studies using stable isotopes," in *Structural and Functional Aspects of Transport in Roots Developments in Plant and Soil Sciences*, eds B. C. Loughamn, O. Gašparíková, and J. Kolek (Dordrecht: Springer), 151–155. doi: 10.1007/978-94-009-0891-8_29

Dansgaard, W. (1953). The abundance of O18 in atmospheric water and water vapour. *Tellus* 5, 461–469. doi: 10.1111/j.2153-3490.1953.tb01076.x

Dawson, T. E., and Ehleringer, J. R. (1991). Streamside trees that do not use stream water. *Nature* 350, 335–337. doi: 10.1038/350335a0

Fischer, B. M. C., Rinderer, M., Schneider, P., Ewen, T., and Seibert, J. (2015). Contributing sources to baseflow in pre-alpine headwaters using spatial snapshot sampling. *Hydrol. Process.* 29, 5321–5336. doi: 10.1002/hyp.10529

Fischer, B. M. C., Stähli, M., and Seibert, J. (2017). Pre-event water contributions to runoff events of different magnitude in pre-alpine headwaters. *Hydrol. Res.* 48, 28–47. doi: 10.2166/nh.2016.176

Goldsmith, G. R., Allen, S. T., Braun, S., Engbersen, N., González-Quijano, C. R., Kirchner, J. W., et al. (2018). Spatial variation in throughfall, soil, and plant water isotopes in a temperate forest. *Ecohydrology* 12:e2059. doi: 10.1002/eco.2059

Good, S. P., Mallia, D. V., Lin, J. C., and Bowen, G. J. (2014). Stable isotope analysis of precipitation samples obtained via crowdsourcing reveals the spatiotemporal evolution of superstorm Sandy. *PLoS One* 9:e91117. doi: 10.1371/journal.pone.0091117

Kerstel, E. R. T., van Trigt, R., Reuss, J., and Meijer, H. A. J. (1999). Simultaneous determination of the 2H/1H, 17O/16O, and 18O/16O isotope abundance ratios in water by means of laser spectrometry. *Anal. Chem.* 71, 5297–5303. doi: 10.1021/ac990621e

Klaus, J., and McDonnell, J. J. (2013). Hydrograph separation using stable isotopes: review and evaluation. *J. Hydrol.* 505, 47–64. doi: 10.1016/j.jhydrol.2013.09.006

Kockmann, N. (2014). "Chapter 1 - History of distillation," in *Distillation*, eds A. Górak and E. Sorensen (Boston, MA: Academic Press), 1–43. doi: 10.1016/B978-0-12-386547-2.00001-6

Koeniger, P., Marshall, J. D., Link, T., and Mulch, A. (2011). An inexpensive, fast, and reliable method for vacuum extraction of soil and plant water for stable isotope analyses by mass spectrometry: vacuum extraction of soil and plant water for stable isotope analyses. *Rapid Commun. Mass Spectrom.* 25, 3041–3048. doi: 10.1002/rcm.5198

Koutsouris, A. J., and Lyon, S. W. (2018). Advancing understanding in data-limited conditions: estimating contributions to streamflow across Tanzania's rapidly developing Kilombero Valley. *Hydrol. Sci. J.* 63, 197–209. doi: 10.1080/02626667.2018.1426857

Lis, G., Wassenaar, L. I., and Hendry, M. J. (2008). High-precision laser spectroscopy D/H and 18O/16O Measurements of microliter natural water samples. *Anal. Chem.* 80, 287–293. doi: 10.1021/ac701716q

Lyon, S. W., Desilets, S. L. E., and Troch, P. A. (2009). A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ18O. *Hydrol. Process.* 23, 2095–2101. doi: 10.1002/hyp.7326

McDonnell, J. J. (2014). The two water worlds hypothesis: ecohydrological separation of water between streams and trees?: the two water worlds hypothesis. *Wiley Interdiscip. Rev.* 1, 323–329. doi: 10.1002/wat2.1027

Millar, C., Pratt, D., Schneider, D. J., and McDonnell, J. J. (2018). A comparison of extraction systems for plant water stable isotope analysis. *Rapid Commun. Mass Spectrom.* 32, 1031–1044. doi: 10.1002/rcm.8136

Munksgaard, N. C., Cheesman, A. W., Wurster, C. M., Cernusak, L. A., and Bird, M. I. (2014). Microwave extraction-isotope ratio infrared spectroscopy (ME-IRIS): a novel technique for rapid extraction and in-line analysis of δ18O and δ2H values of water in plants, soils and insects. *Rapid Commun. Mass Spectrom.* 28, 2151–2161. doi: 10.1002/rcm.7005

Orlowski, N., Breuer, L., Angeli, N., Boeckx, P., Brumbt, C., Cook, C. S., et al. (2018). Inter-laboratory comparison of cryogenic water extraction systems for stable isotope analysis of soil water. *Hydrol. Earth Syst. Sci.* 22, 3619–3637. doi: 10.5194/hess-22-3619-2018

Orlowski, N., Breuer, L., and McDonnell, J. J. (2016a). Critical issues with cryogenic extraction of soil water for stable isotope analysis: issues with cryogenic soil water extraction. *Ecohydrology* 9, 1–5. doi: 10.1002/eco.1722

Orlowski, N., Pratt, D. L., and McDonnell, J. J. (2016b). Intercomparison of soil pore water extraction methods for stable isotope analysis: intercomparison of soil pore water extraction methods. *Hydrol. Process.* 30, 3434–3449. doi: 10.1002/hyp.10870

Peters, L. I., and Yakir, D. (2008). A direct and rapid leaf water extraction method for isotopic analysis. *Rapid Commun. Mass Spectrom.* 22, 2929–2936. doi: 10.1002/rcm.3692

Rinderer, M., Kollegger, A., Fischer, B. M. C., Stähli, M., and Seibert, J. (2012). Sensing with boots and trousers — qualitative field observations of shallow soil moisture patterns. *Hydrol. Process.* 26, 4112–4120. doi: 10.1002/hyp.9531

Sklash, M. G., Farvolden, R. N., and Fritz, P. (1976). A conceptual model of watershed response to rainfall, developed through the use of oxygen-18 as a natural tracer. *Can. J. Earth Sci.* 13, 271–283. doi: 10.1139/e76-029

Spangenberg, J. E. (2012). Caution on the storage of waters and aqueous solutions in plastic containers for hydrogen and oxygen stable isotope analysis: stable isotope variation of water stored in plastic bottles. *Rapid Commun. Mass Spectrom.* 26, 2627–2636. doi: 10.1002/rcm.6386

Sprenger, M., Herbstritt, B., and Weiler, M. (2015). Established methods and new opportunities for pore water stable isotope analysis: pore water stable isotope analysis. *Hydrol. Process.* 29, 5174–5192. doi: 10.1002/hyp.10643

Vendramini, P. F., Sternberg, L., and da, S. L. (2007). A faster plant stem-water extraction method. *Rapid Commun. Mass Spectrom.* 21, 164–168. doi: 10.1002/rcm.2826

Volkmann, T. H. M., Kühnhammer, K., Herbstritt, B., Gessler, A., and Weiler, M. (2016). A method for in situ monitoring of the isotope composition of tree xylem water using laser spectroscopy. *Plant Cell Environ.* 39, 2055–2063. doi: 10.1111/pce.12725

von Freyberg, J., Studer, B., and Kirchner, J. W. (2016). A lab in the field: high-frequency analysis of water quality and stable isotopes in streamwater and precipitation. *Hydrol. Earth Syst. Sci. Discuss.* 2016, 1–32. doi: 10.5194/hess-2016-585

Wassenaar, L. I., Ahmad, M., Aggarwal, P., van Duren, M., Pöltenstein, L., Araguas, L., et al. (2012). Worldwide proficiency test for routine analysis of δ2H and δ18O in water by isotope-ratio mass spectrometry and laser absorption spectroscopy. *Rapid Commun. Mass Spectrom.* 26, 1641–1648. doi: 10.1002/rcm.6270

Wassenaar, L. I., Hendry, M. J., Chostner, V. L., and Lis, G. P. (2008). High resolution pore water δ 2H and δ18O measurements by H2O (liquid) -H2O

(vapor) equilibration laser spectroscopy. *Environ. Sci. Technol.* 42, 9262–9267. doi: 10.1021/es802065s

West, A. G., Goldsmith, G. R., Brooks, P. D., and Dawson, T. E. (2010). Discrepancies between isotope ratio infrared spectroscopy and isotope ratio mass spectrometry for the stable isotope analysis of plant and soil waters. *Rapid Commun. Mass Spectrom.* 24, 1948–1954. doi: 10.1002/rcm.4597

West, A. G., Patrickson, S. J., and Ehleringer, J. R. (2006). Water extraction times for plant and soil materials used in stable isotope analysis. *Rapid Commun. Mass Spectrom.* 20, 1317–1321. doi: 10.1002/rcm.2456

**Conflict of Interest Statement:** 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.

*Copyright © 2019 Fischer, Frentress, Manzoni, Cousins, Hugelius, Greger, Smittenberg and Lyon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 17 July 2019
doi: 10.3389/feart.2019.00184

# Mobile Monitoring—Open-Source Based Optical Sensor System for Service-Oriented Turbidity and Dissolved Organic Matter Monitoring

*Robert Schima[1,2]*, Stephan Krüger[3], Jan Bumberger[2], Mathias Paschen[1], Peter Dietrich[2,4] and Tobias Goblirsch[2]*

[1] Chair of Ocean Engineering, Faculty of Mechanical Engineering and Marine Technology, University of Rostock, Rostock, Germany, [2] Department of Monitoring and Exploration Technologies, UFZ - Helmholtz Centre for Environmental Research, Leipzig, Germany, [3] Chair of Soil Resources and Land Use, Institute of Soil Science and Site Ecology, Technische Universität Dresden, Tharandt, Germany, [4] Center for Applied Geoscience, Eberhard-Karls-University of Tübingen, Tübingen, Germany

The protection and sustainable use of aquatic resources require a better understanding of fresh water sources, limnic ecosystems, and oceans. The effects of global change, intensive use of natural resources and the complex interactions between humans and the environment show different effects at different scales. Current research approaches are not sufficient to appropriately take account of the heterogeneity and dynamics of aquatic ecosystems. A major challenge in applied environmental research is to extend methods for holistic monitoring and long-term observation technologies with enhanced resolution over both space and time. In this study, turbidity and the content of dissolved organic matter (DOM) are key parameters, as they are of importance for assessing the health of aquatic ecosystems and the state of ecosystem services (e.g., the provision of drinking water). Photonics and optical sensors as well as integrated circuits and open-source based components open interesting possibilities to overcome the current lack of adaptive and service-oriented sensor systems. An open source based optical sensor system was developed, which enables a user-specific, modular and adaptive *in-situ* monitoring of the turbidity and the dissolved organic substance content almost in real time. Quantification is based on attenuation or transmission measurements with two narrowband LEDs and corresponding detectors in the ultraviolet (DOM content) and infrared range (turbidity) of the electromagnetic spectrum. The developed *in-situ* sensor system shows a very high agreement with the results obtained using a laboratory photometer but with less methodological effort. First tests carried out in the area close to the city of Leipzig (Saxony, Germany) show promising results. The *in-situ* sensor system is able to acquire the optical attenuation with a sampling rate up to $0.1\ \mathrm{Hz}$. Due to the fact that data is visualized directly with the help of web services, even the quality of data collection can be improved by assisting the selection of sampling points or a direct spatio-temporal data feedback. What this approach illustrates is the fact that open-source technologies and microelectronics can now be used to implement resilient and promising sensor systems that can set new standards in terms of performance and usability within applied environmental research.

**Keywords: photonic sensing, *in-situ* measurements, assisted monitoring, attenuation sensor, internet of things, water quality**

OPEN ACCESS

***Edited by:***
*Peter M. Marchetto, University of Minnesota Twin Cities, United States*

***Reviewed by:***
*Tim van Emmerik, Delft University of Technology, Netherlands*
*Ahmed M. ElKenawy, Mansoura University, Egypt*

****Correspondence:***
*Robert Schima*
*robert.schima@uni-rostock.de*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *28 February 2019*
***Accepted:*** *01 July 2019*
***Published:*** *17 July 2019*

***Citation:***
*Schima R, Krüger S, Bumberger J, Paschen M, Dietrich P and Goblirsch T (2019) Mobile Monitoring—Open-Source Based Optical Sensor System for Service-Oriented Turbidity and Dissolved Organic Matter Monitoring. Front. Earth Sci. 7:184. doi: 10.3389/feart.2019.00184*

# 1. INTRODUCTION

The conservation and use of aquatic resources necessitate a better understanding of freshwater sources, limnic ecosystems, and oceans to sustainably secure the livelihood of a steadily growing world population. A major challenge is seen in the fact that global change and the consequences of human action show different effects at different scales (Chapman, 1996). State of the art research approaches and scientific measurement methods are limited to large-scale measurement campaigns carried out by scientific or governmental institutions. Due to the high methodical effort and the relatively low spatio-temporal coverage, such approaches are not yet feasible enough to provide a monitoring solution to address the heterogeneity and dynamics of aquatic ecosystems in an appropriate manner (Wiemann et al., 2018).

In the field of applied environmental research, different approaches exist to measure and visualize processes in the environment and their effects on the ecosystem. These approaches extend over several scales. Starting with satellite-based earth observation, airborne, or unmanned aerial vehicles (UAVs), environmental data acquisition is also achieved by deploying sensor networks, autonomous systems, or conducting classical manual field sampling using appropriate sampling and measurement technologies for the monitoring.

What is urgently needed, however, are methods that go beyond classical established environmental research. With the help of miniaturized, integrated sensors, user groups outside science could also be integrated into the process of collecting environmental data (Citizen Science). This requires suitable sensor systems that are both easy to use and easy to process and visualize the data later on. In current research, there are approaches that investigate new strategies of environmental data collection with a similar motivation (Hut et al., 2016; Brewin et al., 2017; Seibert et al., 2019). It is particularly important to develop approaches for a wide range of users when determining, controlling and ensuring good water quality in the long term (Lockridge et al., 2016). Under these aspects, an efficient sensor system must not only provide scientifically valid data. For the later utilization of the data, appropriate interfaces must be created for visualization, online availability of the data and for the creation of a spatial reference for measurement (e.g., via GPS).

Therefore, this paper presents and evaluates an approach for such a monitoring based on an optical measuring system for the determination of dissolved organic carbon and turbidity.

## 1.1. Water Quality Assessment

Especially with regard to the process dynamics and heterogeneity of aquatic ecosystems, a comprehensive monitoring of these effects remains to be a challenging issue. This results in a strong pull to develop adaptive survey and monitoring strategies as well as tools to observe even complex ecosystems of large scale and over a longer period of time. In this connection, it is of particular interest to map the heterogeneity as well as the dynamics of processes in a discrete manner to address the spatio-temporal interdependencies of aquatic systems appropriately. Although the performance of sensors and sensor systems has considerably increased during the last years, the integration of data as well as the provision of gathered information have to be improved to achieve a more wide spread use of optical sensors and sensor systems in practice. In terms of optical tools for aquatic monitoring a broad review is given by Moore et al. (2009). The motivation of this work is the development and implementation of an *in-situ* sensor probe prototype for the optical detection of dissolved organic carbon (DOC) in aquatic media.

### 1.1.1. Dissolved Organic Carbon (DOC)

Dissolved organic carbon is a sum parameter and represents all organic compounds dissolved in water. It plays a key role in the assessment of water status and load (Guo et al., 1995). In the field of water and environmental research, the interest in observing and documenting short and long-term trends in DOC concentrations of surface and drinking water is of great importance (Kolka et al., 2008). Current field devices for the recording of DOC are very reduced with regard to their modifiability by the user and are not designed for near real-time data processing. Another limitation of current system is the insufficient temporal-spatial resolution. The implementation of mobile monitoring strategies is therefore difficult to achieve.

### 1.1.2. Turbidity

Turbidity describes how clear a selected volume of water is. In this context, turbidity refers to the presence of suspended solids. The more turbid a water is, the less light can be transmitted through it as a result of the suspended solids. With regard to ecosystem processes, increased turbidity means, for example, that less light is available for photosynthesis (Bass et al., 1995). The interaction therefore has an effect on aquatic fauna, but also on fish abundance and distribution. Furthermore, turbidity is often related to the nutrient content, which strongly influences an ecosystem (Chapman, 1996). In addition to the ecological aspects, turbidity is of great importance as a corrective in applied measurement techniques. Knowledge of turbidity is important to enable appropriate corrective measures when working with optical instruments in general (turbidity correction).

## 1.2. Sensing as a Service

To this end, an open-source based, modifiable *in-situ* probe would represent a promising possibility to realize a mobile and cross-scale monitoring approach, which allows a higher spatial and temporal resolution with lower maintenance and acquisition costs. For this purpose this work aims at the proof of concept and the development of an optical sensor probe as part of a sensor system for the detection of DOC under field conditions based on a spectrometric measurement.

As another aspect of this work, it will be investigated to what extent cost-effective open-source based platforms are suitable for realizing complex environmental information and sensor systems. Raspberry Pi, Arduino and other technological developments, especially in the field of IoT and miniaturized automation, have fundamentally changed the working methods of different user groups and the way inventions can be realized with very limited resources.

In addition to the construction of a field-capable sensor unit, the development of powerful and fast data processing structures is another motivation of this work in order to provide reliable monitoring data close to real-time. To achieve this goal web based services were developed and implemented. In this connection, the definition of standards as well as the establishment of suitable interfaces were key elements to create an holistic process chain from the acquisition of a single measurement to the provision of reliable information for decision making, which is called *Sensing as a Service*.

In contrast to common monitoring approaches using stationary observation data, mobile monitoring approaches provide the potential of a higher spatial resolution in terms of identifying the heterogeneity of an ecosystem. For a service-oriented implementation of the necessary computation, especially during the field measurements, an abstract data model was formulated called the *Object Specific Exposure* (OSE) (Schima et al., 2017; Schima, 2018). The term exposure is used to describe the relationship between a state variable, e.g. the measured concentration of organically dissolved carbon, and the respective period or reference space of the measurement. A mobile measurement series along a river section must therefore be interpreted differently than a stationary measurement in the same period. A simple averaging, however, is not sufficient, since the integral information contained with regard to the temporal and spatial peculiarities would be lost. This approach serves as a holistic representation of an exposure, e.g., an object exposed to a concentration ($c$) of a substance at a specific position ($x, y$), at a specific depth ($z$) and for a specific amount of time ($t$) in an aqueous media as shown in Equation (1).

$$\mathrm{OSE} = \int_{x_0}^{x_i} \int_{y_0}^{y_i} \int_{z_0}^{z_i} \int_{t_0}^{t_i} c(x, y, z, t)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\, \mathrm{d}t \qquad (1)$$

In view of mobile sampling, it is necessary to establish a corresponding spatial and temporal reference for each measurement of a state variable [e.g., concentration ($c$)]. The exposure is then composed integrally (see Equation 1). The OSE formula is a considerable help for the development of mobile sensors since applied as a design and system paradigm, it ensures that all descriptive data of a measurement conversion is generated at the sensor level. The optical *in-situ* measurement of absorption is thus supplemented by information such as location, water depth, time, and system status. Furthermore, a strict monitoring paradigm ensures that, for example, a measured value is only collected at full minute intervals. This allows the direct comparison of data from different sensor systems with no need for any further data harmonization.

At the same time, the usability of the collected data is increased. Following the OSE paradigm and using standardized interfaces for data transmission, real-time environmental information systems can be set up. This makes it possible, for example, to create a map display in the field while collecting the data or to visualize time series on a smartphone or tablet. The joint development of hardware and software leads to a so-called *Assisted Monitoring*, which is particularly useful for mobile applications as shown in the following.

#### 1.2.1. Assisted Monitoring: Mobile App and Web Service

The approach of an *Assisted Monitoring* aims at providing the user with reliable information on the target environmental parameters and sensor system states during data collection (see **Figure 1A**). To do so, the system consists of an optical sensor head that is connected to a control and communication unit. Via a Bluetooth interface it is possible to configure the measuring system with the help of a mobile terminal device and an app (see **Figure 1B**), to start measurements manually or automatically and to send the measured values via a web interface, e.g., for online visualization (see **Figure 1C**).

The mobile app empowers the user to modify the data acquisition and to configure the sensor system. Therefore, single measurements or an automated monitoring can be easily achieved. The app provides an initial data evaluation and a real-time data visualization. During the measurement, data can be stored locally or send to a web service.

On the server side, various services take over the forwarding, processing and visualization of the data. A browser based web service provides a dashboard for real-time data visualization including a web map service for initial interpretation of the data, e.g., in the field for *ad-hoc* adjustments of the sampling strategy. Here again, the sensor system development according to the OSE paradigm allows a straight data stream processing since every single measurement is geo referenced and time synchronized. The data stream process requires the following process stack per each measurement:

1. Sensor level:
   (a) Routine to acquire all data from the sensor system (e.g., temperature, pressure, absorption,...)
   (b) Routine to generate a message string containing all predefined values of interest
   (c) Routine to configure communication module according to the used communication protocol
2. Server level:
   (a) Message-oriented middleware (MOM) – A message broker to handle the incoming messages, e.g., RabbitMQ
   (b) A data base to store the data, e.g., InfluxDB
   (c) A front end to query and visualize the data, e.g., Grafana

In order to feed conventional environmental data processing, a data export function can be used to export the data in an appropriate format, e.g., as a comma separated values file.

## 2. MATERIALS AND METHODS

### 2.1. Measurement Principle

The underlying measuring principle is an optical transmission measurement (respectively attenuation) of water due to different ingredients and dilutes (Preisendorfer, 1976; Bass et al., 1995; Hoge et al., 1995). This measuring principle is well-established in laboratory analysis, so that regulations exist for the detection of certain substances in water. To comply with normed methods and standards, the use of DIN standards is common in Germany.

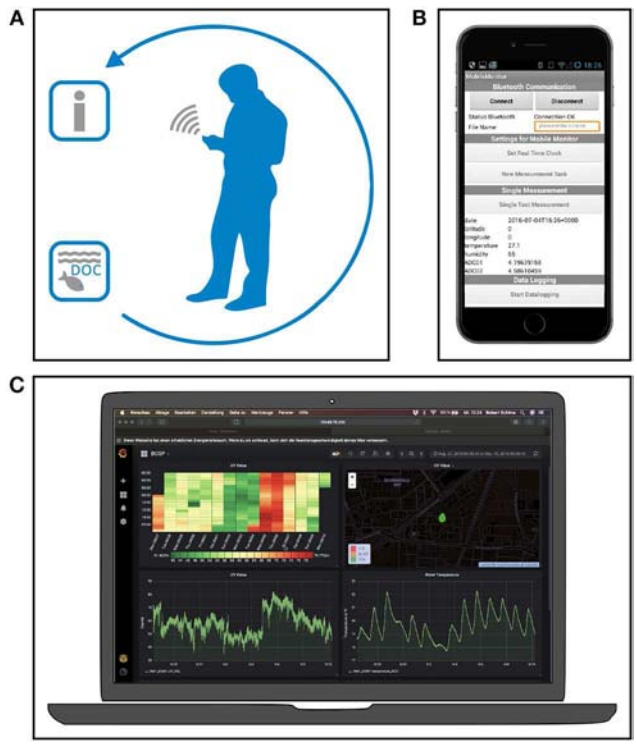



**FIGURE 1 |** *Assisted Monitoring* based on mobile apps and web services will lead to new perspectives in environmental monitoring and observation. The development of service-oriented infrastructures will accelerate the data acquisition toward the provision of reliable information during field measurements **(A)**. The visualization of data near real-time and the usage of mobile devices **(B)** in the field will increase the spatio-temporal resolution and the quality of environmental monitoring in general **(C)**.

DIN (Deutsches Institut für Normung) is the German Institute of standards. The relevant standards for this study are the DIN 38404-3 (2005) for attenuation in the UV wavelength range, DIN EN 1484 (1997) for DOC analysis and DIN EN ISO 7027 (2000) for turbidity approximation.

With respect to DIN 38404-3 (2005) the spectral absorption coefficient $\alpha_\lambda$ or rather the Extinction $E_\lambda$ at $\lambda = 254$ nm is used as a summarizing method for the determination of dissolved organic carbon in aquatic media. Therefore, the attenuation of UV light passing through a sample is detected by absorption. This measured absorbance of the water sample serves as a measure of the concentration of chromophorically dissolved organic substances in water, such as DOC. The general basis for such quantitative absorption measurement is the Beer-Lambert law according to Equation (2):

$$E_\lambda = lg\left(\frac{I_0}{I_1}\right) = \epsilon_\lambda\, c\, d \tag{2}$$

Here, $E_\lambda$ represents the Extinction, $I_0$ the intensity of emitted light in $\mathrm{W\,m^{-2}}$, $I_1$ the intensity of transmitted light in $\mathrm{W\,m^{-2}}$, $\epsilon_\lambda$ the extinction coefficient in $\mathrm{m^2\,mol^{-1}}$, $c$ the concentration of absorbing material in $\mathrm{mol\,l^{-1}}$ or $\mathrm{km\,m^{-3}}$, and $d$ the path length in m. Based on this considerations, the development of an optical *in-situ* sensor probe for measuring the attenuation will be described in the following.

## 2.2. Hardware Description

### 2.2.1. Sensor Probe

According to DIN 38404-3 (2005) the spectral absorption coefficient $\alpha_\lambda$ at $\lambda = 254$ nm can be used to calculate the DOC content in aquatic media. Therefore, the probe consists of an ultraviolet light emitting diode (UV-LED) as emitter and a UV photodiode as detector. Since it is not possible to filter the medium in an *in-situ* measurement as recommended by DIN EN 1484 (1997), for an *in-situ* measurement of DOC the influencing turbidity must be compensated or determined. As shown in Huang et al. (1992) and Liu and Dasgupta (1994), a spectrometric measurement at $\lambda = 860$ nm can be used to compensate the turbidity.

Dual wavelength measurement is implemented with two sensor units, a UV sensor unit for measuring DOC at $\lambda = 254$ nm and an infrared (IR) sensor unit for measuring turbidity at $\lambda = 860$ nm. To perform the photometric measurement, each sensor unit includes an emitter and detector as shown in **Figure 2A**. An UV-LED (UVTOP250-BL-TO18 from ROITHNER LASERTECHNIK GmbH, Vienna, Austria) ($\lambda_{max} = 254$ nm) is installed on the emitter side of the UV unit. An IR LED (SFH 4557 from OSRAM Opto Semiconductors GmbH, Regensburg, Germany) ($\lambda_{max} = 850$ nm) is installed for the emitting IR unit. In addition, the IR emitting sensor head includes a constant current source at $I_{LED} = 100$ mA and a voltage regulator. The UV emitting sensor head contains a constant current source at $I_{LED} = 20$ mA. The detector for the UV sensor unit is a UV silicon carbide photodiode (SIC01S-C18 from ROITHNER LASERTECHNIK GmbH, Vienna, Austria) with peak spectral sensitivity at $\lambda = 254$ nm. An IR LED (SFH 4850 E7800 from OSRAM Opto Semiconductors GmbH, Regensburg, Germany) is used as a detector for the IR channel. Both detector heads also include a 16-bit analog-to-digital converter (ADS1115 from Texas Instruments Incorporated, Dallas, USA). In addition, the UV receive sensor head contains a single supply instrumentation amplifier. For controlling the system, a Raspberry Pi is used. A transistor and an optocoupler are installed for each channel to control the UV LED and IR LED. A Python script running on the Raspberry Pi is responsible for managing the switching characteristics and measurement. The emitter and detector sensor heads are connected to the control unit via 4-pole cable (LifYDY 4 x 0.10 qmm of kabeltronik Arthur Volland GmbH, Denkendorf, Germany). The entire system is supplied by a battery pack (see **Figure 2**). With one battery charge, data can be recorded for up to 24 h. For later application the sensors are placed inside the media. The control unit (topside unit) is designed to ensure user interaction via Bluetooth, GPS positioning and data transmission via WiFi.

The housing of the probe is made of stainless steel. The system consists of two cylindrical sensor heads, one for the emitter and one for the detector. Both are mounted opposite of each other on a connecting bridge. Each sensor head has a sapphire glass window at the narrow end of the cylinder. The wide ends of the

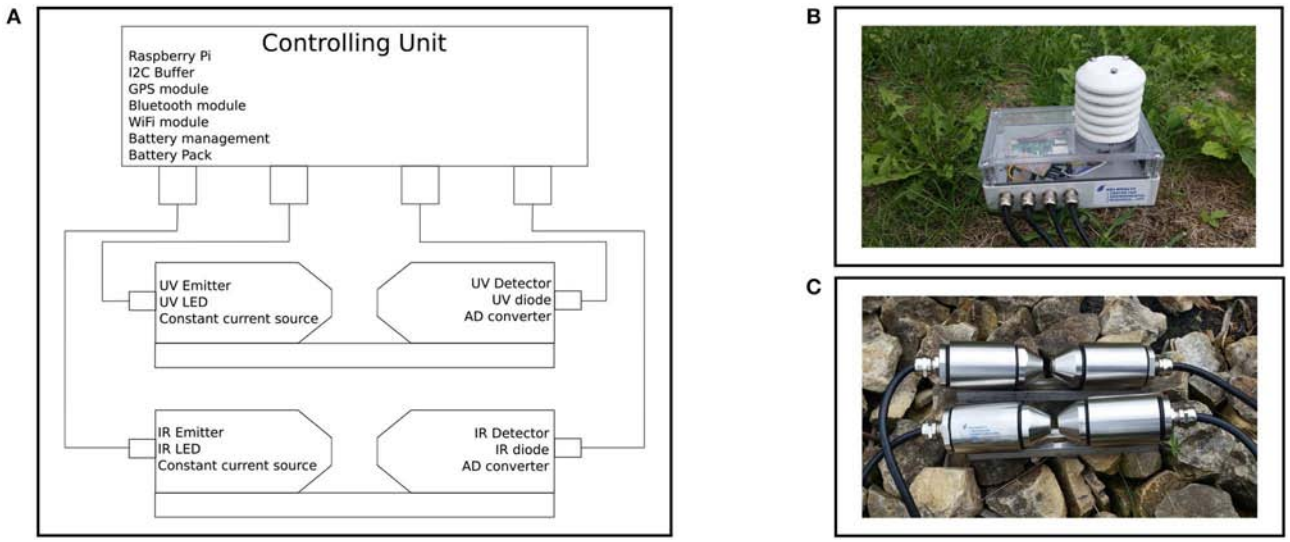



**FIGURE 2 | (A)** shows a schematic drawing of the sensor system. **(B)** shows the control unit with the Raspberry Pi, GPS receiver, and WiFi module as well as the battery pack. Four cables are connected to the emitters and detector units of the UV channel (DOM) and IR channel (turbidity); the white attachment on the top of the control unit is used to measure temperature and humidity. The illustration in **(C)** shows the IR sensor unit at the top and the UV sensor unit at the bottom, with both receiving and transmitting sensor heads mounted opposite each other on a metal guide. Inside there is an LED as emitter and a photodiode as detector in combination with an AD-converter. A constant current source and voltage regulator are integrated on the emitter side for stable operation. The control unit and the sensor modules are connected by cables.

heads have a waterproof cable gland. Including the cable gland, the sensor measures a height of 4.4 cm, a length of 20.8 cm and a width of 3.9 cm, and weighs 119.2 g (see **Figure 2C**). The Python script running on the Raspberry Pi controls the time, duration, and iteration of the measurement. Both, the UV and IR units are activated simultaneously. Accordingly, as soon as the LED is supplied with power, it begins to emit light that goes through the medium and hits the detector (diode). Due to the relationship between the concentration of the measured compounds and the intensity of the transmitted light, the light intensity incident on the receiving diode varies (cf. Lambert-Beer law). In the photo diode, the incident light generates a current dependent on the intensity of the light, so that the voltage can be measured via a trans-impedance amplifier. This voltage is converted into a digital signal by the analog-to-digital converter and evaluated by the control unit (Raspberry Pi). Once the DOC and turbidity units are calibrated, the turbidity and consequently the content of DOC can be determined based on the transmission measurement.

## 2.3. Laboratory Experiments

### 2.3.1. Sensor Calibration

Both sensor units are calibrated directly in the medium. Potassium hydrogen phthalate as recommended in DIN EN 1484 (1997) is used to prepare a dilution series for DOC calibration. Even though formazine is the standard solution for the calibration of turbidity probes according to DIN EN 1484 (1997), the use of formazine is avoided due to its toxic properties. Liu and Dasgupta (1994) used milk in their experiments to produce turbidity. Since milk is an organic product, it influences the DOC measurement and therefore cannot be used for this experimental set-up. Clifford et al. (1995) used the inorganic component Fuller's Earth. Besides milk, Fuller's Earth is unsuitable due to its adsorption properties, which can lead to complications with potassium hydrogen phthalate. Talcum powder (Talcum Powder -350 MESH from Sigma-Aldrich Co. LLC., St. Louis, USA) ($Mg_3Si_4O_{10}(OH)_2$) is used as a non-adsorbent, inorganic, non-water-soluble material that does not influence the DOC measurement, in order to cause turbidity in this experimental setup. To find out how the turbidity and DOC measurement units react to different concentrations, both are mounted upside down with the receiving diode on top in a glass vessel to avoid the influence of stray light during measurement. To avoid mutual interference between the two sensor units, each sensor is inserted into the calibration vessel separately.

To prepare the dilution series, distilled water with a content of total organic carbon (TOC) of $c = 0.003\,\mathrm{mg\,l^{-1}}$ is used as the base medium for calibration. A glass vessel is filled with $V = 0.8\,\mathrm{l}$ distilled water. Due to this low TOC concentration, no filtering of the distilled water is required. It is intended to measure DOC concentrations of $c_{\mathrm{DOC}} = 0, 5, 15, 25, 50, 75, 100\ \mathrm{mg\,l^{-1}}$. Therefore, a stock solution is prepared with $c = 1000\,\mathrm{mg\,l^{-1}}$ potassium hydrogen phthalate. To produce the intended DOC concentrations, the following amount of stock solution is added to the $V = 0.8\,\mathrm{l}$ zero solution (see **Table 1**).

In order to be able to measure turbidity in a comparable way, formazine has become established as the turbidity standard liquid in laboratory analysis. There are different turbidity units, all of which refer to dilutions of this turbidity standard liquid, but some of which reflect different phenomena. This study refers to the turbidity unit FAU (Formazine Attenuation Units), since a transmitted light measurement, as defined in DIN EN ISO 7027 (2000), is carried out to determine this unit. Since this

**TABLE 1 |** Volume of stock solution added to $V = 0.8\,\mathrm{l}$ zero solution to produce several DOC concentrations.

| Resulting DOC concentration $c_{DOC}$ ($\mathrm{mg\,l^{-1}}$) | 0 | 5 | 15 | 25 | 50 | 75 | 100 |
|---|---|---|---|---|---|---|---|
| Added volume of stock solution $V_{add.}$ (ml) | 0 | 4.0 | 12.2 | 20.5 | 42.1 | 64.8 | 88.9 |

is the same measuring arrangement as in the present sensor design, the greatest comparability is given here. For each DOC concentration, five turbidity levels of 0, 20, 40, 60, 80 FAU are generated. To calculate the amount of talcum powder needed to produce the different turbidity levels, several concentrations of talcum powder and distilled water are produced and measured with a laboratory photometer (Spektralfotometer CADAS 50s from Hach Lange GmbH, Berlin, Germany). This results in a linear function of the talcum powder concentration and the resulting turbidity unit. Consequently, the following quantities of talcum powder are added to gradually increase turbidity from 0 to 80 FAU. Since different amounts of stock solution are added at each DOC concentration, the amount of talcum powder added varies depending on the total volume of the sample to be generated.

The measurements are performed by starting a Python script. In this study the measurement duration is $t_m = 10\,\mathrm{s}$. With a measurement interval of $t_i = 0.25\,\mathrm{s}$, 40 single point measurements are performed. For the IR channel, an interruption of $t_d = 30\,\mathrm{s}$ is set between the turbidity measurements to avoid drifting of the turbidity sensor due to heating.

The experimental procedure is as follows. First, a sample with the required DOC concentration is prepared and thoroughly mixed. Even if a magnetic stirrer is used, it is recommended to stir the sample occasionally by hand with a glass stirrer to keep the sample homogeneous. Shortly after stirring, a small amount of the sample is taken with a pipette and placed in a quartz cuvette. The sample is shaken thoroughly five times, the turbidity is measured with the laboratory photometer and the values averaged. The sample is then mixed again by hand in the glass vessel, the IR sensor is mounted inside and the measurement is carried out. Shortly after the IR sensor has been removed, the sample is thoroughly mixed, the UV sensor is placed inside and a measurement is performed. Then the sapphire glass window of the sensor head is cleaned, the next turbidity level is prepared and the procedure is repeated for all turbidity conditions. After measuring all turbidity conditions for one DOC concentration, the entire probe is cleaned, a new sample is prepared with the required DOC concentration and the process starts again. As a result, the voltage output from the IR and DOC sensors is measured for each combination of DOC concentration and turbidity. The results and measured values of the calibration are listed in the form of tables in the **Supplementary Material** of this article.

For IR and UV sensor evaluation, linear regression is performed using the least squares method and the coefficients of determination. For this purpose, the sensor signals are examined with regard to the correlation to the DOC concentration for all turbidity levels. Subsequently, a multiple linear regression according to the least squares method is performed by statistical software (IBM SPSS Statistics Version 21.0.0.0.0). Since the transmission has an exponential relationship, the extinction $E$ (cf. Lambert-Beer law) of the detector signal is used for the comparison according to Skrabal (2009). The dependent variable is the extinction of the UV sensor unit $E_{UV}$, the independent variables are the DOC concentration and the turbidity. To increase the linear relationship between dependent and independent variables, the calibration data set is revised. Therefore, all calibration points near the zero signal are eliminated, in this case all points with $E_{UV} \geq 0.164$. Only the revised data set is used for the evaluation due to the fact that all other calibration points are outside the linear range.

For the revised data set, a statistical, software-based linear regression analysis is performed using the least squares method. The resulting calibration curves are determined with respect to their coefficients of determination. Linear equations are calculated for the following relationships:

Linear relationship of the IR detector signal $E_{IR}$ as dependent variable and the turbidity values $c_{FAU}$ in FAU as independent variable. $m_1$ in $\mathrm{FAU^{-1}}$ and $n_1$ are constants. Therefore, all calibration points are included (see Equation 3).

$$E_{IR} = m_1 \cdot c_{FAU} + n_1 \tag{3}$$

Regarding the linear relationship of the UV detector signal $E_{UV}(c_{DOC})$ as dependent variable and the DOC concentration $c_{DOC}$ in $\mathrm{mg\,l^{-1}}$ as independent variable, only calibration points with turbidity $c_{FAU} = 0\,\mathrm{FAU}$ are included (see Equation 4). $m_2$ in $\mathrm{l\,mg^{-1}}$ and $n_2$ are constants.

$$E_{UV}(c_{DOC}) = m_2 \cdot c_{DOC} + n_2 \tag{4}$$

For the linear relationship of the UV detector signal $E_{UV}$ as dependent variable and the turbidity values $c_{FAU}$ in FAU as independent variable, only calibration points with $c_{DOC} = 0\,\mathrm{mg\,l^{-1}}$ are included (see Equation 5). $m_3$ in $\mathrm{FAU^{-1}}$ and $n_3$ are constants.

$$E_{UV}(c_{FAU}) = m_3 \cdot c_{FAU} + n_3 \tag{5}$$

### 2.3.2. Turbidity Compensation

In order to be able to distinguish between the extinction as a result of turbidity and DOC content in the field, it is necessary to determine a suitable turbidity compensation method.

A practical approach is to calculate additional calibration curves by linear regression using the least squares method. The linear relationship between the UV detector signal $E_{UV}(c_{DOC})$ as a dependent variable and the DOC concentration $c_{DOC}$ as an independent variable for each turbidity value of the calibration data set. In addition to Equation (4) there are further linear functions for $c_{FAU} = 20\,\mathrm{FAU}$, $c_{FAU} = 40\,\mathrm{FAU}$, and $c_{FAU} = 60\,\mathrm{FAU}$. Turbidity values greater than 80 FAU are not included in the revised calibration data set because they are outside the measurable range. After the turbidity value has been measured by the IR sensor unit, it is checked which calibrated turbidity

value is closest. The corresponding calibration curve is used to calculate the compensated DOC value. All negative DOC Values are set to zero. Using the calibration points, compensated DOC values can now be calculated. Two more turbidity compensation methods are described in the **Supplementary Material**. However, they provide qualitatively inferior results, which is why they will not be discussed further here.

## 2.4. Field Measurements

The overall objective of this work is to provide an instrument for service-oriented *in-situ* monitoring. For this purpose, the basic feasibility will be demonstrated by means of a case study on site. To this end, two exemplary monitoring campaigns were carried out in the summer of 2016 in the urban area of Leipzig, Saxony, Germany at Lake Cospuden (see **Figure 3A**) and along the Elstermühlgraben, Pleiße and Elster watercourses (see **Figure 3B**). Lake Cospuden is an artificial lake located south of Leipzig, Saxony, Germany. It is originated from a residual mining pit that was flooded. During re-cultivation, a recreational area with beach and landscape park was created around the lake. Due to the intensive use for local recreation, the ecological condition of the lake is also of interest. For this reason, two measurement campaigns were undertaken in summer 2016 by using a small boat to investigate the situation regarding the turbidity and DOC content.

Due to the service-oriented system architecture, the methodological effort during the field measurements is moderate. As described above, the sensor system consists of the sensor probe, which is connected to the topside unit via a cable and a connector.

The controlling unit can be accessed via a USB connection or via the power supply 12 V . . . 32 V (DC). In the present case, a USB Power Bank (TL-PB20100, Portable Power Station, 20 100 mA h. Capacity, TP-Link Technologies Co., Ltd., USA) with a capacity of 20 100 mA h was used. Thus, the measuring system can be operated for ∼48 h. The sensor probe can be easily mounted using the metal brackets integrated in the sensor housing at the top and bottom.

The user thus has the option of simply placing the sensor in the water on the cable, attaching it with a rope or attaching the sensor to a device carrier or frame. Conventional measuring systems are usually larger and heavier than the sensor system presented here. With the sensor system it is therefore possible to introduce new strategies and technologies for data acquisition in addition to classical monitoring through the use of research vessels.

As a proof of concept, two monitoring campaigns were carried out within the framework of the case study. For this purpose, a monitoring with a canoe was carried out, which enables data acquisition in very shallow river basins and small streams. The sensor installation for this case is shown in **Figure 4**. The sensor probe is attached to a bracket that allows the depth and orientation of the sensor to be changed from on-board.

After mounting the sensor probe, the system can be switched on without further user interaction. As long as a stable Internet connection exists, e.g., via an access point provided by a mobile phone, the data collected by the system is automatically transferred to the server and processed to a real-time data visualization via standardized web services. This is achieved with the help of the standardized data format JSON, which enables direct post-processing with the help of appropriate libraries and plug-ins.

The measurement system also works without an Internet connection and stores the data on an SD card. This prevents any loss of data as a result of a missing Internet connection.

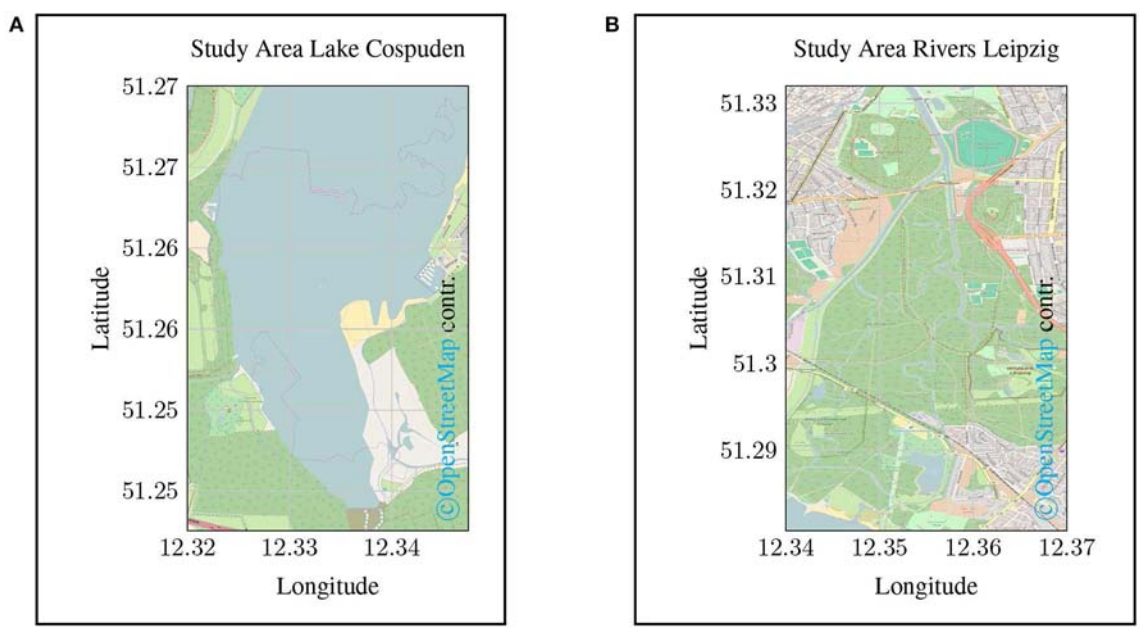



**FIGURE 3 |** Map of the study areas sampled in the case study. **(A)** shows the opencast mining lake Cospuden in the south of Leipzig. **(B)** shows the course of the Elstermühlgraben, a section of the Pleiße up to its confluence with the Elster near Leipzig.

# 3. RESULTS

## 3.1. Calibration of the Sensor

By calibrating the IR sensor unit, the signal was determined for all DOC concentrations and the corresponding turbidity values. For all measurement series, the coefficients of determination for the linear regression of the measurement signal by the IR sensor and the measured turbidity by the laboratory photometer are significantly higher than $R^2 = 0.9$. Consequently, there is a strong linear correlation between the decrease of the IR detector signal and the respective turbidity. Even with turbidity values greater than 80 FAU, with $U \approx 920$ mV the sensor unit is not yet in the detection limit range (zero level at $U \approx 21$ mV as shown in **Figure 5**).

To ensure that a changing DOC concentration has no effect on the IR detector signal, it is useful to perform correlation analysis for both channels. The results for the IR detector are shown in **Table 2** indicating that there is no correlation between the DOC concentration and the IR sensor signal.

The calibration of the UV sensor unit measures the signal for all DOC concentrations and several turbidity levels, as well as the IR sensor unit. For DOC concentrations at $c_{DOC} = 0\,mg\,l^{-1}$, $c_{DOC} = 5\,mg\,l^{-1}$, and $c_{DOC} = 75\,mg\,l^{-1}$, the scale of determination of the sensor signal and turbidity is relatively high at $R^2 \geq 0.9$. While the measured signals correlate with $R^2 \approx 0.8$ at DOC concentrations of $c_{DOC} = 15\,mg\,l^{-1}$, $c_{DOC} = 25\,mg\,l^{-1}$, and $c_{DOC} = 50\,mg\,l^{-1}$, the signal does not correlate

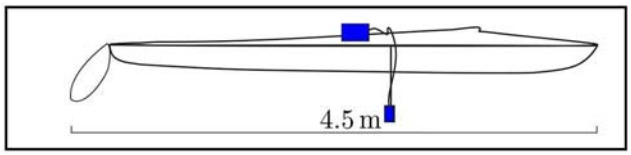



**FIGURE 4 |** Sensor installation on a canoe for mobile monitoring. The canoe is 4.50 m long. The sensor was positioned in the middle of the canoe at a depth of 0.35 m below the water surface. For user-friendly handling of the sensor, it is important that it can be installed quickly and easily. Due to the low weight, the installation can be adapted to the respective installation situation.

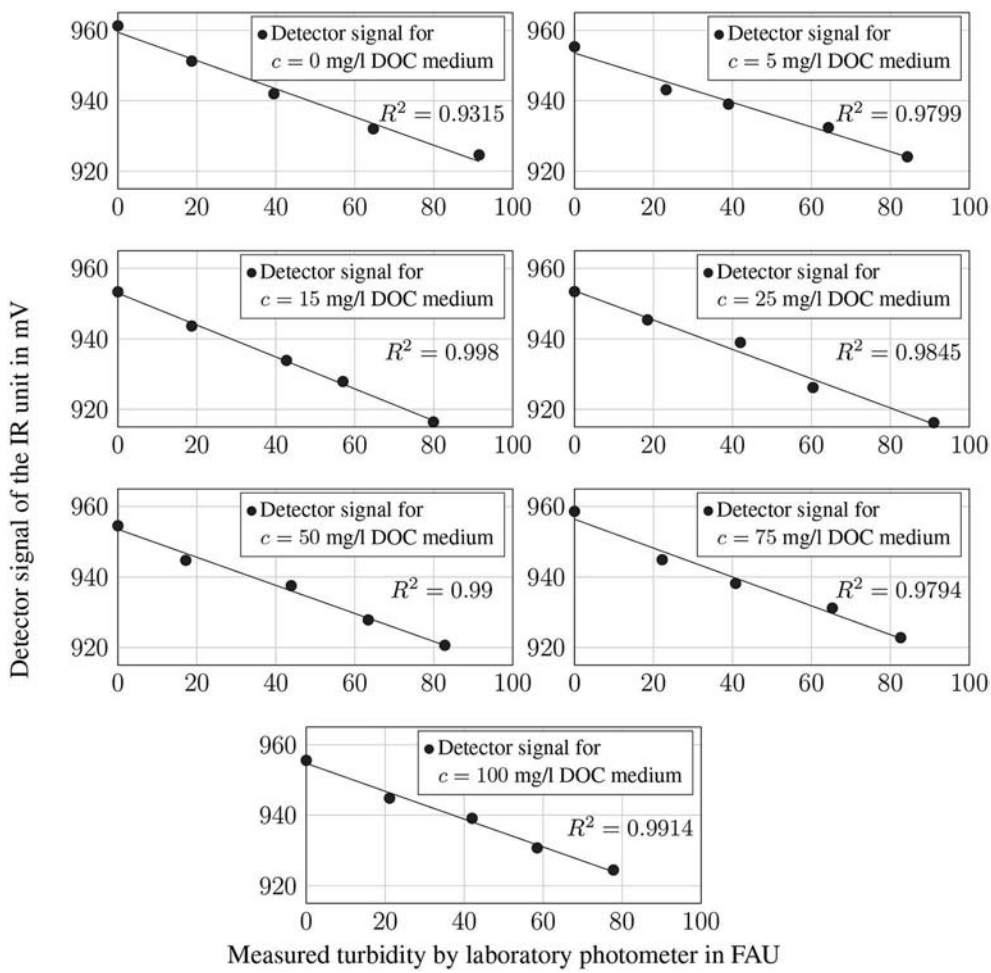



**FIGURE 5 |** Measured detector signal of the IR sensor in mV for all DOC concentrations and different turbidity values in comparison with the turbidity values measured using a laboratory photometer. The trend line was generated by linear regression.

with the turbidity level at $c_{DOC} = 100\,mg\,l^{-1}$. Consequently, there is a strong linear correlation between the sensor and the turbidity caused at low DOC concentrations, which decreases with increasing turbidity. In all cases, the measured signal shows a less pronounced linearity for values close to the zero level $U \approx 3.050$ mV (see **Figure 6**).

In contrast to the correlation between the IR sensor signal and the DOC concentration, the UV sensor signal correlates strongly with the DOC concentration. With correlation coefficients between −0.88 and −0.97 for the first four turbidity levels, the sensor shows a strong linear correlation with the DOC concentration at low turbidity values. Only for turbidity around 80 FAU the correlation coefficient decreases in magnitude so that the relationship between the sensor signal and the DOC concentration is less pronounced (see **Table 3**).

The extinction $E_{UV}$ estimated by the UV sensor unit, is to be compared in the following with the DOC concentration $c_{DOC}$ and the respective turbidity $c_{FAU}$. As shown in **Figure 7**, only

**TABLE 2 |** Correlation coefficients between the IR sensor signals and the different DOC concentrations for every turbidity value.

| $c_{DOC}$ in mg/l | Turbidity in FAU | Correlation coefficient |
|---|---|---|
| 0,5,15,25,50,75,100 | 0 | –0.047 |
| 0,5,15,25,50,75,100 | ≈ 20 | –0.275 |
| 0,5,15,25,50,75,100 | ≈ 40 | –0.063 |
| 0,5,15,25,50,75,100 | ≈ 60 | 0.010 |
| 0,5,15,25,50,75,100 | ≈ 80 | 0.242 |

**TABLE 3 |** Correlation coefficients between the UV sensor signals and the different DOC concentrations and the corresponding turbidity values.

| $c_{DOC}$ in mg/l | Turbidity in FAU | Correlation coefficient |
|---|---|---|
| 0,5,15,25,50,75,100 | 0 | –0.965 |
| 0,5,15,25,50,75,100 | ≈ 20 | –0.878 |
| 0,5,15,25,50,75,100 | ≈ 40 | –0.920 |
| 0,5,15,25,50,75,100 | ≈ 60 | –0.900 |
| 0,5,15,25,50,75,100 | ≈ 80 | –0.728 |

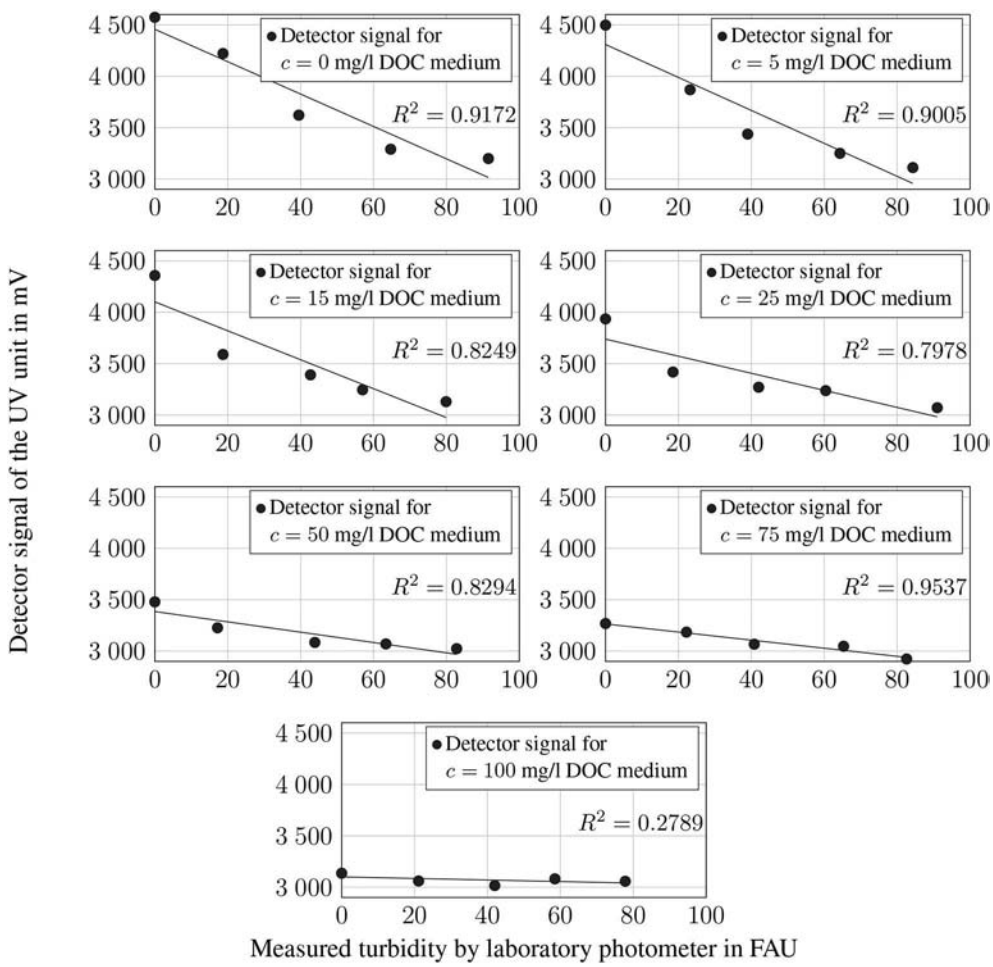



**FIGURE 6 |** Measured detector signals of the UV sensor in mV for all DOC concentrations and different turbidity values in comparison to the values determined by laboratory photometers. The trend line was generated by linear regression.

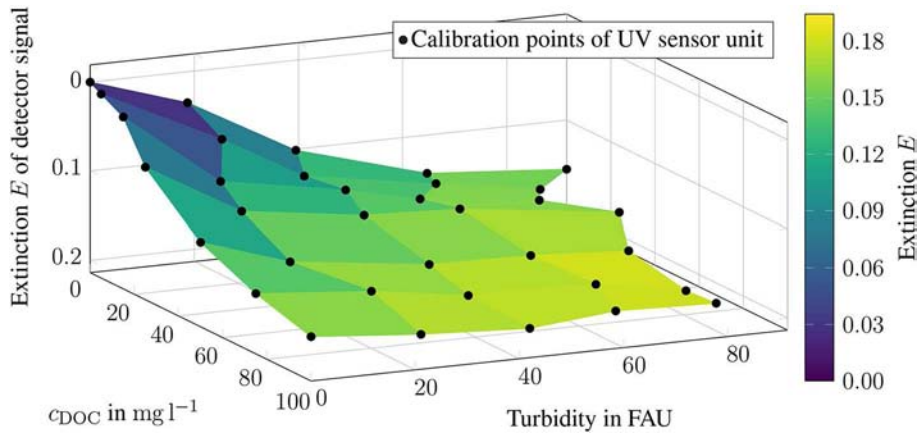



**FIGURE 7 |** Measured calibration points of the UV sensor represented as extinction values $E$, plotted on the respective DOC concentration and turbidity. The color yellow represents high detector signals, while dark blue represents low detector signals.

low extinctions $E_{UV}$ show a linear relationship. The higher the extinction $E_{UV}$, the flatter the curve. All calibration points in the blue area of the diagram are close to the zero signal. A linear relationship between the absorbance and the DOC concentration is only given for samples without turbid material. For turbidities of 20 FAU, only the $c_{DOC} \leq 80\,\mathrm{mg\,l^{-1}}$ range can be described as linear. If the turbidity value rises to 60 FAU, only $c_{DOC} \leq 25\,\mathrm{mg\,l^{-1}}$ can be assigned to the linear range. At 80 FAU the gradient becomes so flat that a linear relationship is no longer given (see **Figure 7**). Therefore, the adjusted calibration plot contains 22 of the original 35 calibration points. Compared to the unadjusted calibration data, $R^2$ has increased from 0.765 to 0.872. Thus, the linear relationship between $c_{DOC}$, $c_{FAU}$ and $E_{UV}$ has also increased (see **Figure 8**).

### 3.1.1. Turbidity Compensation

Based on the revised calibration points, a set of calibration curves can be determined. Equation (6) characterizes the linear relationship between the IR sensor signal and various turbidity values and has a measure of determination of $R^2 = 0.943$.

$$E_{IR} = 0.0029 + 0.00019 \cdot c_{FAU} \tag{6}$$

With $R^2 = 0.948$, Equation (7) shows the linear relation between the UV sensor signal and different DOC concentrations.

$$E_{UV}(c_{DOC}) = 0.0067 + 0.0018 \cdot c_{DOC} \tag{7}$$

The calibration curve (see Equation 8) shows the relation between the UV sensor signal and different turbidity values with $R^2 = 0.932$.

$$E_{UV}(c_{FAU}) = 0.0098 + 0.0018 \cdot c_{FAU} \tag{8}$$

However, despite the knowledge of turbidity, it is by far not trivial to deduce the correct content of dissolved organic material in water. For clarification, an approach for turbidity compensation is shown below.

For the different turbidity levels, the relationship between $E_{UV}$ and $c_{DOC}$ is shown in **Figure 9**.

The following calibration curves are calculated and converted into $c_{DOC}$. By inserting the UV detector signal value $E_{UV}$ into the equation with the corresponding turbidity, $c_{DOC}$ can be calculated directly. The differences between compensated and produced DOC values are greatest at $c_{DOC} = 15.25\,\mathrm{mg\,l^{-1}}$.

$$c_{DOC} = \frac{E_{UV}(0\,\mathrm{FAU}) - 0.0067}{0.0018} \tag{9}$$

$$c_{DOC} = \frac{E_{UV}(20\,\mathrm{FAU}) - 0.0666}{0.0015} \tag{10}$$

$$c_{DOC} = \frac{E_{UV}(40\,\mathrm{FAU}) - 0.1078}{0.0016} \tag{11}$$

$$c_{DOC} = \frac{E_{UV}(60\,\mathrm{FAU}) - 0.145}{0.0002} \tag{12}$$

The compensated measured DOC concentrations correlate closely with more than 90 % regarding the produced DOC concentration of the calibration media. Compared to the uncompensated values, the correlation increased by more than 45 % (see **Table 4**).

## 3.2. Field Measurements

The advantages of the proposed approach were demonstrated based on a case study. For better readability, the results of the measurement campaigns are presented separately. However, for the sake of completeness, all results and measured values of the monitoring campaigns are listed in the form of tables in the **Supplementary Material** of this article.

### 3.2.1. Monitoring Campaign Lake Cospuden, Leipzig

In **Figure 10** the raw values of the UV and IR channel in V are plotted as a map. The results of the first campaign carried out on July 2, 2016 are shown in **Figures 10A,C**. In **Figures 10B,D** the map view of the monitoring data of the second campaign carried out on July 22, 2016 are given.

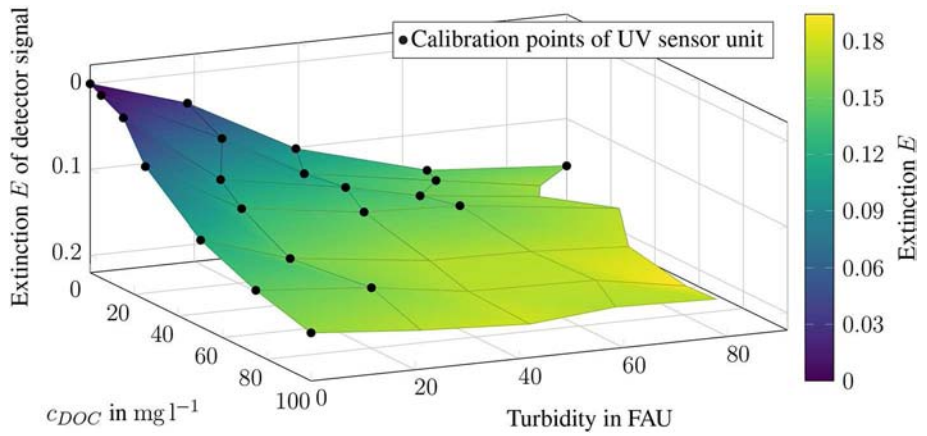



**FIGURE 8 |** Revised calibration points of the UV sensor unit represented as extinction *E*, plotted on the respective DOC concentration and corresponding turbidity. The color yellow represents high detector signals, while dark blue represents low detector signals.

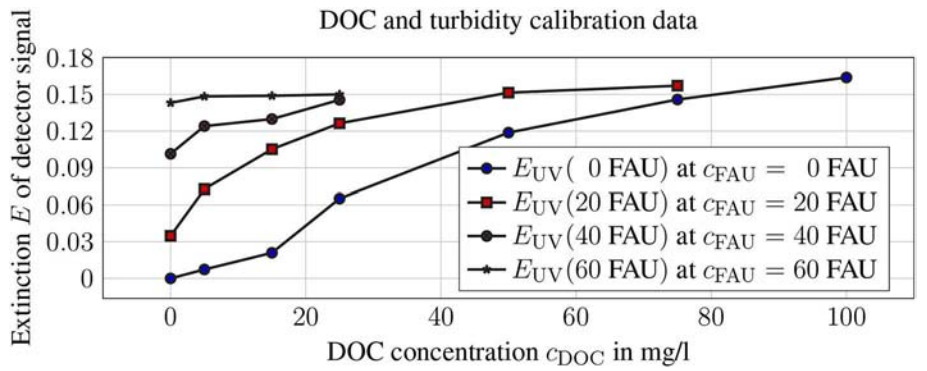



**FIGURE 9 |** Extinction $E_{UV}$ of the UV detector signal in relation to DOC concentration and the corresponding turbidity values.

**TABLE 4 |** Correlation coefficients using compensated and uncompensated values.

| Compensation | Correlation coefficient to produced DOC |
|---|---|
| Uncompensated | 0.461 |
| Compensated | 0.949 |

In addition to proving the field suitability of the developed measuring device, the field tests have also shown that direct feedback in the field, e.g., in the form of a web map display, is of great advantage. Thus, initial statements about distribution patterns can already be made during data collection. Following the idea of *Assissted Monitoring*, it was possible, for example, to sample areas that appear particularly interesting again or more intensively. Conventional monitoring strategies offer this possibility only to a limited extent.

Even though the data obtained represent the raw measured values of the IR and UV channels, some ecosystem relationships can still be identified as shown in **Figure 10**. Since no comparative samples were taken during the measurement campaigns for the laboratory analysis of water quality, only quantitative peculiarities will be discussed in the following. This is sufficient for the evaluation of the prototype and the proof of field suitability. With regard to the measurement results of the UV channel, it is noticeable that the values collected during the first measurement campaign (July 2, 2016) were significantly higher than in the second campaign on July 22, 2016 (see **Figures 10A,B**). With regard to the calibration tests previously carried out in the laboratory, it can therefore be deduced that the extinction as a result of an increased DOC concentration must have been considerably higher at the time of the second measurement campaign.

A similar situation applies to the IR channel. The results for the IR channel also show that there are clear differences in the water status with regard to turbidity (see **Figure 10D**). As shown in the previous laboratory test, the spectral attenuation in the UV range can also be influenced as a result of increased turbidity. It is not yet possible to determine conclusively from the initial measurements whether the measurement results are

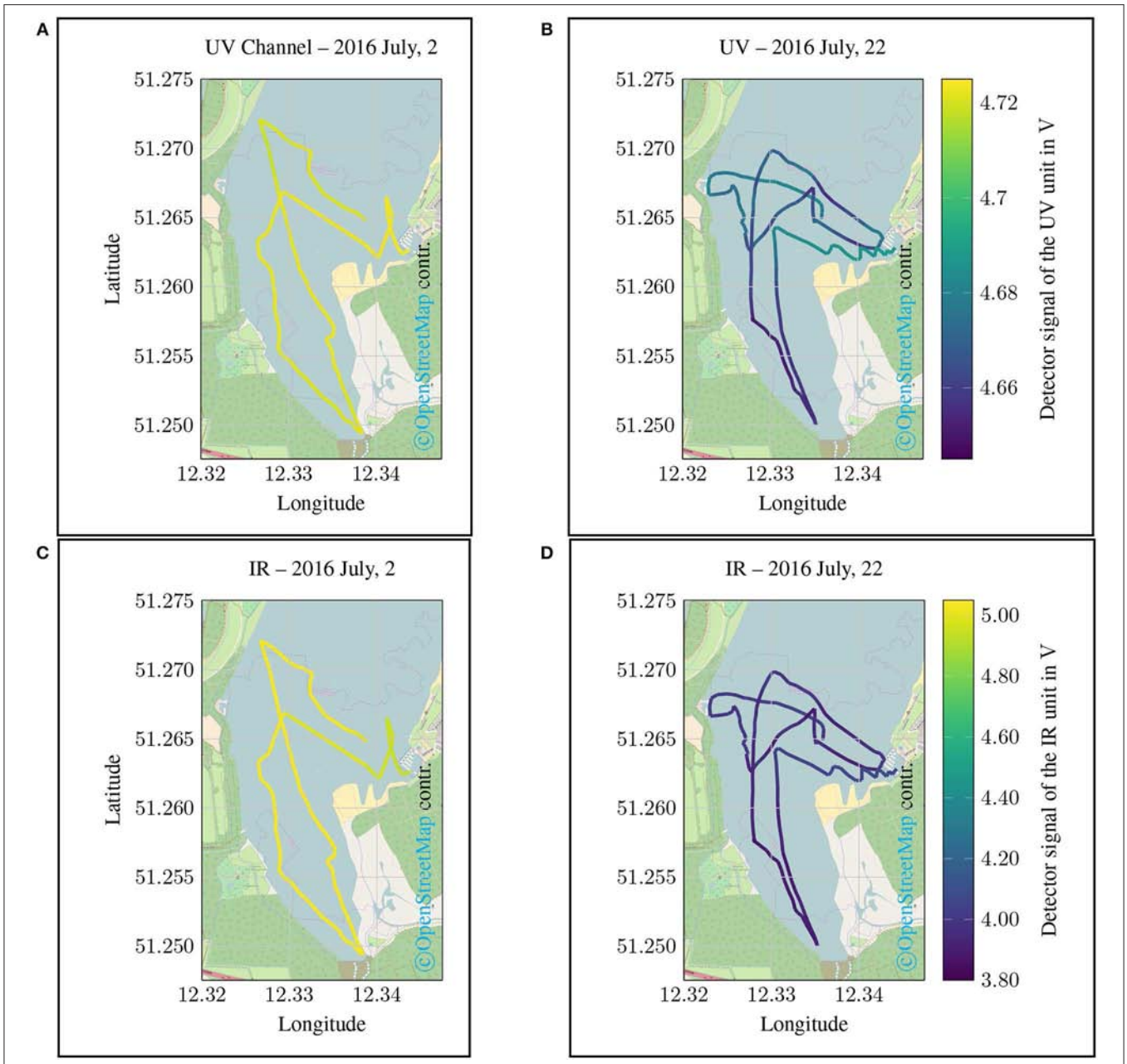



**FIGURE 10 |** Spatial plot of the UV detector Signal, which corresponds to the DOC content. The data were gathered using a small sailing boat at the Lake Cospuden close to the city of Leipzig (Lake Cospuden, Leipzig, Saxony, Germany). **(A)** and **(C)** are showing the results of the campaign carried out on 2 July, 2016. **(B)** and **(D)** are containing the results of the campaign carried out on 22 July, 2016. Shown are the pure measurements in V of the optical UV channel ($\lambda = 254\,\mathrm{nm}$). Low values in this representation mean a stronger light attenuation due to increased DOC content.

due to increased turbidity or an increased DOC concentration. However, supplementary reference measurements at selected points could be used to describe these correlations more precisely in further studies.

### 3.2.2. Monitoring Campaign Urban Streams, Leipzig

In a second campaign it should be investigated whether the developed measuring system can also be used in very small water systems since small bodies of water such as streams or ditches can only be measured with great effort using conventional methods. However, mobile monitoring approaches for small bodies of water pose an additional challenge. For this purpose, a canoe trip along the Floßgraben, the Pleiße and to the Elster in the urban area of Leipzig was carried out (2016 August, 19). These changed requirements for the monitoring task should show that the measurement system is suitable for a holistic monitoring approach. The results of the measurement campaign are shown in **Figure 11**.

For the this campaign, the results of the UV channel show a greater dynamic (see **Figure 11A**) than the values for turbidity (see **Figure 11B**). Due to the fact that the temperature is also recorded in the measuring system, statements can also be made regarding the thermodynamic properties of the water body. It should be noted that temperature differences in the size of ($\triangle\ \vartheta 4\,^{\circ}$C) occur particularly during the measuring campaign (see **Figure 12**).

## 4. DISCUSSION

The monitoring of DOC content by determining the spectral absorption coefficient at $\lambda = 254$ nm is a proven method. Even if there are several established probes on the market, an inexpensive *in-situ* probe for large area applications is not yet available. The prototype of the *in-situ* probe presented in this paper shows a general functionality and promising characteristics. The optical UV sensor unit for measuring DOC values and the IR sensor unit for measuring turbidity can be classified as suitable for photometric measurements. The hardware required for setting up the sensors is manageable thanks to the LEDs used. The costs are also lower compared to previously installed mercury vapor lamps. At the same time, energy efficiency has been increased and applicability improved by reducing the size.

There are some limitations with regard to the laboratory analytical evaluation and calibration of the sensor system. Since the relative standard deviation of the compensated values is higher than 100 %, it is not yet possible to measure precisely compensated DOC concentrations. However, since the correlation between the compensated measured DOC values and the real DOC values is significantly higher than the correlation between the uncompensated and the real DOC values, the turbidity compensation method has led to an improvement of the sensor readings. Possible causes for the high standard deviations lie in the calibration method or the electronic components of the probe used. In addition, formazine could be used as a turbid material because talcum powder solutions were difficult to homogenize. The accuracy of the compensated DOC values can be improved by adding additional turbidity values to the calibration. The range in which DOC concentration and absorbance are linear is very small for this prototype. A light source with a higher radiant power allows measurements of higher DOC concentration and turbidity values. Thus, at higher concentrations, the measurement signal reaches the saturated range and the linear range of the calibrated range becomes larger. By using a "UVLUX250-5" LED, the radiation power can be increased from 0.3 mW to more than 3 mW compared to the currently installed LED. Reducing the distance between the emitter and detector can also increase the maximum measurable concentrations. A further limitation can be found in the lack of verification by means of accompanying measurements during the mobile monitoring campaigns. This was not included in the basic feasibility study presented in this paper. For later studies, however, reference measurements should be made

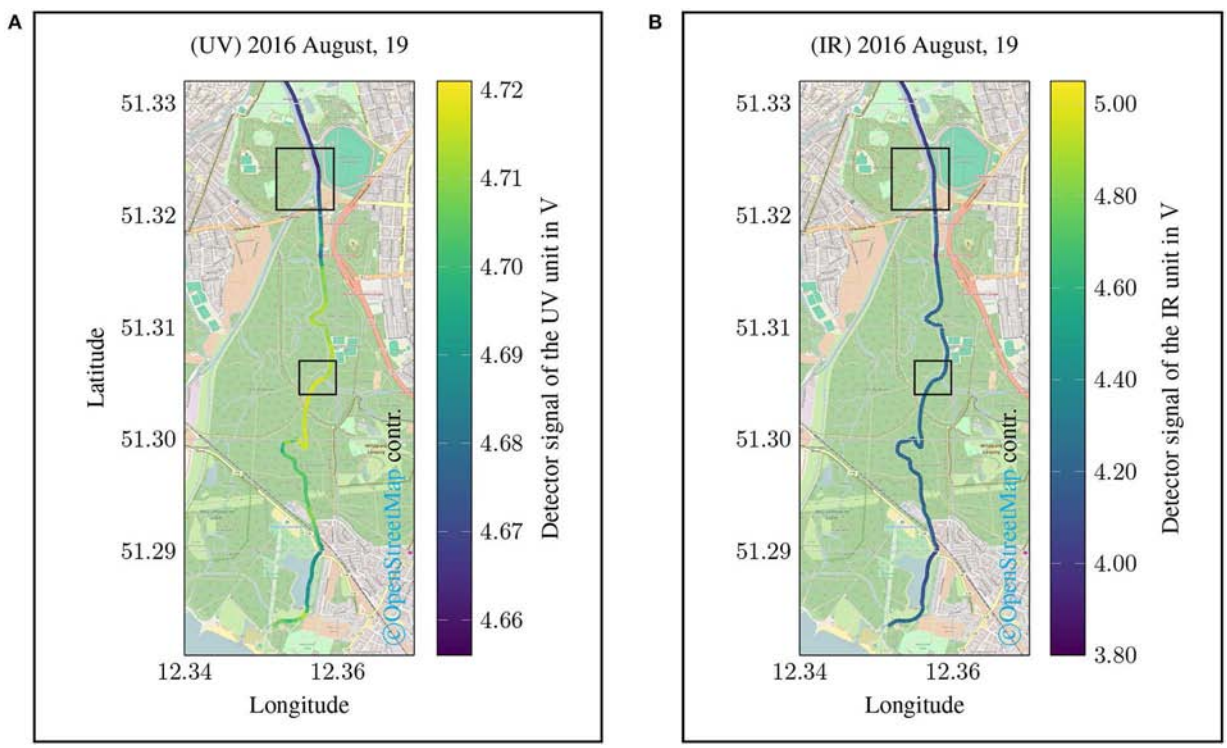



**FIGURE 11 |** Spatial data plot of a monitoring campaign along a river course (Floßgraben, Pleiße, Elster) in the city area of Leipzig, Saxony, Germany (2016 August, 19). **(A)** shows the detector signal of the UV channel corresponding to the DOC content. The results of the IR channel are shown in **(B)**. For both representations, lower values indicate a stronger attenuation due to an increased DOC content or increased turbidity.

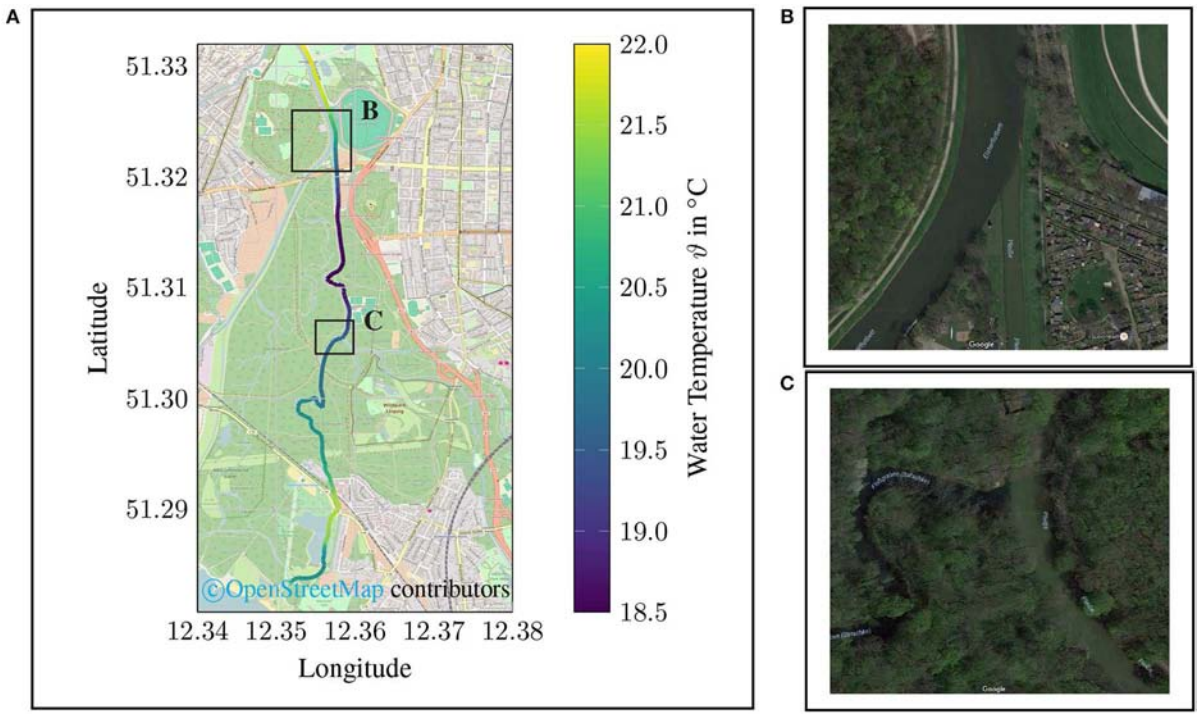



**FIGURE 12 |** Temperature data of mobile monitoring campaign along a river course (Elstermühlgraben, Pleiße, Elster) in the city area of Leipzig, Saxony, Germany (2016 August, 19). In **(A)** the spatial plot of the water temperature is given. Two squares are indicating points of special interest. **(B)** marks the estuary of the Elstermühlgraben into the Pleiße and **(C)** the estuary of the Pleiße into the Elster. (**B** retrieved from https://goo.gl/maps/qQyVoWTnV4J2; **C** retrieved from https://goo.gl/maps/QLphRa7QgGA2).

using established laboratory analytical methods and appropriate sampling, and the results and sensor performance should be evaluated accordingly.

The differences shown between the individual days (cf. **Figures 10A,B**) cannot be clearly justified in the context of the present study, as comparative measurements are also lacking here. Nevertheless, the prove was made that the presented measurement system is able to show spatial differences with little methodical effort and in near real time. Compared to conventional sampling and subsequent laboratory analysis, there is a considerable advantage in terms of time, method and therefore also in financial terms.

What has not yet been satisfactorily solved in the context of this study is the integration of sufficient temperature compensation to eliminate signal deviations due to temperature changes. Both the semiconductor elements used (emitter and detector) and the remaining electronics show temperature dependencies. This should be taken into account in subsequent studies.

In summary, the work shows that a measurement of DOC with a dual wavelength method at $\lambda = 254$ nm and $\lambda = 860$ nm can be realized by using LEDs, photodiodes and several low cost components. Additional experiments are necessary to improve the accuracy of the measurement. The use of more suitable components and an improvement of the calibration method are possible next steps. However, the open-source-based approach also offers the possibility of extending the range of functions and making additional optimization.

## 5. CONCLUSION

In addition to sensor and system development, the introduction of a holistic sampling theorem (OSE) is an essential part of this work (Schima et al., 2017). Through the combination of the monitoring paradigm and the consistent sensor development from the beginning, an unprecedented measurement system was achieved with regard to its functionality. Integrated functions for controlling the emitter intensity by changing the forward current or the pulse width modulation as well as the adjustable detector sensitivity by changing the integration time enable sophisticated functions such as autocalibration routines or adaptive system behavior. Therefore, this paper describes the structure of a modular and adaptive optical sensor concept, which opens the possibility of a fast and service-oriented environmental monitoring.

However, in order to show the feasibility of the sensor system, different initial field measurements campaigns were carried out in the area of the city of Leipzig, Saxony, Germany. A major advantage of the system is seen in its small size and low weight. This allows to carry out appropriate measurements even with small boats such as canoes or small sailing boats. Especially for shallow waters, where such measurements are often missing, this is an important addition to conventional monitoring strategies. In addition, a comparative study has shown that the sensor system presented delivers results that correspond to the laboratory analytical standard method with a laboratory photometer.

Thus, the measurement quality for the proposed application is considered acceptable.

The results confirm that the prototype of the sensor system represents a promising approach that meets the requirements and specifications outlined in the introduction to this work. Therefore, the sensor system provides a fast and useful method to support and improve future environmental monitoring applications, especially for the investigation of areas not yet accessible with conventional monitoring technologies.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

RS, MP, PD, and TG contributed conception and design of the study. RS was responsible for the design of the sensor system and the development of the electronics. The laboratory study was conceptually developed by RS. The laboratory work was carried out by SK. The results of the laboratory tests were evaluated by SK and RS. TG organized the database, the web service and the visualization. RS carried out the field tests and performed the statistical analysis. RS and TG wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## ACKNOWLEDGMENTS

We acknowledge financial support by Deutsche Forschungsgemeinschaft and Universität Rostock within the funding programme Open Access Publishing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00184/full#supplementary-material

## REFERENCES

Bass, M., Van Stryland, E., Williams, D. R., and Wolfe, W. L. (eds.). (1995). *Handbook of Optics, Volume I: Fundamentals, Techniques, and Design, 2 Edn*, Vol. 1. New York, NY: McGraw-Hill, Inc.

Brewin, R. J. W., Hyder, K., Andersson, A. J., Billson, O., Bresnahan, P. J., Brewin, T. G., et al. (2017). Expanding aquatic observations through recreation. *Front. Mar. Sci.* 4:351. doi: 10.3389/fmars.2017.00351

Chapman, D. (ed.) (1996). *Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring, 2 Edn.* London: E & FN Spon.

Clifford, N. J., Richards, K. S., Brown, R. A., and Lane, S. N. (1995). Laboratory and field assessment of an infrared turbidity probe and its response to particle size and variation in suspended sediment concentration. *Hydrol. Sci. J.* 40, 771–791. doi: 10.1080/02626669509491464

[Dataset] DIN 38404-3 (2005). *Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung - Physikalische und Physikalisch-Chemische Kenngrößen (Gruppe C) - Teil 3: Bestimmung der Absorption im Bereich der UV-Strahlung, Spektraler Absorptionskoeffizient (C 3).*

[Dataset] DIN EN 1484 (1997). *Wasseranalytik - Anleitungen zur Bestimmung des Gesamten Organischen Kohlenstoffs (TOC) und des Gelösten Organischen Kohlenstoffs (DOC).*

[Dataset] DIN EN ISO 7027 (2000). *Wasserbeschaffenheit - Bestimmung der Trübung.*

Guo, L., Santschi, P. H., and Warnken, K. W. (1995). Dynamics of dissolved organic carbon (DOC) in oceanic environments. *Limnol. Oceanogr.* 40, 1392–1403. doi: 10.4319/lo.1995.40.8.1392

Hoge, F. E., Vodacek, A., Swift, R. N., Yungel, J. K., and Blough, N. V. (1995). Inherent optical properties of the ocean: retrieval of the absorption coefficient of chromophoric dissolved organic matter from airborne laser spectral fluorescence measurements. *Appl. Opt.* 34, 7032–7038.

Huang, J., Liu, H., Tan, A., Xu, J., and Zhao, X. (1992). A dual-wavelength light-emitting diode based detector for flow-injection analysis process analysers. *Talanta* 39, 589–592. doi: 10.1016/0039-9140(92)80065-L

Hut, R., Tyler, S., and van Emmerik, T. (2016). Proof of concept: temperature-sensing waders for environmental sciences. *Geosci. Instrum. Methods Data Syst.* 5, 45–51. doi: 10.5194/gi-5-45-2016

Kolka, R., Weishampel, P., and Fröberg, M. (2008). "Measurement and importance of dissolved organic carbon," in *Field Measurements for Forest Carbon Monitoring*, ed C. M. Hoover (Dordrecht: Springer), 171–176.

Liu, H., and Dasgupta, P. K. (1994). Dual-wavelength photometry with light emitting diodes. Compensation of refractive index and turbidity effects in flow-injection analysis. *Anal. Chim. Acta* 289, 347–353. doi: 10.1016/0003-2670(94)90011-X

Lockridge, G., Dzwonkowski, B., Nelson, R., and Powers, S. (2016). Development of a low-cost arduino-based sonde for coastal applications. *Sensors* 16:528. doi: 10.3390/s16040528

Moore, C., Barnard, A., Fietzek, P., Lewis, M. R., Sosik, H. M., White, S., et al. (2009). Optical tools for ocean monitoring and research. *Ocean Sci.* 5, 661–684. doi: 10.5194/os-5-661-2009

Preisendorfer, R. W. (1976). *Hydrologic Optics. Volume 1. Introduction*. Technical Report, Honolulu, HI: U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Pacific Marine Environmental Laboratory.

Schima, R. (2018). *A fast method for mobile in-situ monitoring of optical properties in aquatic environments* (Unpublished thesis). University of Rostock, Rostock, Germany.

Schima, R., Goblirsch, T., Salbach, C., Franczyk, B., Aleithe, M., Bumberger, J., et al. (2017). "Research in progress: implementation of an integrated data model for an improved monitoring of environmental processes," in *Business Information Systems Workshops.* Lecture Notes in Business Information Processing (Cham: Springer), 332–339. doi: 10.1007/978-3-319-52464-1_30

Seibert, J., Strobl, B., Etter, S., Hummer, P., and van Meerveld, H. J. I. (2019). Virtual staff gauges for crowd-based stream level observations. *Front. Earth Sci.* 7:70. doi: 10.3389/feart.2019.00070

Skrabal, P. (2009). *Spektroskopie: Eine Methodenübergreifende Darstellung vom UV- bis zum NMR-Bereich* (Zürich: vdf Hochschulverlag AG).

Wiemann, S., Karrasch, P., and Bernard, L. (2018). *Ad-hoc* combination and analysis of heterogeneous and distributed spatial data for environmental monitoring - design and prototype of a web-based solution. *Int. J. Digital Earth* 11, 79–94. doi: 10.1080/17538947.2017.1326984

**Conflict of Interest Statement:** 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.

*Copyright © 2019 Schima, Krüger, Bumberger, Paschen, Dietrich and Goblirsch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

METHODS
published: 31 July 2019
doi: 10.3389/feart.2019.00195

# Particle Seeded Grains to Identify Highly Irregular Solid Boundaries and Simplify PIV Measurements

*William Basham, Ralph Budwig* and Daniele Tonina*

*Center for Ecohydraulics Research, University of Idaho, Boise, ID, United States*

Particle image velocimetry (PIV) is a non-invasive technique for measuring velocity fields. It is especially powerful when coupled with refractive index-matching (RIM) to map velocity fields around solid objects. The solid objects are typically removed from the flow field with a masking approach before performing the PIV analysis and mapping the velocity field, thus defined as an *a priori* method. However, applying this method, with a mask of the correct shape and at the correct location, is difficult, time consuming, and would be potentially unfeasible for packed bed of irregular shaped grains. To address this problem, we present the proof-of-concept of a novel approach to delineate highly irregular granular particles (grains) of varying size and shape and improve PIV processing for flows around grains in laboratory studies. The present technique makes use of seeding transparent RIM solids with light scattering particles during their fabrication. The RIM of the solids preserves the optical fidelity of images and the laser light sheet. Whereas the seeding in the solids can provide image contrast between solid (seeded) and fluid (non-seeded) as well as a strong zero-velocity signal in the solid. The fluid may then be seeded as well, allowing PIV spatial correlations to be performed with high confidence over the entire image. We tested the seeded RIM solid approach with both irregular individual solid pieces as well as with a volume of irregular grains. The new technique effectively obtains the fluid velocity field and solid boundary locations in both cases. Applications of the present method may range from studies of interstitial processes within a simulated sediment bed, such as those of aquifers, soils, sediments and the hyporheic zone, to near bed flow hydraulics.

**Keywords: fluid-solid boundaries, hydrogel, hyporheic flow, irregular granular particles, PIV, porous media flow, refractive index matching**

OPEN ACCESS

***Edited by:***
*Theresa Blume,*
*German Research Centre for Geosciences, Germany*

***Reviewed by:***
*Flavia Tauro,*
*Università degli Studi della Tuscia, Italy*
*Rolf Hut,*
*Delft University of Technology, Netherlands*
*Salvatore Manfreda,*
*University of Basilicata, Italy*

****Correspondence:***
*Ralph Budwig*
*rbudwig@uidaho.edu*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *31 October 2018*
***Accepted:*** *11 July 2019*
***Published:*** *31 July 2019*

***Citation:***
*Basham W, Budwig R and Tonina D (2019) Particle Seeded Grains to Identify Highly Irregular Solid Boundaries and Simplify PIV Measurements. Front. Earth Sci. 7:195. doi: 10.3389/feart.2019.00195*

## INTRODUCTION

Particle image velocimetry (PIV) is often conducted to study the motion of fluid around solid objects, whose presence, however, interferes with velocity spatial correlations that are performed to obtain the fluid velocity field. The typical approach applied to improve the spatial correlation with a solid object in the image is to impose an *a priori* digital mask over the regions where solid material is located and remove it (Gui et al., 2003; Adrian and Westerweel, 2011). Because this region is effectively removed from the flow field before quantifying the velocity distribution, it is very important that the location and extent of the solid is well known, and identifiable to

avoid distorting the real velocity field (thus, we defined it as *a priori*). This approach, known as masking, performs well when the solid objects have well defined shapes, e.g., a circular cylinder or a sphere and their locations are known. However, when objects have unidentified irregular shapes that occupy a large portion or potentially, as in the case of packed beds or sediments, most of the view field, this technique becomes less effective, and slower due to tedious point by point border identification and masking operations. The identification of the solid-liquid boundaries becomes extremely difficult, if not impractical, in a porous media of irregular and heterogeneous grains, such as sediments and soils. These limitations become even more restrictive when PIV is coupled with RIM because both fluid and solid objects may become indistinguishable. In the RIM method, the refractive index of two transparent materials are matched such that light is not refracted or reflected at the interface between the materials making the solid disappear in the fluid (Budwig, 1994; Wiederseiner et al., 2011; Wright et al., 2017).

For porous media with regular shaped and sized grains (usually spheres and cylinders), masking has been successfully applied for RIM-coupled PIV (e.g., Hassan and Dominguez-Ontiveros, 2008; Arthur et al., 2009; Satake et al., 2015; Harshani et al., 2017). In these laboratory studies, the authors imposed regular masking shapes onto the images by an *a priori* approach before conducting the PIV. Thus, the shape and location of the solids were assumed to be known and then the masking was applied. Alternatively, Dijksman et al. (2017) presented a method for finding the solid-liquid borders of spherical grains using variations in pixel light intensity from grain to liquid. They commented that "*Finding border voxels at the edge of the grains is a challenge.*" Their approach had several steps, one of which included a spherical analytical fitting of the border, which will not apply when grains are of irregular shape.

In this work, we describe an alternative approach to the *a priori* masking technique and that of Dijksman et al. (2017). The proposed method will both identify the boundary between the fluid and the solid surface and facilitate velocity vector determination with minimal interference. The concept of the proposed approach is to fabricate transparent (with the same refractive index as the fluid) solid objects seeded with light scattering particles. Seeded RIM solids have two key advantages. They may be distinct from the surrounding *unseeded* fluid when illuminated by the laser light sheet, which allows the identifying of their boundaries. The fluid is then seeded allowing PIV analysis of the entire field of view regardless of the presence of fluid and solids (no need to create a mask before performing PIV). In this way, the spatial correlation image analysis will obtain zero-velocity vectors in regions with seeded RIM solids and generally non-zero velocities in regions with moving fluid. The seeded RIM solid region has an additional important and key property: near zero (temporal) fluctuations of the velocity as the velocity does not change among succeeding images (because the seeded solid grains do not move). This last property allows differentiation between solids and potentially very slow moving fluids. Consequently, the seeded RIM solid method proposed here will allow finding both the solid surface outline and the velocity vectors simultaneously. It is particularly well suited to irregular shaped solid surfaces, such as grains, as well as to the complex pore flows in a packed bed of grains.

The present method was designed for physical modeling of flow through and around irregular shape grains in laboratory experiments. We demonstrated the method in a cm scale test cell, but it could be applied to larger physical models, e.g., flumes. PIV has been used for field studies of rivers as a means of mapping the surface velocity (Large Scale PIV known as LSPIV, see for example, Fujita et al., 1997; Muste et al., 2008; Tauro et al., 2016). To the author's knowledge, it has not been used for field studies within the water column of a river or stream, but it has been used within the water column of the ocean (Kakani and Dabiri, 2008; Kakani et al., 2017).

Seeded RIM solids have been used in a few previous studies for regular shaped solid pieces but not for other applications. Bellani et al. (2012) used a seeded spherical solid hydrogel piece in order to study its rotation and motion in a turbulent flow. They identified three points in the seeded solid to determine position and rotation, but did not discuss the use of the seeded solid to identify the solid-liquid boundary. Byron and Variano (2013) created a seeded ellipsoidal agarose piece to study its interactions with the fluid flow. They measured velocity of the seeded ellipsoidal piece but did not discuss the use of the seeded solid to identify the solid-liquid boundary.

In the present study, we test and apply the seeded RIM solid method in a set of experiments with two, single, irregular grains, and with irregular and a heterogeneous grain packed bed in a flow through the cell. The results demonstrate its potential to identify solid-liquid boundaries and velocity fields for these two limiting cases: a single irregular solid object and many grains of irregular shape and size.

## METHODS

Hydrogel with a refractive index that matches water (Weitzman et al., 2014) was selected as transparent RIM solid. The seeded RIM solid methods were tested in a 7 cm high, 5 by 5 cm square base flow cell, which was operated in two modes: (a) with two pieces of hydrogel grains, one seeded and one unseeded, and (b) with a packed bed of seeded hydrogel grains, which may simulate porous media, like soils and sediments. **Figure 1** shows a schematic diagram of the flow cell and a photograph of the flow cell filled with hydrogel grains and water. Hydrogel is visible in air as shown in the photograph where a few hydrogel grains were left on the top of the cell without water. However, submerged hydrogel grains are invisible because their refractive index matches that of water. Slabs of hydrogel were made following the recipe provided by Weitzman et al. (2014) with additional information from Menter (2016). Pieces were then broken from the slab for mode "a" experiments and they were irregular in shape similar to natural coarse sediment grains. The approximate width of these grains was 1 cm. For mode "b" experiments, the slab was pressed through a sieve with 8 mm openings into a sieve with 2 mm. As hydrogel was pressed through the sieves, it fractured into grains of irregular shapes, and sizes ranging between approximately 2 and 8 mm.

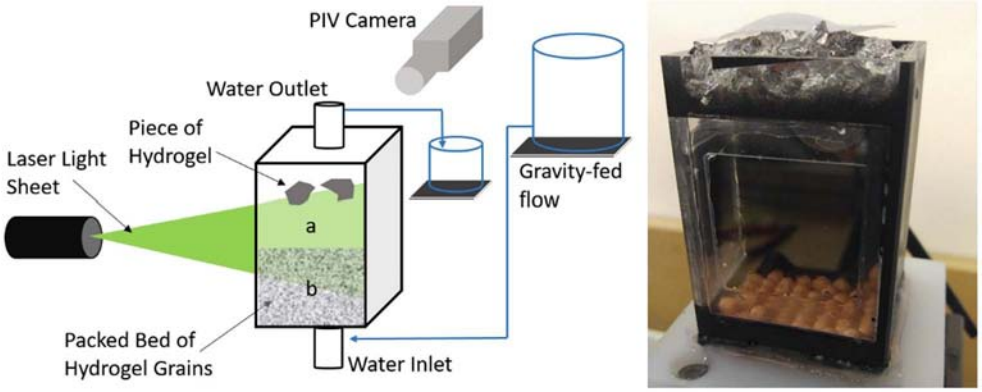



**FIGURE 1 |** Flow cell schematic (on the left) operated in two modes: mode "a" with pieces of hydrogel, and mode "b" with a packed bed of hydrogel. Note that each mode was tested separately with only one mode being in the test cell at a time. The flow cell is 5 cm by 5 cm footprint and 7 cm high. Photograph of the flow cell (on the right) with the lid removed and filled with hydrogel grains and water. Additional hydrogel grains were added on the top and are in air. Those in the air are visible but those fully submerged are invisible in water as their refractive index matches that of the water.

Only one hydrogel piece (the left piece in **Figure 2**) for experiments (a) and all the hydrogel grains for the packed bed of experiments (b) were seeded with 4 μm Nylon particles. The seeding was conducted during the hydrogel fabrication process by dispersing the Nylon particles into a portion of the water used to form the hydrogel. Particle dispersion into the water was facilitated by placing the beaker of water on the tray of a water filled ultrasonic cleaner. Our intent was to have the second hydrogel piece (the right piece in **Figure 2**) to be completely unseeded for experiment (a), so we could make a comparison between seeded and unseeded solids. Thus, the hydrogel piece on the right-hand-side of the flow cell was not seeded intentionally, but, nevertheless, contained a sparse seeding of background particles present in the de-ionized water used to make the hydrogel. Degassed reverse osmosis filtered water was delivered to the test cell from a head tank to create water flow over the hydrogel pieces or through the packed bed.

The PIV approach for this study used dual Nd:YAG lasers for light sheet production and a CCD camera with 1200 × 1600 pixels and a 180 mm focal length macro lens, which were required to obtain the field of view to capture the small flow passages among grains (interstitial flows). DaVis software was used for imaging and processing of the images. Image pairs were acquired at a rate of 3 pairs per second. Preprocessing of images included rotation and shift correction to diminish vibration effects as wells as subtraction of sliding average to reduce background noise. Image pairs were processed by a spatial correlation method down to a 12 × 12-pixel (0.11 mm × 0.11 mm) interrogation cell (IC) size. The only post processing conducted was outlier detection and removal.

The particle image velocimetry images and velocity results were then used to identify the fluid-solid boundaries by three methods, which may also be used in combination: (1) by contrast between seeded solid and unseeded fluid, (2) by applying a near zero-velocity threshold, and (3) by combining near zero-velocity with comparison between root mean square of the fluctuations of the in-plane velocity within the solid and the fluid.

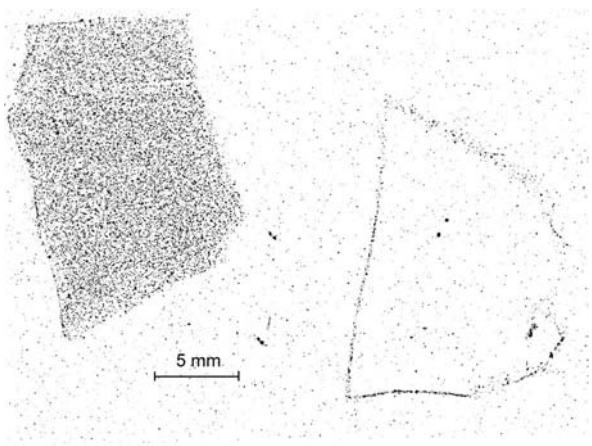



**FIGURE 2 |** Image of hydrogel pieces in test cell for mode "a" experiment. The image of the piece on the left illustrates the potential for determination of the solid-liquid boundary by method 1, i.e., contrast between the heavily seeded grain and the lightly seeded fluid around the grain.

## RESULTS AND DISCUSSION

**Figures 2**, **3** show the results for mode "a" experiments. **Figure 2** shows the hydrogel pieces, which were rigidly held in place from behind (by two sewing needles inserted into each grain) and illuminated by the laser light sheet in the test cell as shown in **Figure 1**. The hydrogel piece on the left was seeded, whereas the piece on the right had only a background level of particles as described in section "Materials and Methods." The water flowing around the pieces was not seeded but had background particles similar to the hydrogel piece on the right. **Figures 2**, **3** demonstrate that the solid material may be identified in the proposed three ways: (1) by contrast between seeded solid and unseeded fluid (**Figure 2**), (2) by applying a near zero-velocity threshold (**Figure 3**), and (3) by combining near-zero velocity

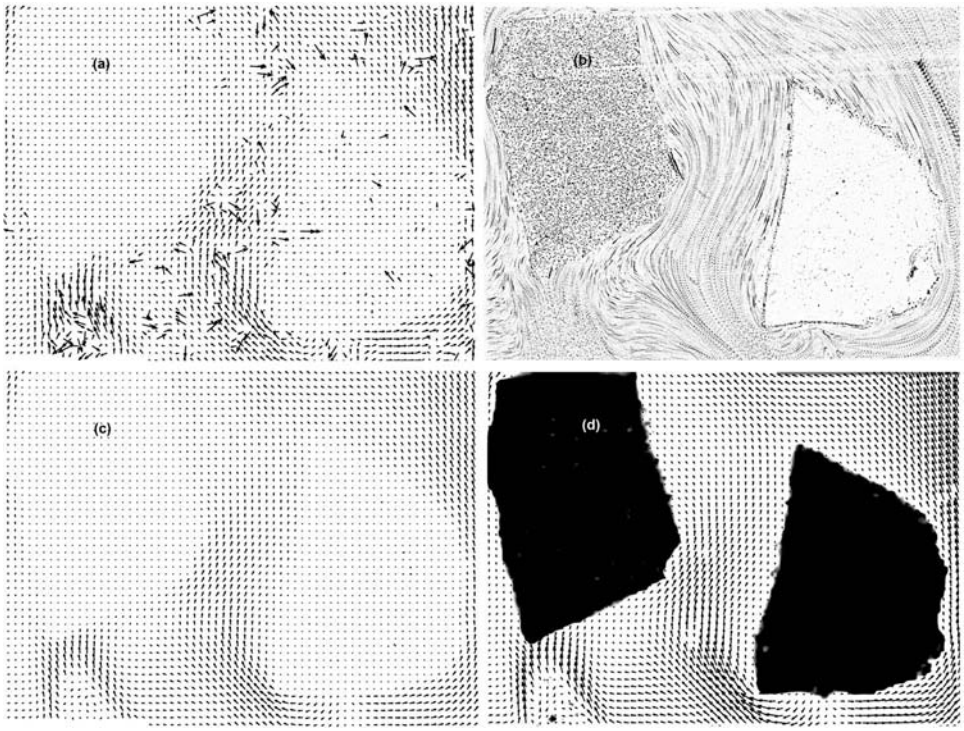



**FIGURE 3 |** Flow around hydrogel pieces in test cell for mode "a" experiment. **(a)** Instantaneous vector field. **(b)** Time series showing particle-pathlines. **(c)** Average over 100 instantaneous fields. **(d)** Applying near zero velocity threshold to reveal outline of pieces (method 2 for finding the solid-liquid boundary). A velocity vector with length of one grid spacing had a velocity of 4.6 mm/s. The approximate vertical velocity in the cell away from the grains was 5 mm/s. All vector plots show a vector for every other interrogation cell (IC).

with comparison between root mean square of the fluctuation of the in-plane velocity within the solid and the fluid. Whereas the first method does not require PIV but only light contrast between seeded solid and unseeded fluid, the last two methods require measurements of the velocity field and are based on the premises that (1) PIV predicted velocities within the solid are near-zero, and (2) because velocity in the solid does not change (stays near-zero) its temporal fluctuations are also near-zero.

For the first method, the left piece of hydrogel was heavily seeded and was made distinct when immersed in lightly seeded fluid (**Figure 2**). The right solid piece has the same level of seeding as the fluid and normally would not be distinct, but the deposition of background particles on the surface of the hydrogel has defined the surface-fluid boundary over most of the perimeter of the hydrogel piece. **Figure 2** also demonstrates that the laser light sheet illumination remained uniform even with the heavy particle seeding in the left piece of hydrogel.

For the second method, PIV was used to obtain the velocity field around the central plane of the hydrogel pieces including, instantaneous vector field (**Figure 3a**), time series showing particle-pathlines (**Figure 3b**), average over 100 instantaneous fields (**Figure 3c**), and applying near zero-velocity threshold to reveal the outline of pieces (**Figure 3d**). In addition, a video of the particle motion has been included in the **Supplementary Video S1**. The left solid piece in **Figure 3a** with dense seeding correctly shows zero-velocity vectors within the solid, while the right hydrogel piece with sparse seeding contains several spurious non-zero vectors, and which were not generated in seeded hydrogel because of the strong zero-velocity signal.

The outline of the pieces in the plane of the laser light sheet (i.e., the location of the fluid-solid boundary) may be determined by applying a near zero-velocity threshold to the PIV velocity results. **Figure 3d** shows the resulting piece outlines. The velocity threshold, $v_t$, used to locate the boundary was 0.1 mm/s. All interrogation cells with velocity less than 0.1 mm/s were classified as solid and set to a black background color. This method also worked for the grain on the right since it had enough background seeding to provide zero velocity levels. However, this approach erroneously identifies as solid a small location in the flow field near the left bottom of the figure, a small black spot, because of its very low in-plane velocity.

To better constrain the near zero-velocity threshold method, we tested the third method to improve the ability to distinguish between solid material, and regions of low in-plane velocity. This method additionally uses the root-mean-square (rms) level of the in-plane velocity fluctuations to discriminate between solid and fluid. In the present flow, solid material had low rms levels and fluid had high rms levels. This will be the case for most flows with some level of unsteadiness due to instability or turbulence. For example, the rms level in the heavily seeded solid piece divided by the rms level of the fluid at the small black spot near the left bottom of **Figure 3d** had a ratio of 1 to 25. Thus, based on

high rms level, the black spot at the left bottom of **Figure 3d** is actually fluid and it should be converted from solid (black color) to fluid (white color).

The velocity threshold, $v_t$, of 0.1 mm/s was identified from the velocity distribution of the single grain experiment, mode "a" (**Figure 4A**, vertical dotted line). Both the frequency distribution, FD, (solid line) and the cumulative frequency distribution, CFD, (dashed line) show two groups of velocities. We classified the first group as slow velocities belonging to the seeds within the solid. The rms values also show an increase for values of velocities larger than $v_t$ (**Figure 4C**), as expected because rms for the fluid velocities should be larger than those of the solid. A similar behavior is visible in the rms of the velocity for the packed bed experiment, mode "b" (**Figure 4D**), with rms increasing beyond the $v_t$ threshold. However, the velocity CDF is smooth and does not show the bi-modal characteristic observed in mode "a" (**Figure 4B**), because of the large fraction of solid in the system. The selected $v_t$ is small enough to be near-zero velocity but large enough to have most of the low velocity points (**Figure 4B**). Some low velocities are actually slow moving fluid particles approaching the solid boundary. These slow moving fluid locations can be identified by their large rms, with rms values larger than twice the minimum rms value quantified in the solid (green triangle marker points in **Figure 4D**).

Besides identifying the solid, the technique can help simultaneously quantifying and visualizing the flow field. The results of the time series particle pathlines reveal the flow patterns around the hydrogel pieces (**Figure 3B**). The flow at the top of the right piece appears to be emanating from the surface of the piece, which would violate the fluid, solid-surface boundary condition, and but it is an artifact of the three-dimensional characteristics of the flow. The reason for this apparently unphysical behavior is that the laser light sheet intersected the face of the hydrogel piece, where its face was strongly sloped upward, creating a significant vertical velocity component very near the surface. Nevertheless, the velocity went to zero at the hydrogel surface, as it must, to satisfy normal and tangential boundary conditions.

Regions with low magnitude in-plane velocity vectors were observed upstream of both pieces. The flow entered the cell through a small diameter inlet tube (**Figure 1**) without the aid of flow straighteners installed in the cell. This, along with the blocking effect of the two solid pieces, generated secondary flow patterns in the velocity field upstream of the pieces including out-of-plane velocities that were not resolved with the two dimensional PIV.

Flow velocities naturally approach zero velocity as the fluid gets closer to the boundary (the no-slip condition). These slow velocities may cause overestimates of the solid size. This effect

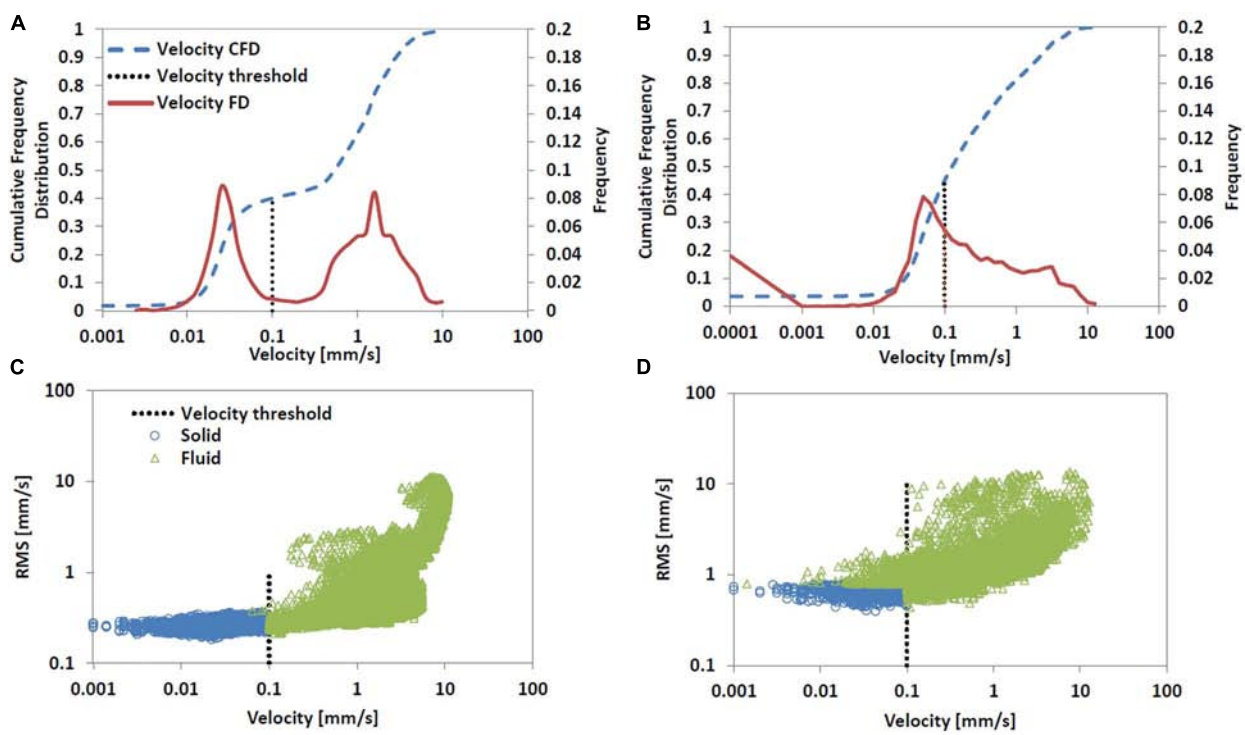



**FIGURE 4 |** Cumulative frequency distribution, CFD, (dashed lines), and frequency distribution (solid line) of the entire velocity field including solid and fluid are shown in the upper graphs. The root mean square velocity levels (symbols) are shown in the lower graphs with blue points as solid and green points as fluid. Individual grain, mode "a" experiments, left graphs **(A,C)**; and packed bed, mode "b" experiments, right graphs **(B,D)**. The vertical dotted line identifies the velocity threshold of 0.1 mm/s that separates solid from fluid points. For the packed bed shown in **(D)**, the threshold based on rms (twice the minimum rms, 0.39 mm/s) quantified several slow moving fluid locations that were below the velocity threshold. For the single grain shown in **(C)**, the threshold based on rms (twice the minimum rms, 0.18 mm/s) quantified very few slow moving fluid locations and only closely adjacent to the velocity threshold.

was shown by comparing the digitized boundaries of the seeded grain identified by method (1) (**Figure 5** solid red line) and by the combined method (2) and (3) (**Figure 5** dashed blue line). Because of the highly irregular shape of the grain, we used the contrast method (1) to provide the reference size to compare that from the combined methods (2) and (3), because we can visually see the grain. The combined methods (2) and (3) yield an increase in the grain perimeter of 1.1% and in planar area of 2.8%. Most of the error is located near the downstream side of the grain were very low velocities of low rms fluctuations were formed. However, the overall error is small.

The results for the mode "b" experiment with a seeded hydrogel packed bed are shown in **Figure 6**, which includes an instantaneous image of the seeding hydrogel grains infused with seeded flowing water (**Figure 6a**), a time series showing particle-pathlines (**Figure 6b**), and a plot with near-zero velocity threshold applied to reveal the outline of the hydrogel grains (**Figure 6c**). In addition, a video of the particle motion has been included in the **Supplementary Video S2**. The instantaneous image of particles shown in **Figure 6a** demonstrates that it is difficult to distinguish the location and extent of grains. We also took images of the seeded grains infused with reverse osmosis filtered water without added seeding. It was nevertheless difficult to distinguish the precise grain location and extent because of the presence of background particles in the filtered water. Thus, it would be impossible to apply the masking technique for this RIM packed bed study because the location and extent of the mask shape was unknown. The time series image shown in **Figure 4B** reveals stationary particles at grain locations and dark areas of particle-pathlines at locations where water was moving through the pore spaces. Pathlines moving around small solids or portions of large grains may appear to end in unconnected pores, which indicates a pathline has exited the plane, thus revealing a complex three-dimensional flow as water moves out of the plane through pores (Rubol et al., 2018).

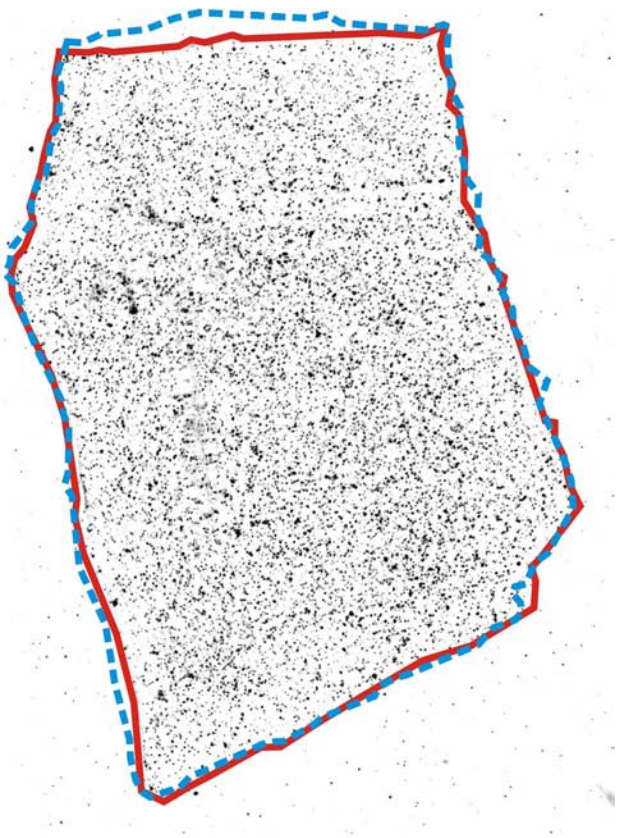


**FIGURE 5 |** Boundaries of the seeded particle in mode "a" experiment identified based on method (1) by contrast with the surrounding fluid without particles, red solid line, and based on methods (2) and (3), which had similar results, blue dashed line.

The in-plane pore water velocity vector plot (averaged over 100 instantaneous vector fields) is shown in **Figure 6c** along with blacked out interrogation cell's (IC's) at locations where the velocity was less than 0.1 mm/s. Velocity vectors in **Figure 6c** were plotted for every fourth IC and the scale was such that a velocity vector with length of one grid spacing had a velocity magnitude of 1.8 mm/s. The blacked out regions in **Figure 6c** reveal the extent of the grains in the plane of the laser light sheet. The image shown in **Figure 6c** is rich with details indicating the complexity of the pore spaces and multiple irregular grain shapes captured in the illuminated plane. The pore flow regions are in excellent agreement with the particle-pathline image shown in **Figure 6b**. The image in **Figure 6c** also reveals that pore spaces were small compared to regions of solid grain material (as it is typical because sediment porosity may range between 0.2 and 0.42), making it important to reduce the IC size to capture the details of the pore flow. The resolution of the present velocity threshold method used to determine the solid-liquid boundary location is set by the interrogation cell size. For the packed bed flow shown in **Figure 4**, the interrogation cell side length was 0.11 mm. We applied a smoothing function to interpolate velocity between neighboring interrogation cells.

We again tested the rms level method on the packed bed flow to better constrain the identification of solids. The method revealed that only one region was erroneously identified as solid but should have been slow moving fluid. This region was in the upper left area of the image and it is identified with a red arrow in **Figure 6c**. The ratio of the rms level of this region to the level within the seeded grains was 17 to 1, indicating that it was fluid rather than solid.

We were not able conduct a comparison of solid planar area between methods for mode "b" experiments, because the seeding level in fluid and solid were so close (see **Figure 6a**) such that it was impossible to distinguish between the two by method (1). Thus, the actual solid planar area of the irregular grains, as would be determined by method (1), was not available for comparison. Consequently, the mode "b" solid-liquid boundary results should be viewed as proof-of-concept. The primary method for identifying the solid-liquid boundary for a packed bed of irregular grains should be by contrast (method 1); since the flow over the packed bed of irregular grains is complex and this may affect threshold settings for methods (2) and (3).

**FIGURE 6 |** Flow through a packed bed of hydrogel grains. **(a)** Instantaneous image of particles. **(b)** Time series of particle motion. **(c)** Applying near zero-velocity threshold to reveal outline of grains (method 2 for finding the solid-liquid boundary). A velocity vector with length of one grid spacing had a velocity of 1.8 mm/s. The velocity in the interstitial spaces away from the boundary was 2 mm/s. The vector plot shows a vector for every fourth IC. Red arrows shows a solid identified by method (2), which instead should be a slow moving fluid when checked with rms value (method 3).

## SUMMARY AND CONCLUSION

In our test case, the seeded RIM grain was hydrogel and the fluid with matched refractive index was filtered pure water. The contrast between seeded grain and unseeded fluid was used to identify the solid-liquid boundary. The fluid was then seeded, allowing PIV analysis of the entire field of view regardless of the presence of fluid and solids (with no need to create a mask before performing PIV). In addition, the solid-liquid boundary was identified by considering locations with both near-zero velocity and low root mean square levels of the velocity fluctuations (i.e., low standard deviations of the velocity) in the field of view. Furthermore, the laser light sheet was not attenuated or distorted by seeding the RIM solid material and was able to uniformly illuminate the central plane of an entire packed bed of grains. We have also demonstrated that hydrogel may be fractured into grains for the study of packed bed flows. Other RIM liquid-solid pairs have the solid material with the potential for seeding as reported in Budwig (1994), Wiederseiner et al. (2011), and Wright et al. (2017). A likely candidate for solid seeding, in addition to the presently tested hydrogel, would be silicone rubber, which is paired with aqueous solution of sodium chloride and glycerol (Shuib et al., 2011). Others could be fluorinated ethylene propylene, FEP, or tetrafluoroethylene–hexafluoropropylene–vinylidene fluoride (THV), which, during three-dimensional (3D) printing, can be mixed with seeding. We believe that 3D printing, by both curing and melting techniques, will enable the proposed technique to be used with these and other solids. By curing method, the seeding can be mixed in the fluid and by melting process, the solid is shaped by applying the melted filament at micrometers thick layers, between which the seeding could be added (e.g., Guo et al., 2017).

The present seeded RIM solid approach is particularly attractive for irregular solid boundaries, like those found in packed beds. An *a priori* masking would be impossible for the present packed bed case and for any packed bed of irregular RIM grains. The present approach may be used to determine both the solid-liquid boundary as well as the velocity vector field. Additionally, by performing PIV in multiple planes across the test cell, the complete topography of the irregular grain packed bed as well as the pore flow map may be determined.

As in the work of Byron and Variano (2013), the method, in addition, could be used to identify the location of the solids and their motion. We also suggest, but we did not directly test it in this work, that the seeded RIM solid method is extremely powerful in studying moving mixtures of fluids and solid particles because the PIV spatial correlation analysis can be performed to the entire image regardless of where solids and fluid are. The mixture motion would be fully resolved. We suggest that the solid location and motion could be detected with pattern recognition analysis applied over several frames because the seeded RIM solid velocity field would follow that of a rigid body and would form a coherent structure. Thus, this physical modeling technique could be applied to other scientific fields beyond granular beds, such as flows within sediments and soils, and fluidized beds of mixture of fluid and particles. Other fields may include sediment transport, particle mobility analysis, and flow field around rough boundary like streambeds.

The ability to use irregular shaped grains of different sizes from fractions of a millimeter to centimeter sizes is significant, because these grains may be fabricated to mimic natural soil particles that are highly heterogeneous in sizes and shapes. Whereas masking is an effective technique when the shape and location of grains are known, it is not possible when grains have a range of unknown shapes and sizes. In the present study, we created and studied a packed bed of irregular grains, but we did not attempt to mimic the actual shape and size distributions of a real sediment bed, though there is the potential for this in future studies. Consequently, this technique provides an effective method for studying natural porous media flows, because it allows both the identification of the shape and size of the grains and the quantification of the flow field around them. This technique will pave the way to explore the interstitial processes within a simulated sediment bed, such as those of aquifers, soils, sediments, and the hyporheic zone (Tonina, 2012).

## AUTHOR CONTRIBUTIONS

WB, RB, and DT designed the experiments. WB and RB conducted the experiments and analyzed the data. RB and DT wrote the manuscript. WB reviewed the manuscript. All authors discussed the results.

## FUNDING

This research was funded by the National Science Foundation award number EAR1559348. Any opinions, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the supporting agency.

## ACKNOWLEDGMENTS

We thank the three reviewers and the Associate Editor Theresa Plume for their valuable and insightful comments, which improve our contribution.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00195/full#supplementary-material

Two video files have been included showing seeding particle motion in a plane of the test cell, (i) over the two irregular hydrogel grains (mode "a" experiment), and (ii) through the packed bed of hydrogel grains (mode "b" experiment). The original PIV image frames have been placed in the Hydroshare archive (Basham et al., 2019).

## REFERENCES


Adrian, R., and Westerweel, J. (2011). *Particle Image Velocimetry*. New York, NJ: Cambridge, 404–405.

Arthur, J., Ruth, D., and Tachie, M. (2009). PIV measurements of flow through a model porous medium with varying boundary conditions. *J. Fluid Mech.* 629, 343–374. doi: 10.1017/s0022112009006405

Basham, W., Budwig, R., and Tonina, D. (2019). Data for particle seeded grains to identify highly irregular solid boundaries and simplify PIV measurements. *HydroShare.* doi: 10.4211/hs.1f0242ee442945ca87de348ad2dd4315

Bellani, G., Byron, M., Collignon, A., Meyer, C., and Variano, E. (2012). Shape effects on turbulent modulation by large nearly neutrally buoyant particles. *J. Fluid Mech.* 712, 41–60. doi: 10.1017/jfm.2012.393

Budwig, R. (1994). Refractive index matching methods for liquid flow investigations. *Exp. Fluids* 17, 350–355. doi: 10.1007/bf01874416

Byron, M., and Variano, E. (2013). Refractive-index-matched hydrogel materials for measuring flow-structure interactions. *Exp. Fluids* 54, 1456–1461.

Dijksman, J., Brodu, N., and Behringer, R. (2017). Refractive index matched scanning and detection of soft particles. *Rev. Sci. Instrum.* 88:051807. doi: 10.1063/1.4983047

Fujita, I., Muste, M., and Kruger, A. (1997). Large-scale particle image velocimetry for flow analysis in hydraulic engineering applications. *J. Hydraul. Res.* 36, 397–414. doi: 10.1080/00221689809498626

Gui, L., Wereley, S., and Kim, Y. (2003). Advances and applications of the digital mask technique in particle image velocimetry experiments. *Meas. Sci. Technol.* 14, 1820–1828. doi: 10.1088/0957-0233/14/10/312

Guo, T., Lembong, J., Zhang, L. G., and Fisher, J. P. (2017). Three-dimensional printing articular cartilage: recapitulating the complexity of native tissue. *Tissue Eng. Part B* 23, 225–236. doi: 10.1089/ten.TEB.2016.0316

Harshani, H., Galindo-Torres, S., Scheuermann, A., and Muhlhaus, H. (2017). Experimental study of porous media flow using hydro-gel beads and LED based PIV. *Meas. Sci. Technol.* 28:015902. doi: 10.1088/1361-6501/28/1/015902

Hassan, Y., and Dominguez-Ontiveros, E. (2008). Flow visualization in a pebble bed reactor experiment using PIV and refractive index matching techniques. *Nuclear Eng. Des.* 238, 3080–3085. doi: 10.1016/j.nucengdes.2008.01.027

Kakani, K., and Dabiri, J. (2008). In situ field measurements of aquatic animal-fluid interactions using a self-contained underwater velocimetry apparatus (SCUVA). *Limnol. Oceanogr. Methods* 6, 162–171. doi: 10.4319/lom.2008.6.162

Kakani, K., Sherlock, R., Sherman, A., and Robison, B. (2017). New technology reveals the role of giant larvaceans in oceanic carbon cycling. *Sci. Adv.* 3:e1602374. doi: 10.1126/sciadv.1602374

Menter, P. (2016). *Acrylimide Polymerization – a Practical Approach. Bio-Rad Laboratories, Tech Note 1156*. Available at: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_1156.pdf (accessed October 24, 2018).

Muste, M., Fujita, I., and Hauet, A. (2008). Large-scale particle image velocimetry for measurements in the riverine environment. *Water Resour. Res.* 44:W00D19. doi: 10.1029/2008WR006950

Rubol, S., Tonina, D., Vincent, L., Sohm, J., Basham, W., Budwig, R., et al. (2018). Seeing through porous media: an experimental study for unveiling interstitial flows. *Hydrol. Process.* 32, 402–407. doi: 10.1002/hyp.11425

Satake, S., Aoyagi, Y., Unno, N., Yuki, K., Seki, Y., and Enoeda, M. (2015). Three-dimensional flow measurement of a water flow in a sphere-packed pipe by digital holographic PTV. *Fusion Eng. Des.* 9, 1864–1867. doi: 10.1016/j.fusengdes.2015.03.023

Shuib, A., Hoskins, P., and Easson, W. (2011). Experimental investigation of particle distribution in a flow through a stenosed artery. *J. Mech. Sci. Technol.* 25, 357–364. doi: 10.1007/s12206-010-1232-4

Tauro, F., Porfiri, M., and Grimaldi, S. (2016). Surface flow measurements from drones. *J. Hydrol.* 540, 240–245. doi: 10.1098/rsif.2014.1283

Tonina, D. (2012). Surface water and streambed sediment interaction: the hyporheic exchange, in *Fluid Mechanics of Environmental Interfaces*, (eds.) Gualtieri, C. *and Mihailoviæ, D.T.* London: CRC Press, 255–294. doi: 10.1201/b13079-13

Weitzman, J., Samuel, L., Craig, A., Zeller, R., Monismith, S., and Koseff, J. (2014). On the use of refractive-index-matched hydrogel for fluid velocity measurement within and around geometrically complex solid obstructions. *Exp. Fluids* 55, 1862–1873.

Wiederseiner, S., Andreini, N., Epely-Chauvin, G., and Ancey, C. (2011). Refractive-index and density matching in concentrated particle suspensions: a review. *Exp. Fluids* 50, 1183–1206. doi: 10.1007/s00348-010-0996-8

Wright, S., Zadrazil, I., and Markides, C. (2017). A review of solid–fluid selection options for optical based measurements in single phase liquid, two phase liquid–liquid and multiphase solid–liquid flows. *Exp. Fluids* 58, 108. doi: 10.1007/s00348-017-2386-y


**Conflict of Interest Statement:** 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.

*Copyright © 2019 Basham, Budwig and Tonina. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

MINI REVIEW
published: 11 September 2019
doi: 10.3389/feart.2019.00221

# Low-Cost Environmental Sensor Networks: Recent Advances and Future Directions

*Feng Mao*, Kieran Khamis, Stefan Krause, Julian Clark and David M. Hannah*

*School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, United Kingdom*

The use of low-cost sensor networks (LCSNs) is becoming increasingly popular in the environmental sciences and the unprecedented monitoring data generated enable research across a wide spectrum of disciplines and applications. However, in particular, non-technical challenges still hinder the broader development and application of LCSNs. This paper reviews the development of LCSNs over the last 15 years, highlighting trends and future opportunities for a diverse range of environmental applications. We found air quality, meteorological and water-related networks were particularly well represented with few studies focusing on sensor networks for ecological systems. Furthermore, we identified bias toward studies that have direct links to human health, safety and livelihoods. These studies were more likely to involve downstream data analytics, visualizations, and multi-stakeholder participation through citizen science initiatives. However, there was a paucity of studies that considered sustainability factors for the development and implementation of LCSNs. Existing LCSNs are largely focused on detecting and mitigating events which have a direct impact on humans such as flooding, air pollution or geo-hazards, while these applications are important there is a need for future development of LCSNs for monitoring ecosystem structure and function. Our findings highlight three distinct opportunities for future research to unleash the full potential of LCSNs: (1) improvement of links between data collection and downstream activities; (2) the potential to broaden the scope of application systems and fields; and (3) to better integrate stakeholder engagement and sustainable operation to enable longer and greater societal impacts.

**Keywords: sensor network, low-cost, environment, monitoring, internet of thing, information and communication technology**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology,*
*Netherlands*

***Reviewed by:***
*Lutz Breuer,*
*University of Giessen, Germany*
*Peter M. Marchetto,*
*University of Minnesota Twin Cities,*
*United States*
*Chet Udell,*
*Oregon State University,*
*United States*

****Correspondence:***
*Feng Mao*
*f.mao@bham.ac.uk*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *19 March 2019*
***Accepted:*** *13 August 2019*
***Published:*** *11 September 2019*

***Citation:***
*Mao F, Khamis K, Krause S, Clark J and Hannah DM (2019) Low-Cost Environmental Sensor Networks: Recent Advances and Future Directions. Front. Earth Sci. 7:221. doi: 10.3389/feart.2019.00221*

## INTRODUCTION

In recent years there has been a marked increase in the use of low-cost sensor networks (LCSNs) in the environmental sciences to address both pure research questions and applied management issues (Benedetti et al., 2010; Ojha et al., 2015; Prasad, 2015). As sensor networks with low-cost components in the setup, the rise of LCSNs has been driven by a number of factors including: the reduced cost of microcontrollers, communication modules and environmental sensors (Fisher et al., 2015), and the open science movement, which has seen the research community readily sharing designs, underlying software and firmware and data (Pearce, 2013). While there are some trade-offs with regards to robustness, calibration requirements and accuracy of low cost sensors

when compared to high–end commercial sensors (Castell et al., 2017), the potential for greatly increased spatial coverage will facilitate new insights into environmental process dynamics (Krause et al., 2015). In addition to the low-cost, a key advantage of open source electronics and "DIY" sensor networks is that end-users can fully customize the network applications with potential to employ adaptive monitoring or real-time feedback and control (Blaen et al., 2016). This also enables specific monitoring or research requirements to be achieved in a number of contexts, such as smart earth and smart agriculture (Hart and Martinez, 2006; Ojha et al., 2015; Bakker and Ritts, 2018).

The technical aspects of low-cost sensor network design and application are now relatively well understood, thanks to the rapid development of information and communication technologies. However, recent research suggests that remaining challenges are largely focused around non-technical factors such as stakeholder engagements, socio-economic contexts, financial and operational mechanisms (Mao et al., 2018). These non-technical issues have already started to hinder the potential benefits these sensor networks can provide society. For example, the potential for risk reduction, resilience building, and adaptive management are frequently overlooked (Paul et al., 2018). These points are salient given the potential of low-cost sensor networks to address the inadequate data coverage in low- and mid-income countries (e.g., Strigaro et al., 2019), particularly as this lack of information remains a major challenge for policy makers in these regions (UN, 2015). Hence, there is an urgent need to better understand these emerging challenges and identify possible opportunities for future research.

Given the above, this study aims to quantitatively and systematically review and synthesize the contemporary literature on environmental LCSNs, in order to analyze current research foci and identify knowledge gaps. Reviewed publications are assessed in three non-technical dimensions that are believed to be critical for successful implementation of low-cost sensor networks and maximize their societal benefits (Mao et al., 2018) – first, clear workflow from data collection to data processing and provision (Paul et al., 2018); second, consideration of stakeholder groups (e.g., end-users and operators) in designing, using or managing sensor networks; and third, sustainable and adaptive setup of the sensor network. In doing so we sought to address four specific hypotheses, namely that: ($H_1$) studies using LCSNs have a bias toward fields that have a *in situ* sensor monitoring tradition, such as meteorology; ($H_2$) the predominate focus has been on data collection, with limited effort dedicated to other downstream data activities such as data visualization, analytics or real-time control; ($H_3$) most studies focus on technically orientated single end-users (i.e., scientists), without considering the high potential for multi-stakeholder participation; and, ($H_4$) given $H_3$, the focus in the field has been on the technical feasibility of sensors and networks and the importance of factors such as sustainable operating mechanisms and physical and socio-economic contexts have been neglected.

# METHODS

The use of systematic review procedures to identify the state of the art in a given research field is becoming increasingly popular in the physical (e.g., Bartesaghi Koc et al., 2018), medical (e.g., Hill et al., 2016) and social sciences (e.g., Karpouzoglou et al., 2016). This approach facilitates a rigorous appraisal and synthesis of the literature in a (semi)-quantitative way to address specific hypotheses or research questions (Mulrow, 1987). Furthermore, the analytical tools and search engines now available enable large databases of academic literature to be searched and results categorized in short amounts of time (Xu and Marinova, 2013). However, search criteria must be carefully selected to ensure the pool of literature used is suited to the hypotheses or questions posed. Here we adopt the approach outlined by Pickering and Byrne (2014) which attempts to identify geographical patterns, theoretical trends, and methodological gaps rather than undertake statistical analysis of evidence as is common in the meta-analyses of the medical sciences.

To identify the body of literature for quantitative review we used the Web of Science database, which is the largest online database for searching peer reviewed scientific literature and the most academically orientated of the main search engines (Xu and Marinova, 2013). Our aim was to include papers from two general themes: (1) low-cost environmental sensors networks, and (2) low-cost technologies that have direct relevance to (low-cost environmental) sensor networks. To achieve this, we used the following search criteria:

$$\text{TS} = [(\text{“sensor network}*\text{”}) \text{ AND } (\text{“low-cost” OR “lowcost” OR “opensource” OR “open source” OR “inexpensive”})] \quad (1)$$

where TS represents topic searching title, abstract and keywords that returned the initial pool of papers for consideration ($n$ = 4593). The literature was then filtered to include only papers that were deemed explicitly related to environmental monitoring. To achieve this, we only included papers from Web of Science categories that were related to the geographical, environmental or earth sciences (see **Supplementary Table S1** for list of categories). This step returned 218 articles from 153 journals and conferences proceedings.

These papers were then assessed in turn by examining the abstract or full manuscript to extract: (i) general information (publication year, country and study type); (ii) information on the environmental system studied ($H_1$; i.e., Atmosphere, Hydrosphere, Earth, etc.); (iii) sensor mobility and data transmission/processing level ($H_2$); (iv) user groups ($H_3$); and, (v) sustainability considerations ($H_4$). In order to consider how the existing studies utilize sensor networks, we also checked if the publications were: (1) focused on an environmental application of the technologies described; (2) describing a sensor network rather than a single sensor; or (3) focused on the measurement and collection of environmental data rather than the performance of the network per say. There were 135 publications meeting all the three criteria. More detailed information on this procedure can be found in **Supplementary Table S2**. All

data collation and analysis was conducted using R version 3.5.1 and the *Tidyverse* ecosystem of packages outlined in Wickham and Grolemund (2016).

## RESULTS AND DISCUSSION

The concept of low-cost environmental sensor networks appears to have first emerged in the literature in 2004. Since this date there has been a steady increase in the number of publications per year, with the highest numbers (32 and 33) recorded in 2017 and 2018, respectively (**Figure 1A**). This result was expected as the increase in published studies tracks the growing trend toward open science and the rise of the "makers movement" within the wider scientific community (Baden et al., 2015). Interestingly, 2004 roughly coincides with the development and release of the *Arduino* board an inexpensive, consumer orientated microcontroller board[1] and the increase in publications post 2012 also coincides with the release of the low cost, single board computer, *Raspberry Pi*[2].

[1]https://www.arduino.cc/

[2]https://www.raspberrypi.org/

The global distribution of the analyzed studies displayed a distinct bias toward developed countries (particularly North America and Western Europe) with no studies from Africa and only a limited number from other developing regions (**Figure 1B**). This is concerning as, for example, the low number of African hydrological or meteorological monitoring stations hamper policy development and environmental management (van de Giesen et al., 2014). However, there are some projects underway such as the TAHMO project[3] which aims to install 20,000 low-cost weather stations across Africa.

Most studies were single case studies with few review or conceptual articles captured by our literature search (**Figure 1C**). This disparity is likely to represent the relatively recent development of LCSNs as tools for environmental monitoring applications, particularly those used in peer reviewed scientific studies. The review papers were either focused on more general technological advances in environmental monitoring and not focused solely on low-cost networks (e.g., Rossiter, 2018), or provided a user group perspective on low cost sensor networks (e.g., citizen science; cf. Rai et al., 2017; Paul et al., 2018).

[3]https://tahmo.org

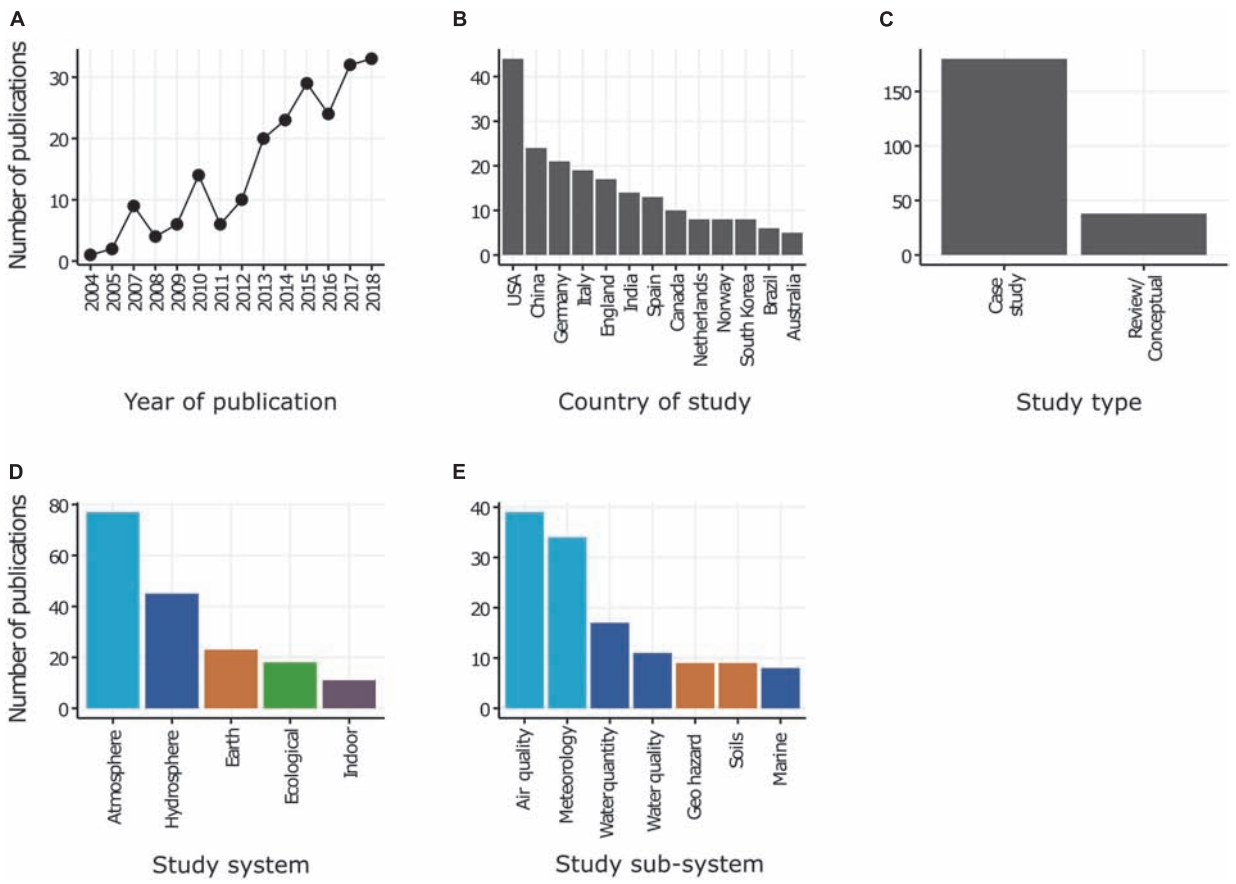



**FIGURE 1 |** Number of publications as a function of: **(A)** year of publication, 2004–2018; **(B)** country in which the study was conducted; **(C)** the type of study or paper; **(D)** broad study system(s), and; **(E)** sub-categories identified within the air, water and earth study systems. In panel **(B)** only countries with >5 studies are displayed.

When considering the study system at a relatively coarse scale the literature appeared to support $H_1$ (i.e., there was a bias toward fields with a history of *in situ* monitoring), with 77 publications focused on applications in the lower atmosphere and 19 on ecological systems (**Figure 1D**). Given the long history of sensor use for *in situ* atmospheric monitoring, particularly for meteorological variables, and limited use of sensors for monitoring ecological systems these results may not be surprising (Hart and Martinez, 2006). However, a larger number of the atmosphere system studies were focused on air pollution ($n$ = 39), rather than displaying a bias toward meteorology as anticipated ($n$ = 34) (**Figure 1E**). This was unexpected given the historical reliance on passive sampling and expensive laboratory equipment for analysis in air quality studies (Snyder et al., 2013). It appears public awareness of health risks (Ali et al., 2015; van Zoest et al., 2018), and the proliferation of low-cost *in situ* sensors (Schneider et al., 2017; Munir et al., 2019) are driving this trend. For water systems more studies focused on quantity ($n$ = 17) as opposed to quality ($n$ = 11). The water quantity studies were predominately focused on flooding (e.g., Horita et al., 2015; Acosta-Coll et al., 2018; Bartos et al., 2018), but studies on water resource management (e.g., Katsiri and Makropoulos, 2016) and the interface between agriculture and water resource monitoring were apparent (e.g., Kim et al., 2011; López et al., 2015). The water quality studies represented a mixture of pollution monitoring networks (e.g., Schneider et al., 2016) and agriculture focused applications (e.g., López Riquelme et al., 2009). Studies on earth systems were evenly distributed between those with a geo-hazard focus, such as landslides and earthquakes (Pumo et al., 2016; Finazzi and Fassò, 2017) and those with a focus on soil properties (e.g., Shaw et al., 2016). A further category was identified that represented studies focused on communication protocols or network architecture. An interesting trend was identified with many of these studies being pre 2012 (e.g., Bengston and Dunbabin, 2007; Walter, 2010), suggesting the field is moving beyond some of the technical aspects of wireless communication protocols and hardware with the focus now on data quality, interpretation and analysis.

When considering how existing studies collect environmental data and how they are utilized (e.g., analysis, decision-making and system control), some distinct patterns are apparent. Firstly, most sensor networks used fixed point sensors and data were transferred wirelessly either to a base station, remote server or the cloud (**Figures 2A,B**). The use of mobile sensors is more common for ecological systems, particularly tracking animal movement (e.g., Davis et al., 2012) and for monitoring air quality (Mead et al., 2013); however, Schneider et al. (2016) outlined a study in which sensors fitted to rafts or kayaks were used to continuously gather water quality data while moving downstream. Wired sensor networks or systems that required direct data download from local storage were associated with either: (1) scientific studies in which networks were maintained to answer a specific research objective or test a new senor type (Barnard et al., 2014; Pohl et al., 2014); or (2) monitoring networks for human infrastructure in urban environments where Ethernet connections were available (Dauwe et al., 2014; Rettig et al., 2014).

Secondly, there was a slight bias with regards to how the data were used with more papers ($n$ = 76) reporting just data collection and storage than with a data analysis component ($n$ = 59) (**Figure 2C**). This result appears to support $H_2$ (i.e., predominate focus is on data collection), however, there appears to be a growing trend toward the development of online analytics and visualization with 23.8% of all pre 2012 studies and 46.6% of those post 2011. Most storage-only-networks were used in scientific studies with analysis conducted offline by researchers. For example, Pohl et al. (2014) used a network of low-cost weather stations to collect information on snow depth at a high spatial-resolution to quantify the influence of landscape factors on snow accumulation. Monitoring networks with online analytics and visualization were more common in recent studies where some degree of human safety or health was related to the sensed parameters. Examples include geo-hazards (Finazzi and Fassò, 2017), flooding (Jones et al., 2015; Acosta-Coll et al., 2018; Bartos et al., 2018) or air quality (Schneider et al., 2017; Kizel et al., 2018). A further type of monitoring network with analytics and visualization was associated with agriculture (Kubicek et al., 2013) and in several studies this was advanced toward automated control of nutrient addition/irrigation to optimize resource use and yield (López et al., 2015; Srbinovska et al., 2015).

Thirdly, the majority of studies (82.9%) involved networks that collected data and were isolated in nature (i.e., not part of a wider dataset or larger network) (**Figure 2D**). These data were then only available to or used by direct stakeholders, for example technicians/scientists (e.g., Pohl et al., 2014) or farmers involved in crop production (e.g., López et al., 2015). More recent studies have collected data to complement existing monitoring efforts (i.e., or have been operating as a sub network within a larger national network). These were in many cases associated with human health (Rogulski, 2018) or climate impacts (Shusterman et al., 2018; Šećerov et al., 2019) or had direct economic implications, for example through flooding (Horita et al., 2015) or fishing livelihoods (Wada et al., 2007). It should be noted that very few studies embraced the principles of open science and open data more generally (however see Rettig et al., 2014; Jones et al., 2015).

Despite stakeholder engagement, especially citizen science, being one of the most significant "innovative" approaches associated with LCSNs there was a paucity of such studies identified in the literature (**Figure 2F**). Given the relatively small number of studies with multiple stakeholders (21.2%) there appears to be strong support for $H_3$ (i.e., most studies focus on technically orientated single end-users). However, there are some interesting examples of multiple stakeholder participation (e.g., Ali et al., 2015; Finazzi and Fassò, 2017). The involvement of citizen scientists can improve the functionalities and impacts of low-cost sensor networks by supporting its operation, enhancing adaptation, information provisioning and resilience building (Horita et al., 2015; Paul et al., 2018). In return, some sensor network applications have tailored

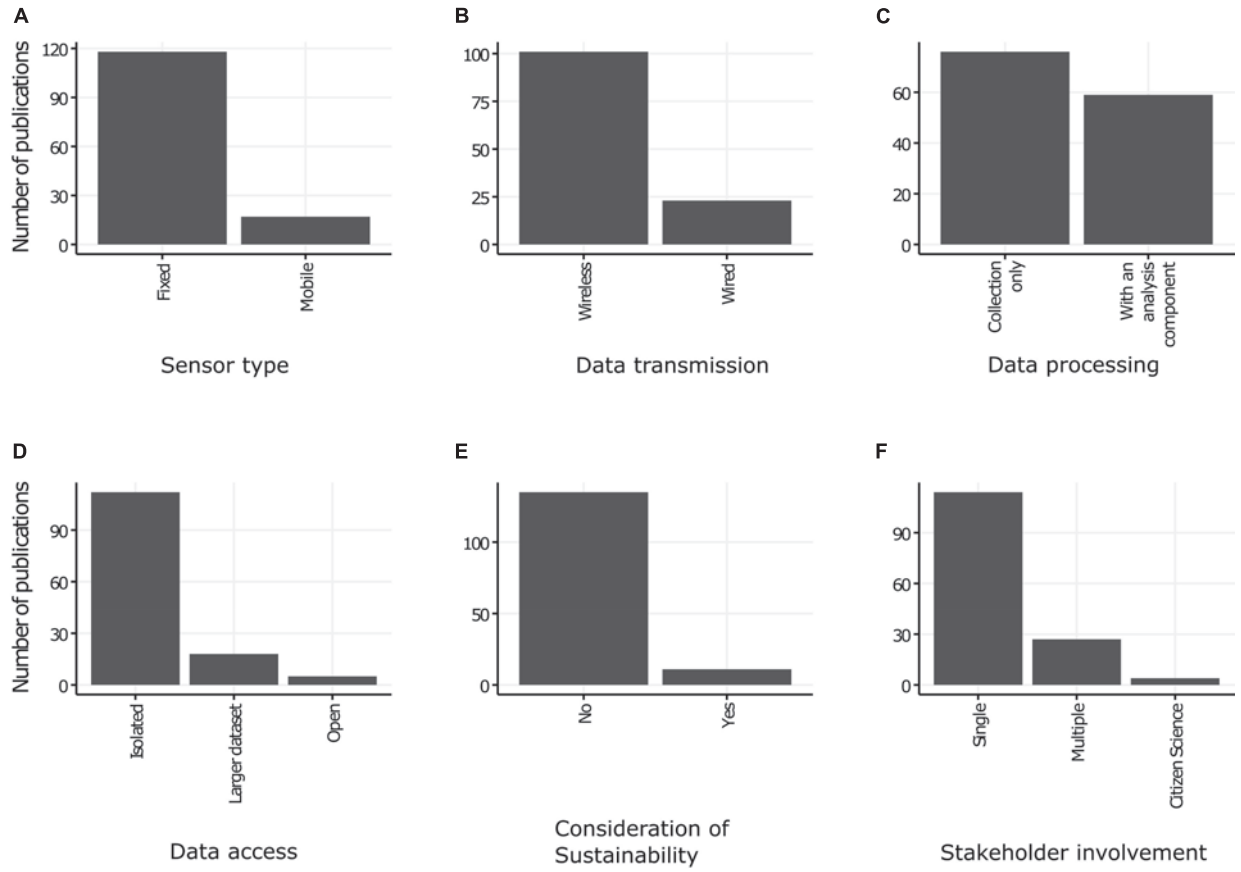



**FIGURE 2 |** The distributions of publications based on sensor and data collection properties. **(A)** Publication number split by sensor/node deployment properties. **(B)** The breakdown between wireless transmission of data to a base station or cloud-based server vs. hard-wired connection including Ethernet and local storage with manual data download. **(C)** The number of studies which highlighted networks for collection and data storage only, as opposed to studies with a data analysis component after collection, including visualization, analytics and control. **(D)** Data access, here the category "Isolated" refers to a dataset only available internally or to direct stakeholders, "Open" refers to an open access data storage platform, and "Larger dataset" refers to a study that contributed to a larger dataset. **(E)** The number of publications that explicitly considered sustainability principles. **(F)** The degree of stakeholder and end-user engagement in the study.

designs to improve the user experience of citizen scientists (Schneider et al., 2016).

The application of low-cost sensor networks has been highlighted as a key area that can transform environmental governance, yet long-term environmental governance requires sustainable and long-term operations of low-cost sensor networks (Bakker and Ritts, 2018; Paul et al., 2018). Despite this, most studies identified in this review do not explicitly consider sustainability (92.5%; **Figure 2E**) and thus support of $H_4$ is strong (i.e., sustainable operating mechanisms and physical and socio-economic contexts have been neglected). One possible explanation for this could be that most studies are from developed regions with significant resources and infrastructure (cf. **Figure 1B**). Sustainability can be achieved through either technical improvements via means such as optimization of energy efficiency (Gleonec et al., 2017; Mazinani and Davarzani, 2017), or innovative soft management/incentive-based approaches (Bakker and Ritts, 2018). Most of the reviewed studies identified with a sustainability element were associated with explicit and direct human benefits, such as monitoring a particular resource (e.g., Wada et al., 2007), protecting property or livelihoods (e.g., Lopes Pereira et al., 2014) or were agricultural in nature and focused on resource use to maximize yields (e.g., Geipel et al., 2015).

## CONCLUSION AND FUTURE OPPORTUNITIES

To summarize, LCSNs are increasing in popularity but there is still a distinct bias toward developed countries, particularly Western Europe and North America, and certain study systems (e.g., atmosphere and hydrosphere). From this systematic literature review, three key challenges and opportunities were identified, which can also guide future technical development of LCSNs. Firstly, data outputs from LCSNs need to be processed and presented to benefit multiple stakeholders including scientists, the general public and policy makers. While there is still a paucity of examples with studies exploring down-stream data activities such as analysis, decision-making

and system control examples exist for certain study purposes (e.g., geo-hazards) from which lessons can be learned for other purposes. Secondly, there is a clear need to improve data integration and sharing. This will involve a move away from isolated datasets to closer alignment with existing monitoring systems to create larger, richer datasets and a concerted effort to make data more open. While this has begun the idea needs to be at the core of future networks to improve system understanding and avoid duplication of effort. Thirdly, the design of LCSNs needs to better integrate stakeholder engagement and sustainable operation to enable longer term and greater societal impacts and environmental benefits.

## AUTHOR CONTRIBUTIONS

FM conceived of the presented idea and designed the research framework with support from KK, JC, SK, and DH. FM and KK reviewed the literature and drafted the manuscript. KK led the data analysis and interpreted the results together with FM. SK and DH provided critical feedback and constructive comments. All authors were involved in the discussion of the results.

## FUNDING

This work was supported by United Kingdom Natural Environment Research Council (NERC) - United Kingdom Economic and Social Research Council - United Kingdom Department for International Development, Grant/Award Number: project NE/K010239/1 (Mountain-EVO); NERC and DFID - Science for Humanitarian Emergencies and Resilience (SHEAR) program, Grant/Award Number: project NE/P000452/1 (Landslide EVO). This research has also been funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 734317 (HiFreq). Funds for open access publication fees were received from the University of Birmingham.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00221/full#supplementary-material

## REFERENCES

Acosta-Coll, M., Ballester-Merelo, F., and Martínez-Peiró, M. (2018). Early warning system for detection of urban pluvial flooding hazard levels in an ungauged basin. *Nat. Hazards* 92, 1237–1265. doi: 10.1007/s11069-018-3249-4

Ali, H., Soe, J. K., and Weller, S. R. (2015). "A real-time ambient air quality monitoring wireless sensor network for schools in smart cities," in *Proceedings of the 2015 IEEE First International Smart Cities Conference (ISC2)*, Guadalajara. 2, 5–10.

Baden, T., Chagas, A. M., Gage, G., Marzullo, T., Prieto-Godino, L. L., and Euler, T. (2015). Open labware: 3-d printing your own lab equipment. *PLoS Biol.* 13:e1002175. doi: 10.1371/journal.pbio.1002086

Bakker, K., and Ritts, M. (2018). Smart earth: a meta-review and implications for environmental governance. *Glob. Environ. Change* 52, 201–211. doi: 10.1016/j.gloenvcha.2018.07.011

Barnard, H. R., Findley, M. C., and Csavina, J. (2014). PARduino: a simple and inexpensive device for logging photosynthetically active radiation. *Tree Physiol.* 34, 640–645. doi: 10.1093/treephys/tpu044

Bartesaghi Koc, C., Osmond, P., and Peters, A. (2018). Evaluating the cooling effects of green infrastructure: a systematic review of methods, indicators and data sources. *Sol. Energy* 166, 486–508. doi: 10.1016/j.solener.2018.03.008

Bartos, M., Wong, B., and Kerkez, B. (2018). Open storm: a complete framework for sensing and control of urban watersheds. *Environ. Sci. Water Res. Technol.* 4, 346–358. doi: 10.1039/c7ew00374a

Benedetti, M., Ioriatti, L., Martinelli, M., and Viani, F. (2010). Wireless sensor network: a pervasive technology for earth observation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 3, 488–496. doi: 10.1109/JSTARS.2010.2052917

Bengston, K. J., and Dunbabin, M. D. (2007). "Design & performance of a networked Ad-Hoc acoustic communications system using inexpensive commercial CDMA modems," in *Proceedings of the Oceans 2007-Europe*, (Aberdeen: IEEE), 6.

Blaen, P. J., Khamis, K., Lloyd, C. E. M., Bradley, C., Hannah, D., and Krause, S. (2016). Real-time monitoring of nutrients and dissolved organic matter in rivers: capturing event dynamics, technological opportunities and future directions. *Sci. Total Environ.* 569–570, 647–660. doi: 10.1016/j.scitotenv.2016.06.116

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., et al. (2017). Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environ. Int.* 99, 293–302. doi: 10.1016/j.envint.2016.12.007

Dauwe, S., Oldoni, D., De Baets, B., Van Renterghem, T., Botteldooren, D., and Dhoedt, B. (2014). Multi-criteria anomaly detection in urban noise sensor networks. *Environ. Sci. Process. Impacts* 16, 2249–2258. doi: 10.1039/c4em00273c

Davis, M. J., Thokala, S., Xing, X., Thompson Hobbs, N., Walsh, D. P., Han, R. Y., et al. (2012). Developing a data-transfer model for a novel Wildlife-tracking network. *Wildl. Soc. Bull.* 36, 820–827. doi: 10.1002/wsb.188

Finazzi, F., and Fassò, A. (2017). A statistical approach to crowdsourced smartphone-based earthquake early warning systems. *Stoch. Environ. Res. Risk Assess.* 31, 1649–1658. doi: 10.1007/s00477-016-1240-8

Fisher, R., Ledwaba, L., Hancke, G., and Kruger, C. (2015). Open hardware: a role to play in wireless sensor networks? *Sensors* 15, 6818–6844. doi: 10.3390/s150306818

Geipel, J., Jackenkroll, M., Weis, M., and Claupein, W. (2015). A sensor web-enabled infrastructure for precision farming. *ISPRS Int. J. Geo Inform.* 4, 385–399. doi: 10.3390/ijgi4010385

Gleonec, P. D., Ardouin, J., Gautier, M., and Berder, O. (2017). "Architecture exploration of multi-source energy harvester for IOT nodes," in *Proceedings of the 2016 IEEE Online Conference on Green Communications (OnlineGreenComm)*, Piscataway, NJ, 27–32. doi: 10.1109/OnlineGreenCom.2016.7805402

Hart, J. K., and Martinez, K. (2006). Environmental sensor networks: a revolution in the earth system science? *Earth Sci. Rev.* 78, 177–191. doi: 10.1016/j.earscirev.2006.05.001

Hill, N. R., Fatoba, S. T., Oke, J. L., Hirst, J. A., Callaghan, A. O., Lasserson, D. S., et al. (2016). Global prevalence of chronic kidney disease – a systematic review and meta-analysis. *PLoS One* 11:e0158765. doi: 10.5061/dryad.3s7rd

Horita, F. E. A., Albuquerque, J. P., de Degrossi, L. C., Mendiondo, E. M., and Ueyama, J. (2015). Development of a spatial decision support system for flood risk management in Brazil that combines volunteered geographic information with wireless sensor networks. *Comput. Geosci.* 80, 84–94. doi: 10.1016/j.cageo.2015.04.001

Jones, A. S., Horsburgh, J. S., Reeder, S. L., Ramírez, M., and Caraballo, J. (2015). A data management and publication workflow for a large-scale, heterogeneous sensor network. *Environ. Monit. Assess.* 187:348. doi: 10.1007/s10661-015-4594-3

Karpouzoglou, T., Dewulf, A., and Clark, J. (2016). Advancing adaptive governance of social-ecological systems through theoretical multiplicity. *Environ. Sci. Policy* 57, 1–9. doi: 10.1016/j.envsci.2015.11.011

Katsiri, E., and Makropoulos, C. (2016). An ontology framework for decentralized water management and analytics using wireless sensor networks. *Desalin. Water Treat.* 57, 26355–26368. doi: 10.1080/19443994.2016.1202144

Kim, Y., Jabro, J. D., and Evans, R. G. (2011). Wireless lysimeters for real-time online soil water monitoring. *Irrig. Sci.* 29, 423–430. doi: 10.1007/s00271-010-0249-x

Kizel, F., Etzion, Y., Shafran-Nathan, R., Levy, I., Fishbain, B., Bartonova, A., et al. (2018). Node-to-node field calibration of wireless distributed air pollution sensor network. *Environ. Pollut.* 233, 900–909. doi: 10.1016/j.envpol.2017.09.042

Krause, S., Lewandowski, J., Dahm, C. N., and Tockner, K. (2015). Frontiers in real-time ecohydrology - a paradigm shift in understanding complex environmental systems. *Ecohydrology* 8, 529–537. doi: 10.1002/eco.1646

Kubicek, P., Kozel, J., Stampach, R., and Lukas, V. (2013). Prototyping the visualization of geographic and sensor data for agriculture. *Comput. Electron. Agric.* 97, 83–91. doi: 10.1016/j.compag.2013.07.007

Lopes Pereira, R., Trindade, J., Gonc,alves, F., Suresh, L., Barbosa, D., and Vazão, T. (2014). A wireless sensor network for monitoring volcano-seismic signals. *Nat. Hazards Earth Syst. Sci.* 14, 3123–3142. doi: 10.5194/nhess-14-3123-2014

López, J. A., Navarro, H., Soto, F., Pavón, N., Suardíaz, J., and Torres, R. (2015). GAIA2: a multifunctional wireless device for enhancing crop management. *Agric. Water Manag.* 151, 75–86. doi: 10.1016/j.agwat.2014.10.023

López Riquelme, J. A., Soto, F., Suardíaz, J., Sánchez, P., Iborra, A., and Vera, J. A. (2009). Wireless Sensor Networks for precision horticulture in Southern Spain. *Comput. Electron. Agric.* 68, 25–35. doi: 10.1016/j.compag.2009.04.006

Mao, F., Clark, J., Buytaert, W., Krause, S., and Hannah, D. M. (2018). Water sensor network applications: time to move beyond the technical? *Hydrol. Process.* 32, 2612–2615. doi: 10.1002/hyp.13179

Mazinani, S. M., and Davarzani, S. (2017). Presenting an optimal algorithm based on firefly algorithm with specific parameters to select the cluster head in wireless sensor networks in order to reduce energy consumption. *QUID Investig. Cienc. Tecnol.* 1, 1936–1941.

Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., et al. (2013). The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmos. Environ.* 70, 186–203. doi: 10.1016/j.atmosenv.2012.11.060

Mulrow, C. (1987). The medical review article: state of the science. *Ann. Intern. Med.* 106, 485–488.

Munir, S., Mayfield, M., Coca, D., Jubb, S. A., and Osammor, O. (2019). Analysing the performance of low-cost air quality sensors, their drivers, relative benefits and calibration in cities—a case study in Sheffield. *Environ. Monit. Assess.* 191:94. doi: 10.1007/s10661-019-7231-8

Ojha, T., Misra, S., and Raghuwanshi, N. S. (2015). Wireless sensor networks for agriculture: the state-of-the-art in practice and future challenges. *Comput. Electron. Agric.* 118, 66–84. doi: 10.1016/j.compag.2015.08.011

Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., et al. (2018). Citizen science for hydrological risk reduction and resilience building. *Wiley Interdiscip. Rev. Water* 5:e1262. doi: 10.1002/wat2.1262

Pearce, J. (2013). *Open-Source Lab: How to Build Your Own Hardware and Reduce Research Costs*. Amsterdam: Elsevier.

Pickering, C., and Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *High. Educ. Res. Dev.* 33, 534–548. doi: 10.1080/07294360.2013.841651

Pohl, S., Garvelmann, J., Wawerla, J., and Weiler, M. (2014). Potential of a low-cost sensor network to understand the spatial and temporal dynamics of a mountain snow cover. *Water Resour. Res.* 50, 2533–2550. doi: 10.1002/2013WR014594

Prasad, P. (2015). Recent trend in wireless sensor network and its applications: a survey. *Sens. Rev.* 35, 229–236. doi: 10.1108/SR-08-2014-683

Pumo, D., Francipane, A., Lo Conti, F., Arnone, E., Bitonto, P., Viola, F., et al. (2016). The SESAMO early warning system for rainfall-triggered landslides. *J. Hydroinform.* 18, 256–276. doi: 10.2166/hydro.2015.060

Rai, A. C., Kumar, P., Pilla, F., Skouloudis, A. N., Di Sabatino, S., Ratti, C., et al. (2017). End-user perspective of low-cost sensors for outdoor air pollution monitoring. *Sci. Total Environ.* 607–608, 691–705. doi: 10.1016/j.scitotenv.2017.06.266

Rettig, A. J., Khanna, S., Heintzelman, D., and Beck, R. A. (2014). An open source software approach to geospatial sensor network standardization for urban runoff. *Comput. Environ. Urban Syst.* 48, 28–34. doi: 10.1016/j.compenvurbsys.2014.05.003

Rogulski, M. (2018). Using low-cost PM monitors to detect local changes of air quality. *Polish J. Environ. Stud.* 27, 1699–1706. doi: 10.15244/pjoes/77075

Rossiter, D. G. (2018). Past, present & future of information technology in pedometrics. *Geoderma* 324, 131–137. doi: 10.1016/j.geoderma.2018.03.009

Schneider, J., Schultz, L. E., Mancha, S., Hicks, E., and Smith, R. N. (2016). "Development of a portable water quality sensor for river monitoring from small rafts," in *Proceedings of the OCEANS 2016 MTS/IEEE Monterey*, (Monterey, CA: IEEE), 1–10. doi: 10.1109/OCEANS.2016.7761392

Schneider, P., Castell, N., Vogt, M., Dauge, F. R., Lahoz, W. A., and Bartonova, A. (2017). Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environ. Int.* 106, 234–247. doi: 10.1016/j.envint.2017.05.005

Šećerov, I. B., Savić, S. M., Milošević, D. D., Arsenović, D. M., Dolinaj, D. M., and Popov, S. B. (2019). Progressing urban climate research using a high-density monitoring network system. *Environ. Monit. Assess.* 191:89. doi: 10.1007/s10661-019-7210-0

Shaw, R., Lark, R. M., Williams, A. P., Chadwick, D. R., and Jones, D. L. (2016). Characterising the within-field scale spatial variation of nitrogen in a grassland soil to inform the efficient design of in-situ nitrogen sensor networks for precision agriculture. *Agric. Ecosyst. Environ.* 230, 294–306. doi: 10.1016/j.agee.2016.06.004

Shusterman, A. A., Kim, J., Lieschke, K. J., Newman, C., Wooldridge, P. J., and Cohen, R. C. (2018). Observing local CO2 sources using low-cost, near-surface urban monitors. *Atmos. Chem. Phys.* 18, 13773–13785. doi: 10.5194/acp-18-13773-2018

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., et al. (2013). The changing paradigm of air pollution monitoring. *Environ. Sci. Technol.* 47, 11369–11377. doi: 10.1021/es4022602

Srbinovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., and Borozan, V. (2015). Environmental parameters monitoring in precision agriculture using wireless sensor networks. *J. Clean. Prod.* 88, 297–307. doi: 10.1016/j.jclepro.2014.04.036

Strigaro, D., Cannata, M., and Antonovic, M. (2019). Boosting a weather monitoring system in low income economies using open and non-conventional systems: data quality analysis. *Sensors* 19:1185. doi: 10.3390/s19051185

UN (2015). *The Millennium Development Goals Report 2015*. New York, NY: UN.

van de Giesen, N., Hut, R., and Selker, J. (2014). The trans-african hydro-meteorological observatory (TAHMO). *Wiley Interdiscip. Rev. Water* 1, 341–348. doi: 10.1002/wat2.1034

van Zoest, V. M., Stein, A., and Hoek, G. (2018). Outlier detection in urban air quality sensor networks. *Water. Air. Soil Pollut.* 229:111. doi: 10.1007/s11270-018-3756-7

Wada, M., Hatanaka, K., and Toda, M. (2007). "Practical use of personal buoy system for fishery using sensor network technologies," in *Proceedings of the Ocean 2007-Europe*, (Vancouver, BC: IEEE), 1–6.

Walter, K. (2010). Development of an early warning information infrastructure using spatial web services technology. *Int. J. Digit. Earth* 3, 384–394. doi: 10.1080/17538947.2010.486871

Wickham, H., and Grolemund, G. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 1st Edn. Newton, MA: O'Reilly Media.

Xu, L., and Marinova, D. (2013). Resilience thinking: a bibliometric analysis of socio-ecological research. *Scientometrics* 96, 911–927. doi: 10.1007/s11192-013-0957-0

**Conflict of Interest Statement:** 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.

*Copyright © 2019 Mao, Khamis, Krause, Clark and Hannah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

BRIEF RESEARCH REPORT
published: 27 September 2019
doi: 10.3389/feart.2019.00244

# Assessing the Sampling Quality of a Low-Tech Low-Budget Volume-Based Rainfall Sampler for Stable Isotope Analysis

*Benjamin M. C. Fischer[1,2]*, Franziska Aemisegger[3], Pascal Graf[3], Harald Sodemann[3,4,5] and Jan Seibert[1,6,7]*

[1] Department of Physical Geography, Stockholm University, Stockholm, Sweden, [2] Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden, [3] Institute for Atmospheric and Climate Science, ETH Zürich, Zurich, Switzerland, [4] Geophysical Institute, University of Bergen, Bergen, Norway, [5] Bjerknes Centre for Climate Research, Bergen, Norway, [6] Department of Geography, University of Zurich, Zurich, Switzerland, [7] Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden

OPEN ACCESS

**Edited by:**
Theresa Blume,
German Research Centre
for Geosciences, Helmholtz Centre
Potsdam, Germany

**Reviewed by:**
Ahmed M. ElKenawy,
Mansoura University, Egypt
Christoff Andermann,
German Research Centre
for Geosciences, Helmholtz Centre
Potsdam, Germany
Natalie Orlowski,
Albert-Ludwigs-Universität Freiburg,
Germany

***Correspondence:**
Benjamin M. C. Fischer
benjamin.fischer@natgeo.su.se

**Specialty section:**
This article was submitted to
Hydrosphere,
a section of the journal
Frontiers in Earth Science

**Received:** 31 January 2019
**Accepted:** 02 September 2019
**Published:** 27 September 2019

**Citation:**
Fischer BMC, Aemisegger F,
Graf P, Sodemann H and Seibert J
(2019) Assessing the Sampling
Quality of a Low-Tech Low-Budget
Volume-Based Rainfall Sampler
for Stable Isotope Analysis.
Front. Earth Sci. 7:244.
doi: 10.3389/feart.2019.00244

To better understand the small-scale variability of rainfall and its isotopic composition it is advantageous to utilize rain samplers which are at the same time low-cost, low-tech, robust, and precise with respect to the collected rainwater isotopic composition. We assessed whether a self-built version of the Kennedy sampler is able to collect rainwater consistently without mixing with antecedent collected water. We called the self-built sampler made from honey jars and silicon tubing the Zurich sequential sampler. Two laboratory experiments show that high rainfall intensities can be sampled and that the volume of water in a water sample originating from a different bottle was generally less than 1 ml. Rainwater was collected in 5 mm increments for stable isotope analysis using three (year 2011) and five (years 2015 and 2016) rain samplers in Zurich (Switzerland) during eleven rainfall events. The standard deviation of the total rainfall amounts between the different rain gauges was <1%. The standard deviation of $\delta^{18}O$ and $\delta^{2}H$ among the different sequential sampler bottles filled at the same time was generally <0.3‰ for $\delta^{18}O$ and <2‰ for $\delta^{2}H$ (8 out of 11 events). Larger standard deviations could be explained by leaking bottle(s) with subsequent mixing of water with different isotopic composition of at least one out of the five samplers. Our assessment shows that low-cost, low-tech rain samplers, when well maintained, can be used to collect sequential samples of rainfall for stable isotope analysis and are therefore suitable to study the spatio-temporal variability of the isotopic composition of rainfall.

**Keywords: rainfall and its isotopic composition, sequential rainwater sampler, laboratory experiments, field test, stable isotopes ($^{18}O$ and $^{2}H$), low-cost/low-tech self-built sampler**

## INTRODUCTION

The stable isotopologues of water ($H_2{}^{16}O$, $H_2{}^{18}O$ $^{1}H_2H^{16}O$, hereafter referred to as isotopes), are valuable tracers to study long-term changes in climatic conditions (Dansgaard et al., 1993), atmospheric processes at the weather system timescale (Pfahl et al., 2012), and are useful to understand how catchments transform rainfall into runoff. The regional variations in the isotopic

composition of rainfall are relatively well understood thanks to the monthly dataset from the Global Network of Isotopes in Precipitation (GNIP; Aragus-Aragus et al., 2000). In contrast, the small-scale spatial variability of the isotopic composition of rainfall (<10 $km^2$) has been much less studied and is often assumed to be homogenous. To better understand the isotopic composition of rainfall it is necessary to collect rainwater which is temporally resolved, and in a spatially distributed way.

Rainfall collection for stable isotope analysis started with holding a bottle of beer and a funnel in the rain (Dansgaard, 2004). Bottles have been used since then, e.g., for crowdsourced snapshot information of super storm Sandy (Good et al., 2014). Despite successful examples of manual sampling by Hrachowitz et al. (2011) or Graf et al. (2019) the need for several people and the right-on-time presence of staff and logistics during an event, makes this type of sampling demanding.

Higher quality information can be obtained by using rain samplers that collect sequential samples of rainwater in either volumetric or temporal intervals. Ideally, the rain sampler should collect rainwater at fixed temporal or volumetric intervals without any mixing of different samples. Furthermore, a sampler should be low-cost, compact, work autonomously, and consume low amounts of electricity. In addition, the sampler needs to be easy to handle, enable fast sample collection, and allow for repair in the field.

Different volume- or time-based rain samplers exist (Laquer, 1990). Rain samplers for stable isotope analysis can be self-built (Prechsl et al., 2014) or commercial (Gröning et al., 2012) cumulative precipitation collectors. However, as shown in several studies, the rainfall isotopic composition changes during the rainfall event (McDonnell et al., 1990; Munksgaard et al., 2012; Aemisegger et al., 2015; Fischer et al., 2017b; Graf et al., 2019). Therefore, it is crucial to collect sequential samples of rainfall with a high temporal resolution to capture these variations in stable isotope composition. Commercially available sequential samplers such as revolver type samplers (e.g., Rücker et al., 2019) usually need electricity to be operated and are costly. Self-built sequential samplers, using open low-budget microcontrollers, e.g., Arduino™ (Aemisegger et al., 2015; Nelke and Selker, 2015; Hartmann et al., 2018; Ankor et al., 2019; Michelsen et al., 2019) are flexible but need energy, and a certain level of electronic knowledge is required. Instead, field-deployed laser spectrometers allow the isotopic composition of rainfall to be measured directly in the field at a high temporal frequency (Berman et al., 2009; Munksgaard et al., 2012; Tweed et al., 2016; von Freyberg et al., 2016). However, the high investment cost and high-tech character make it unfeasible to use this type of sampler to collect rainfall at a high spatial resolution in small catchments. In contrast to such high-tech high cost samplers, the Kennedy sequential sampler (Kennedy et al., 1979), which is used in many hydrological studies (McDonnell et al., 1990; James and Roulet, 2009; Šanda et al., 2014; Fischer et al., 2017b), meets many aforementioned requirements of an ideal sampler. However, it is not clear whether this sampler is able to collect rainwater without any mixing of subsequent samples. Therefore, in this study we built a version of the Kennedy sampler and evaluated its functioning in two ways: (1) a laboratory experiment using deionized water and a salt solution and, (2) a field experiment based on the comparison of the isotopic composition of sequentially sampled rainfall collected by multiple samplers during eleven rainfall events.

# MATERIALS AND METHODS

## The Zurich Sequential Sampler

The Zurich sequential sampler (ZRS-sampler) is an adapted version of the Kennedy type volume-based sequential rainfall sampler (Kennedy et al., 1979) and uses low-cost materials such as 100 ml honey jars, silicon tubing, plastic connectors (which can be potentially 3D-printed), and a plastic box enclosure (**Figure 1**). The ZRS-sampler design resulted from experimenting with different bottles using rubber plugs, tube diameters and materials, radii of transport tubes, and air vents to minimize the mixing of new rainwater collected, with water which was previously collected and stored in the series of interconnected bottles. With too small sample volumes, water droplets in the tubing might introduce memory effects. Therefore, a bottle volume of 100 ml was chosen. To also collect data on the rainfall amounts over time, we directly connected the ZRS-sampler to a tipping bucket rain gauge by attaching a small funnel to each of the two drains at the base of the tipping bucket rain gauge (**Figures 1A,D**, Rain collector II – tipping bucket; 0.2 mm; Davis Instruments Corp., United States, rim height installed at 1.5 m above ground level). To each funnel, a 10 cm silicon tube is connected with a Y-connector (6–7 mm, Kartell, Italy) which is connected to the sampler with a silicon tube (1.5 m). The ZRS-sampler consists of a frame (plastic sheet, 350 × 250 × 3 mm, L × W × H), where 12 × 100 ml screw top glass bottles, each representing 5 mm of rainwater, are attached with their metal screw lids (2 × M3 screws and silicone adhesive to ensure a sealed watertight connection). The different bottles are connected serially to each other using silicon tubing (Ø 9 mm OD). To divert rainwater into a bottle, a bifurcation is made using a Y-connector connected to a 100 mm vertical silicon tube (Ø 9 mm OD) reaching the bottom of the bottles. Each bottle has a smaller second silicon tube attached to the lid (Ø 3 mm OD, 500 mm, small Ø chosen to prevent fractionation from evaporation) acting as an air vent to regulate the atmospheric pressure in each bottle and prevent the water from siphoning. For a correct functioning of the sampler, these air vents always need to be vertical. Once the water level reaches the air vent, a headspace of 0.5 cm filled with air remains, no additional rainwater can enter into the bottle, and water flows to the next empty bottle without mixing with the antecedent collected water. For transport, protection and to minimize solar radiation, each ZRS-sampler is enclosed in a plastic box (UTZ-Rako 400 × 300 × 120 mm). After the last bottle, excess rainfall flows through a tube into the plastic box. It is also possible to connect a second sequential sampler to capture large rainfall amounts.

The cost per sampler is approximately €330 including tipping bucket or €85 when using a 214 $cm^2$ funnel instead of a tipping bucket (for price per sampler, see **Supplementary Table S1**).

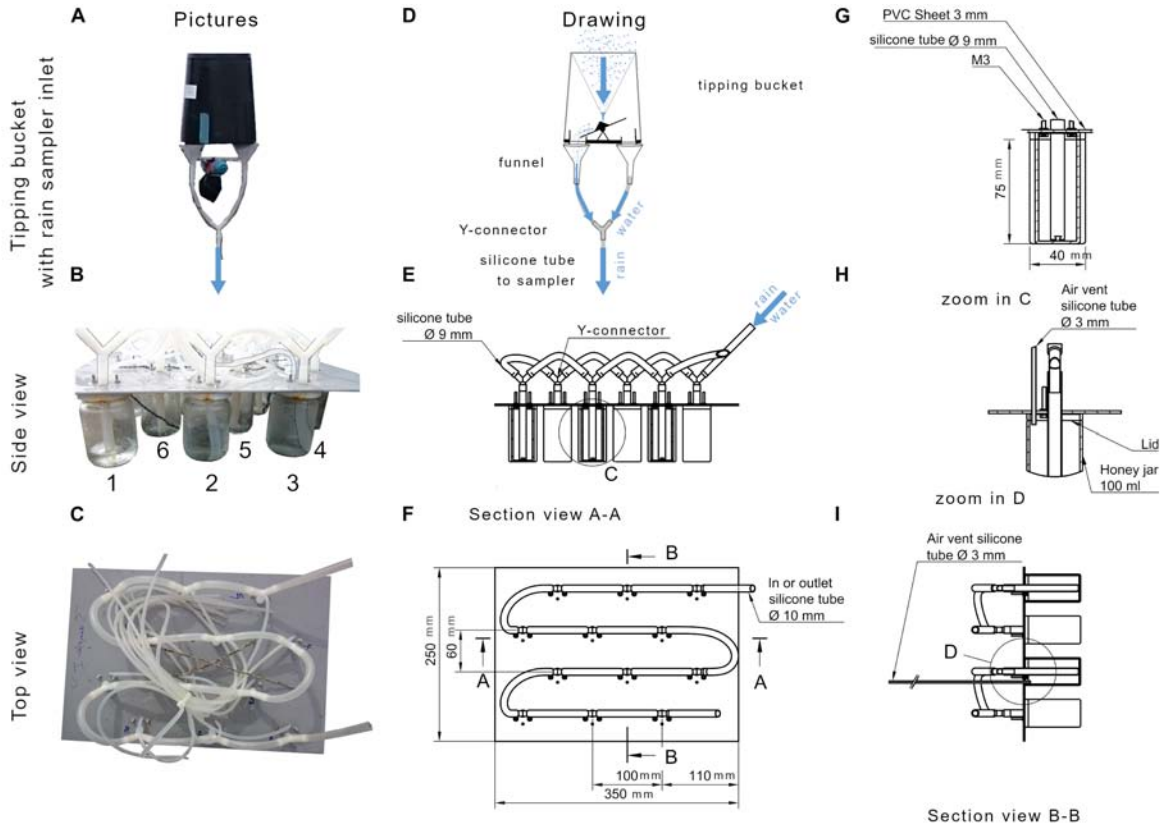



**FIGURE 1 |** Overview of the ZRS-sampler tipping bucket with attached funnels **(A)**, side view (numbers indicate sample bottles, each 100 ml or 5 mm of precipitation), side view **(B)** is an example of a malfunctioning sampler where in rainfall event June 2016, honey jar 1 and 2 (left to right) were leaking (not full) and therefore not properly sampling, resulting in mixing of water of different isotopic composition and top view **(C)**. Drawing of the tipping bucket **(D)**, section view A–A **(E)**, top view **(F)**, zoom in C (bottle and connections) **(G)**, zoom in D (air vent and connection of glass) **(H)**, and section view B–B **(I)**. In case the study requires a high or lower temporal resolution, i.e., rainwater samples representing <5 mm or >5 mm rainfall amount, it is possible to connect honey jars with smaller or larger volumes (having a similar honey jar lid diameter).

## Experimental Setup, Isotope Analysis, and Comparison

The mixing of different water samples was assessed in two laboratory experiments: I. testing the maximum rainfall intensity before sampling errors occur, and II. assessing the mixing, i.e., sampling error and memory effect within the sampler (**Supplementary Material** and **Supplementary Figure S1**, laboratory experiment).

Furthermore, the mixing of different water samples was additionally assessed by collecting rainwater during rainfall events using three (year 2011) and five (years 2015 and 2016) ZRS-samplers installed within a distance of 2 m (**Supplementary Material** and **Supplementary Figure S2**, laboratory experiment).

# RESULTS AND DISCUSSION

## Laboratory Experiment

The rainfall sampler was able to collect water correctly for rainfall intensities up to 1 mm $s^{-1}$ (comparable to a very high rain burst, e.g., pour 1, **Supplementary Video S1**). For higher rainfall intensities, which rarely occur in natural rainfall events, water entered not only in the first empty bottle but in multiple bottles (pour two onward, **Supplementary Video S1**).

The second laboratory experiment revealed that in some bottles, mixing of different water samples occurred, which was visible from the color of the water (**Supplementary Videos S2–S6**). In addition to the color indicator, the electrical conductivity increased or decreased between 0 and 50 μS $cm^{-1}$, to what the electrical conductivity of the originally collected water was (**Supplementary Table S2**). Despite the increase or decrease in electrical conductivity due to mixing of different water samples, from the mass balance (**Supplementary Equation S1**), the volume of water in a water sample originating from a different bottle was generally less than 1 ml (<1% of the sample volume, 10 out of 18 bottles, **Supplementary Table S2**). Mixing of more than 1 ml was due to memory effects, i.e., antecedent water remained in the tubing and mixed with the newly poured water. In 3 out of 18 pours, the water was accidentally poured at rates >1 mm $s^{-1}$, resulting in a non-correct sampling due to trapped air bubbles

in the tubing and consequently a volume of water which mixed with water of a different bottle of more than 4 ml or more than 4% of the sample volume (Test 3 in **Supplementary Table S2** and **Supplementary Videos S2–S6**).

These results indicate that overall the sampler can be used to collect sequential samples of rainfall with a minimal mixing between different bottles. These samples can subsequently be used to determine the isotopic composition of rainfall increments.

## Sampled Rainfall Amount and Its Isotopic Composition of Different Events

The different ZRS-samplers collected rainwater samples of 11 rainfall events ($P_{tot}$ 1 to 30 mm, $S_{D\ Ptot}$ <1%) with an isotopic composition that is aligned along the global meteoric water line, **Figure 2**, **Supplementary Figure S3**, and **Supplementary Table S3**.

The rooftop location was chosen for practical reasons, but does not comply with WMO recommendations for rainfall measurements (WMO, 2008) because of wind exposure. However, the variation of the measured rainfall amounts between the gauges was small for all events ($S_{D\ Ptot}$ <1%). As the different ZRS-samplers received a similar amount of rainwater, it can be assumed that the isotopic composition of the water samples collected by the different samplers should be similar.

For all events with rainfall $P_{tot}$ >5 mm the $\delta^{18}O$ decreased with subsequent samples, i.e., in time (**Figure 2** and **Supplementary Table S3**). For events with low rainfall amounts (e.g., October 15, 2015; October 16, 2015; and October 17, 2015) the last sample bottle in each sampler was partly filled. No difference in the isotopic composition between samplers with and without tipping buckets could be observed (individual samples within the $S_D$ of $\delta^{18}O$ or $\delta^{2}H$ to the mean $\delta^{18}O$ or $\delta^{2}H$ of all samplers, **Supplementary Table S3**). Also, no correlation was found between the rainfall amount and $S_D$ (**Supplementary Figure S5**). Comparing the isotopic composition of the water in the bottles filled at the same time, the $S_{D\ 18O}$ generally ranged from 0.01 to 1.5‰ and $S_{D\ 2H}$ generally ranged from 0.01 to 4‰ (**Supplementary Table S3**). Only for three events, some water samples taken at the same time had an $S_{D\ 18O}$ up to 5.8‰ and $S_{D\ 2H}$ up to 45‰ (**Supplementary Table S3**). This considerable standard deviation, which is almost as large as the temporal variability of the $\delta^{18}O$ from bottle to bottle, can be explained by malfunctioning (leaking bottles by which water samples of different isotopic composition were mixing) of one or some ZRS-sampler(s) during the collection of samples (**Figure 1B**). During some events, one or more of the honey jar lids were not closed tightly enough, and water and air were leaking. This malfunctioning resulted in only partially filled bottles and mixing of different water samples (e.g., ZRS-sampler S-3 for events May 25, 2011 and June 1, 2011, and ZRS-sampler S-5 event November 19, 2015). When removing these outliers (known from field notes and indicated in **Supplementary Table S3** with the letter L), the $S_D$ was generally near the laboratory analytical precision (7 out of 11 events the $S_{D\ 18O}$ <0.3‰ and $S_{D\ 2H}$ <2‰). After October 2015, the $S_{D\ 18O}$ remained higher and the samplers showed signs of aging (after sampling approximately 50 rainfall events over 4 years). The frequent screw movements of the bottles, deformed the thin honey jar lids and the seals degraded. Despite maintenance and repairs, from October 2015 onward, bottles started to leak more often and were more challenging to repair, resulting in larger differences ($S_{D\ 18O}$ >0.3‰ or $S_{D\ 2}$ >2‰) caused by mixing of the water samples with a different isotopic composition. This shows that it is necessary to continuously monitor the state of the sampler, performing maintenance and repairs to achieve a recommended deployment time of approximately 4 years.

# CONCLUSION AND RECOMMENDATIONS

Our assessment of the self-built ZRS-sampler showed that by using honey jars and silicon tubing, it is possible to collect rainwater samples in incremental volumes with minimal mixing between sample bottles (less than 1 ml or 1% of the sample volume). During natural rainfall events, the standard deviation among the different sequential sampler bottles filled at the same time was generally <0.3‰ for $\delta^{18}O$ and <2‰ for $\delta^{2}H$.

Similar samplers can be built from any material, e.g., honey jars, PET, beer, milk, and laboratory glass bottles and any kind of tubing. Notably, the low-cost, low-tech character makes this type of sequential sampler useful for investigating the spatial-temporal variability in the isotopic composition of rainfall.

However, to work correctly, and to minimize both technical and human errors when collecting rainwater samples for stable isotope analysis, from the findings in this study, we recommend:

1. Air vents need to be vertical at all times to function as intended.
2. It is essential to test for leaking bottles. Before sampling rainfall events this can be tested by filling the sampler with water and examining if all bottles fill with water to the level of the air vents. In the field, when samples are picked up, leaking bottles can be detected by checking if bottles are not correctly filled, e.g., water does not reach the air vent. Leaking bottles can be repaired using O-rings or by replacing the metal honey jar lid.
3. Full samplers must be lifted vertically. Tilting might cause mixing or spilling of rainwater collected in the different honey jars (bottles).
4. Before transferring the rainwater into transport bottles (labeled with the sampler number, bottle number and date) slightly shake each bottle to homogenize collected water.
5. To avoid unintentional mixing of rainwater of different rainfall events, remove excess water from the sequential sampler's tubing (shake thoroughly), as well as from honey jars (use a cloth) and rain gauge bucket (**Supplementary Figure S4**). After controlling all tubing connections for leaks, the sampler is ready for a next rainfall event.

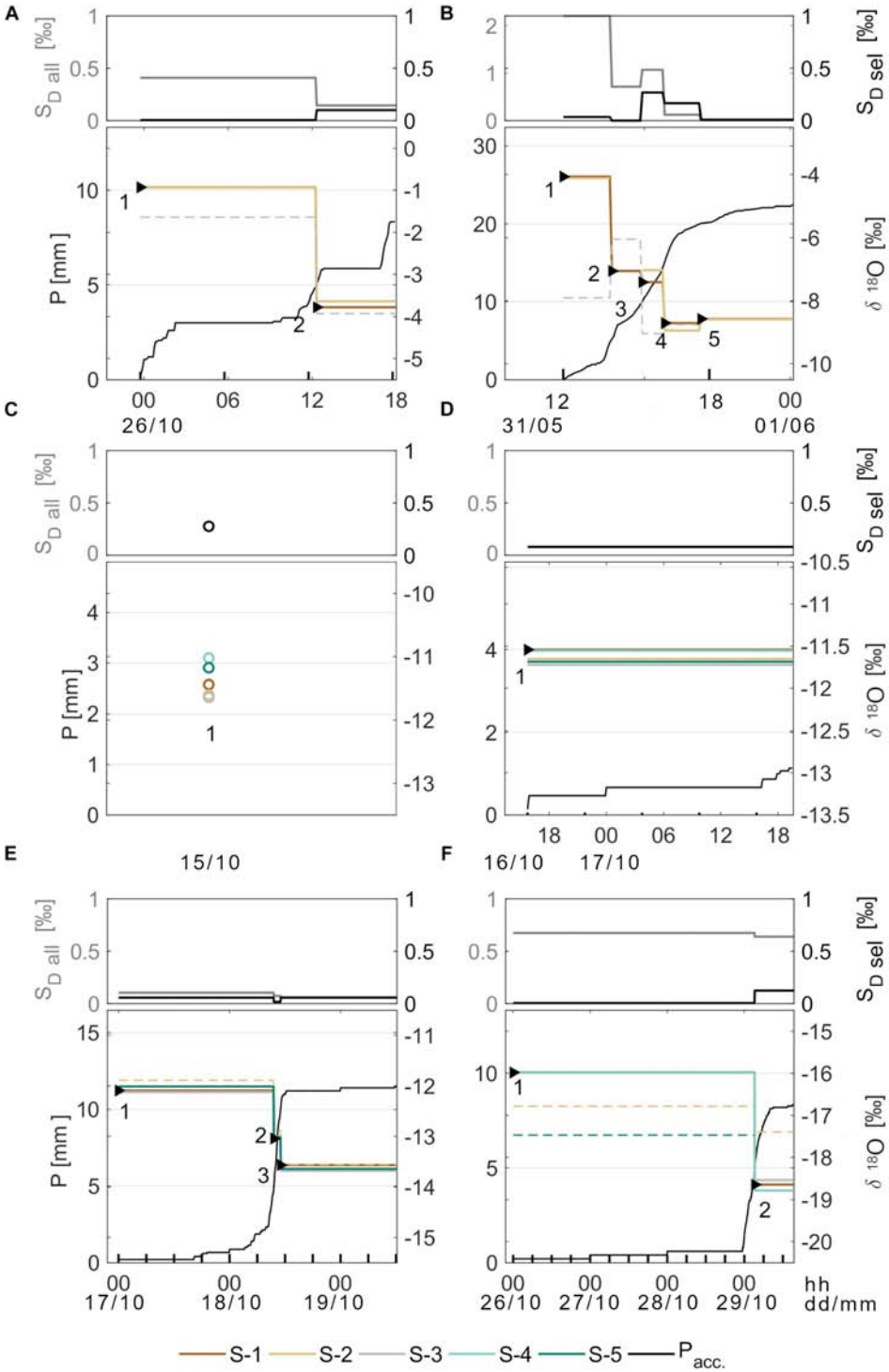



**FIGURE 2 |** Continued

6. It should be noted that during frontal passages with substantial temperature changes or pressure drops during the event, stronger cross-contamination might occur between the samples due to pressure fluctuations within the vials. However, in an experiment we found no evidence for cross-contamination between the samples caused by an expansion of air and water due to temperature fluctuations.

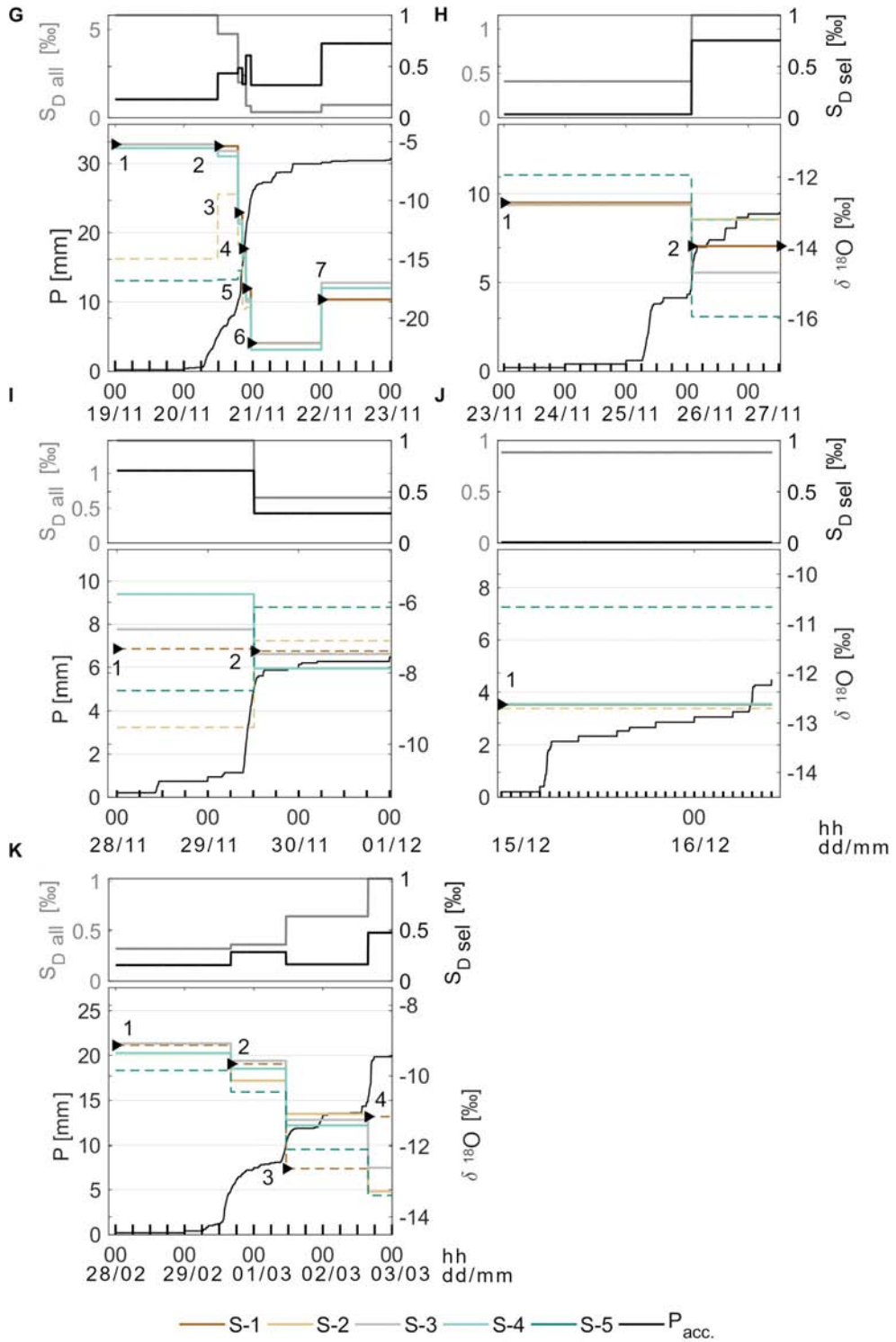



**FIGURE 2 |** For the 11 rainfall events **(A–K)**, the mean accumulated rainfall (black line, left axis) and the $\delta^{18}O$ of rainfall (5 mm increments) of the different rain samplers (S-1 to S-5; colored lines). The triangles indicate the start of a new bottle which are consecutively numbered number 1 to *n*. The solid lines indicate a correct $\delta^{18}O$ sample while a dashed line indicates $\delta^{18}O$ samples, for which the rain samplers malfunctioned due to technical problems. On top of each panel the temporal evolution of the $S_D$ $\delta^{18}O$ all rain samplers (gray line, left *y*-axis) and $S_D$ $\delta^{18}O$ sel rain samplers excluding malfunctioning rain samplers (black line, right *y*-axis).

Furthermore, from our experience using this sampler we recommend the following additional points be considered:

7. The water samples should be collected directly after a rainfall event. A dedicated test in a climate or sprinkling chamber, simulating different environmental changes in a controlled way, could be done to quantify the fractionation and importance of cross-contamination effects in the case of delayed collection of samples.
8. Lids and tubes should be replaced about every 4 years. Instead of using honey jars, more durable threaded caps and wide mouth bottles (100 or 250 ml; Duran or Nalgene) could be used.

## AUTHOR CONTRIBUTIONS

BF designed and built the sampler and wrote the first draft of the manuscript. BF, FA, and PG developed the field concept and design of the project and collected the data. All authors contributed to the manuscript revision, and read and approved the submitted version.

## ACKNOWLEDGMENTS

We thank all the people who helped and contributed to this study: Bruno Kägi, Claudia Schreiner, Michael Hilf, Sandra Röthlisberger, Roland Werner, Peter Isler, and Urs Beyerle. Barbara Herbstritt for the isotope analysis of rainwater samples (2015) and Valentin Mansanarez for lending his video equipment used in the laboratory experiments. In particular, we thank Heini Wernli for support and feedback to an earlier version of the manuscript and Ivan Woodhatch for his help, good humor, and useful input to the different versions of the ZRS-sampler. We thank Tracy Ewen for proofreading the manuscript, and the editor and the three reviewers for their constructive comments and suggestions that helped to significantly improve the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00244/full#supplementary-material

## REFERENCES

Aemisegger, F., Spiegel, J. K., Pfahl, S., Sodemann, H., Eugster, W., and Wernli, H. (2015). Isotope meteorology of cold front passages: a case study combining observations and modeling. *Geophys. Res. Lett.* 42, 5652–5660. doi: 10.1002/2015GL063988

Ankor, M. J., Jonathan, T. J., and Hughes, C. E. (2019). Development of an autonomous, monthly and daily, rainfall sampler for isotope research. *J. Hydrol.* 575, 31–41. doi: 10.1016/j.jhydrol.2019.04.074

Aragus-Aragus, L., Froehlich, K., and Rozanski, K. (2000). Deuterium and oxygen-18 isotope composition of precipitation and atmospheric moisture. *Hydrol. Process.* 14, 1341–1355. doi: 10.1002/1099-1085(20000615)14:8<1341::AID-HYP983<3.0.CO;2-Z

Berman, E. S. F., Gupta, M., Gabrielli, C., Garland, T., and McDonnell, J. J. (2009). High-frequency field-deployable isotope analyzer for hydrological applications. *Water Resour. Res.* 45, 1–7. doi: 10.1029/2009WR008265

Coplen, T. B. (2011). Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Commun. Mass Spectrom.* 25, 2538–2560. doi: 10.1002/rcm.5129

Dansgaard, W. (2004). *Frozen Annals, Greenland Ice Sheet Research*. Copenhagen: The Niels Bohr Institute for Astronomy.

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., et al. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* 364, 218–220. doi: 10.1038/364218a0

Fischer, B. M. C., Stähli, M., and Seibert, J. (2017a). Pre-event water contributions to runoff events of different magnitude in pre-alpine headwaters. *Hydrol. Res.* 48, 28–47. doi: 10.2166/nh.2016.176

Fischer, B. M. C., (Ilja), van Meerveld, H. J., and Seibert, J. (2017b). Spatial variability in the isotopic composition of rainfall in a small headwater catchment and its effect on hydrograph separation. *J. Hydrol.* 547, 755–769. doi: 10.1016/j.jhydrol.2017.01.045

Good, S. P., Mallia, D. V., Lin, J. C., and Bowen, G. J. (2014). Stable isotope analysis of precipitation samples obtained via crowdsourcing reveals the spatiotemporal evolution of superstorm sandy. *PLoS One* 9:e91117. doi: 10.1371/journal.pone.0091117

Graf, P., Wernli, H., Pfahl, S., and Sodemann, H. (2019). A new interpretative framework for below-cloud effects on stable water isotopes in vapour and rain. *Atmos. Chem. Phys.* 19, 747–765. doi: 10.5194/acp-19-747-2019

Gröning, M., Lutz, H. O., Roller-Lutz, Z., Kralik, M., Gourcy, L., and Pöltenstein, L. (2012). A simple rain collector preventing water re-evaporation dedicated for $\delta^{18}$O and $\delta^{2}$H analysis of cumulative precipitation samples. *J. Hydrol.* 448–449, 195–200. doi: 10.1016/j.jhydrol.2012.04.041

Hartmann, A., Luetscher, M., Wachter, R., Holz, P., Eiche, E., and Neumann, T. (2018). Technical note: GUARD – an automated fluid sampler preventing sample alteration by contamination, evaporation and gas exchange, suitable for remote areas and harsh conditions. *Hydrol. Earth Syst. Sci.* 22, 4281–4293. doi: 10.5194/hess-22-4281-2018

Hrachowitz, M., Bohte, R., Mul, M. L., Bogaard, T. A., Savenije, H. H. G., and Uhlenbrook, S. (2011). On the value of combined event runoff and tracer analysis to improve understanding of catchment functioning in a data-scarce semi-arid area. *Hydrol. Earth Syst. Sci.* 15, 2007–2024. doi: 10.5194/hess-15-2007-2011

James, A. L., and Roulet, N. T. T. (2009). Antecedent moisture conditions and catchment morphology as controls on spatial patterns of runoff generation in small forest catchments. *J. Hydrol.* 377, 351–366. doi: 10.1916/j.jhydrol.2009.08.039

Kennedy, V. C., Zellweger, G. W., and Avanzino, R. J. (1979). Variation of rain chemistry during storms at two sites in northern California. *Water Resour. Res.* 15, 687–702. doi: 10.1029/WR015i003p00687

Laquer, F. C. (1990). Sequential precipitation samplers: a literature review. *Atmos. Environ. Part A Gen. Top.* 24, 2289–2297. doi: 10.1016/0960-1686(90)90322-E

McDonnell, J. J., Bonell, M., Stewart, M. K., and Pearce, A. J. (1990). Deuterium variations in storm rainfall: implications for stream hydrograph separation. *Water Resour. Res.* 26, 455–458. doi: 10.1029/WR026i003p00455

Michelsen, N., Laube, G., Friesen, J., Weise, S. M., Bait Said, A. B. A., and Müller, T. (2019). Technical note: a microcontroller-based automatic rain sampler for stable isotope studies. *Hydrol. Earth Syst. Sci.* 23, 2637–2645. doi: 10.5194/hess-23-2637-2019

Munksgaard, N. C., Wurster, C. M., Bass, A., and Bird, M. I. (2012). Extreme short-term stable isotope variability revealed by continuous rainwater analysis. *Hydrol. Process.* 26, 3630–3634. doi: 10.1002/hyp.9505

Nelke, M., and Selker, J. (2015). *OPEnSampler*. Available at: http://www.open-sensing.org/opensampler (accessed January 12, 2018).

Penna, D., Stenni, B., Šanda, M., Wrede, S., Bogaard, T. A., Gobbi, A., et al. (2010). On the reproducibility and repeatability of laser absorption spectroscopy measurements for δ2H and δ18O isotopic analysis. *Hydrol. Earth Syst. Sci.* 14, 1551–1566. doi: 10.5194/hess-14-1551-2010

Pfahl, S., Wernli, H., and Yoshimura, K. (2012). The isotopic composition of precipitation from a winter storm a case study with the limited-area model COSMOiso. *Atmos. Chem. Phys.* 12, 1629–1648. doi: 10.5194/acp-12-1629-2012

Prechsl, U. E., Gilgen, A. K., Kahmen, A., and Buchmann, N. (2014). Reliability and quality of water isotope data collected with a low-budget rain collector. *Rapid Commun. Mass Spectrom.* 28, 879–885. doi: 10.1002/rcm.6852

Rücker, A., Zappa, M., Boss, S., and Freyberg, J. V. (2019). An optimized snowmelt lysimeter system for monitoring melt rates and collecting samples for stable water isotope analysis. *J. Hydrol. Hydromech.* 67, 20–31. doi: 10.2478/johh-2018-0007

Šanda, M., Vitvar, T., Kulasová, A., Jankovec, J., and Císlerová, M. (2014). Run-off formation in a humid, temperate headwater catchment using a combined hydrological, hydrochemical and isotopic approach (Jizera Mountains, Czech Republic): catchment run-off formation: hydrology and isogeochemistry. *Hydrol. Process.* 28, 3217–3229. doi: 10.1002/hyp.9847

Tweed, S., Munksgaard, N., Marc, V., Rockett, N., Bass, A., Forsythe, A. J., et al. (2016). Continuous monitoring of stream δ18O and δ2H and stormflow hydrograph separation using laser spectrometry in an agricultural catchment. *Hydrol. Process.* 30, 648–660. doi: 10.1002/hyp.10689

von Freyberg, J., Studer, B., and Kirchner, J. W. (2016). A lab in the field: high-frequency analysis of water quality and stable isotopes in streamwater and precipitation. *Hydrol. Earth Syst. Sci. Discuss.* 2016, 1–32. doi: 10.5194/hess-2016-585

WMO, (2008). *The Guide to Hydrological Practices (WMO No.168)*, 6th Edn. Geneva: World Meteorological Organization.

**Conflict of Interest Statement:** 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.

*Copyright © 2019 Fischer, Aemisegger, Graf, Sodemann and Seibert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

ORIGINAL RESEARCH
published: 26 November 2019
doi: 10.3389/feart.2019.00313

# Monitoring Atmospheric, Soil, and Dissolved $CO_2$ Using a Low-Cost, Arduino Monitoring Platform (CO2-LAMP): Theory, Fabrication, and Operation

*Joshua M. Blackstock[1]*, Matthew D. Covington[1], Matija Perne[2] and Joseph M. Myre[3]*

[1] Department of Geosciences, University of Arkansas, Fayetteville, AR, United States, [2] Department of Systems and Control, Jožef Stefan Institute, Ljubljana, Slovenia, [3] Computer and Information Sciences, University of St. Thomas, St. Paul, MN, United States

Variability of $CO_2$ concentrations within the Earth system occurs over a wide range of time and spatial scales. Resolving this variability and its drivers in terrestrial and aquatic environments ultimately requires high-resolution spatial and temporal monitoring; however, relatively high-cost gas analyzers and data loggers can present barriers in terms of cost and functionality. To overcome these barriers, we developed a low-cost Arduino monitoring platform (CO2-LAMP) for recording $CO_2$ variability in electronically harsh conditions: humid air, soil, and aquatic environments. A relatively inexpensive $CO_2$ gas analyzer was waterproofed using a semi-permeable, expanded polytetrafluoroethylene membrane. Using first principles, we derived a formulation of the theoretical operation and measurement of $PCO_{2(aq)}$ by infrared gas analyzers submerged in aquatic environments. This analysis revealed that an IRGA should be able to measure $PCO_{2(aq)}$ independent of corrections for hydrostatic pressure. CO2-LAMP theoretical operation and measurement were also verified by accompanying laboratory assessment measuring $PCO_{2(aq)}$ at multiple water depths. The monitoring platform was also deployed at two sites within the Springfield Plateau province in northwest Arkansas, USA: Blowing Springs Cave and the Savoy Experimental Watershed. At Blowing Springs Cave, the CO2-LAMP operated alongside a relatively greater-cost $CO_2$ monitoring platform. Over the monitoring period, measured values between the two systems covaried linearly ($r^2 = 0.97$ and 0.99 for cave air and cave stream dissolved $CO_2$, respectively). At the Savoy Experimental Watershed, measured soil $CO_2$ variability capturing sub-daily variation was consistent with previously documented studies in humid, temperate soils. Daily median values varied linearly with soil moisture content ($r^2 = 0.84$). Overall, the CO2-LAMP captured sub-daily variability of $CO_2$ in humid air, soil, and aquatic environments that, while out of the scope of the study, highlight both cyclical and complex $CO_2$ behavior. At present, long-term assessment of platform design is ongoing. Considering cost-savings, CO2-LAMP presents a working base design for continuous, accurate, low-power, and low-cost $CO_2$ monitoring for remote locations.

**Keywords: Arduino®, carbon dioxide, hydrology, soil carbon, karst, low-cost, critical zone**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology, Netherlands*

***Reviewed by:***
*Ryan D. Stewart,*
*Virginia Tech, United States*
*Sierra Young,*
*North Carolina State University, United States*

****Correspondence:***
*Joshua M. Blackstock*
*jmblack@uark.edu*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *23 March 2019*
***Accepted:*** *08 November 2019*
***Published:*** *26 November 2019*

***Citation:***
*Blackstock JM, Covington MD, Perne M and Myre JM (2019) Monitoring Atmospheric, Soil, and Dissolved $CO_2$ Using a Low-Cost, Arduino Monitoring Platform (CO2-LAMP): Theory, Fabrication, and Operation. Front. Earth Sci. 7:313. doi: 10.3389/feart.2019.00313*

## INTRODUCTION

Carbon exchange within the Earth system is facilitated, in part, by the production, transfer, and uptake of carbon dioxide (Schimel et al., 2001; Brantley et al., 2007). Unraveling biologic and biogeochemical (Broecker and Sanyal, 1998; Davidson et al., 2010; Demars et al., 2015; Florea, 2015), geologic (Lowenstern, 2001; Werner and Cardellini, 2006; Burton et al., 2013; Queiβer et al., 2016), and anthropogenic factors (Olah et al., 2011; Ward et al., 2015; Decina et al., 2016) that influence $CO_2$ concentrations require not only accurate, high-frequency measurements of $CO_2$ concentrations, but widely distributed, and if possible, spatially dense $CO_2$ measurements (Schimel et al., 2001; Hari et al., 2008; McDowell et al., 2008; Richter and Mobley, 2009; Brantley et al., 2016).

Long-term, high-frequency measurements of $CO_2$ concentrations are limited across Earth (McDowell et al., 2008; Andrews et al., 2014) compared to other continuous environmental monitoring in terrestrial and aquatic environments (e.g., air and stream temperature, air pressure, humidity, stream pH; Martin et al., 2017). In turn, the inter- and intra-seasonal variability of $CO_2$ and environmental factors controlling variability across terrestrial ecosystems remains poorly constrained (Serrano-Ortiz et al., 2010; Lombardozzi et al., 2015). Reducing these uncertainties in carbon transfers hinges upon increasing the spatial and temporal coverage of $CO_2$ measurements across the Earth system (Schimel et al., 2001, 2015; Lombardozzi et al., 2015; Bradford et al., 2016).

While the availability of commercial, field-deployable infrared gas analyzers (IRGA) have greatly enhanced measurement capacity, costs due to instrumentation acquisition, maintenance, and in some cases, limited storage capacity and control over measurement frequency using proprietary systems greatly limit the spatial and temporal extent of monitoring (Fisher and Gould, 2012; Martin et al., 2017). Furthermore, ancillary data, such as temperature (in air or water), are needed for environmental correction of $CO_2$ values, but combined sensor, and data logger selection may be limited between proprietary systems stemming from incompatibilities between manufacturers (Fisher and Gould, 2012).

Over the last decade, the availability and use of relatively inexpensive microcontrollers and "microcomputers" for scientific research has increased significantly (Cressey, 2017). Use of these platforms to interface sensors has grown, in part, from the increasing availability of sensors, and the need for customized interfacing to measure, and monitor conditions both in increasingly complex laboratory experiments, and challenging environmental settings (e.g., caves; Pearce, 2012; Beddows and Mallon, 2018). Low-cost $CO_2$ IRGAs (<$150 USD) and low-cost Arduino monitoring platforms (LAMPs) have been specifically used to measure and monitor dissolved $CO_2$ using automated floating chambers (Bastviken et al., 2015) and ambient $CO_2$ (Martin et al., 2017); however, adoption of a low-cost IRGA for electronically harsh conditions, such as high-humidity environments (e.g., caves) or within stream environments (i.e., submerged, direct dissolved $CO_2$ measurement), have been limited. If similar methods for waterproofing $CO_2$ sensors are used (Johnson et al., 2009), adoption of a low-cost IRGA to monitor $CO_2$ in electronically harsh environments should be possible.

We present a low-cost ($250–300 USD), Arduino-based monitoring platform (CO2-LAMP) for measuring atmospheric, soil, and dissolved $CO_2$ concentrations. Included in this study are methods for fabrication, reference measurement (i.e., zero and span reference gases), instrument value corrections and post-processing, and results from field-trial evaluations. As part of the reference measurement and post-processing, a novel presentation of theoretical sensor operation, sensor output, and accompanying empirical experiments were made to verify theoretical instrument output, and applicable environmental corrections. Consequently, the description and corrections are highly relevant to other direct, dissolved gas measurement systems by IRGAs (Johnson et al., 2009; Yoon et al., 2016). Field evaluation comprised: (1) A comparative field trial between the CO2-LAMP and a relatively greater-cost system for monitoring ambient $CO_2$ and dissolved $CO_2$, and (2) monitoring soil $CO_2$ in a shallow soil pit. Lastly, recommendations, and future work with respect to fabrication, improving measurement accuracy, and deployment of the CO2-LAMP are discussed.

## MEASUREMENT OF $CO_2$ IN EARTH'S NEAR-SURFACE ENVIRONMENT

Measurements of $CO_2$ within ambient air, soil, and aqueous environments encompass a range of sampling protocols and gas analyses. While not an exhaustive review, this section provides theoretical principles and practical aspects of measuring $CO_2$ in Earth's near-surface environment used in this study. Moreover, this brief overview presents information on discrete and continuous $CO_2$ measurement methods within air, soil, and aqueous environments with emphasis on the operating principles of direct dissolved $CO_2$ measurements using IRGAs within aqueous environments specifically investigated.

### Analysis of $CO_2$ in Air and Soils

Analysis of ambient $CO_2$ and soil $CO_2$ are routinely conducted by discrete sampling and *in situ* gas analyzers (Jassal et al., 2005; Andrews et al., 2014; Sánchez-Cañete et al., 2017; Jochheim et al., 2018). Discrete sampling is conducted primarily through gas collection into evacuated air-tight or "inert gas flushed" (e.g., helium flushing) containers. Extracted gases are subsequently sampled, typically using an IRGA, gas chromatography (GC), or isotope ratio mass spectrometer (Breecker and Sharp, 2008; Joos et al., 2008). Common *in situ* gas analyzers for measuring $CO_{2(soil)}$ have included the Vaisala GMD20, GMM221, and GMM222, and the Eosense eosGP (Hirano et al., 2003; Jassal et al., 2005; Sánchez-Cañete et al., 2017). Unlike discrete measurements, *in situ* sensors allow greater measurement frequency directly located within air and soil environments. However, *in situ* sensors require continual reference measurements to ensure accountability of sensor drift and offsets during deployment (Moran et al., 2010; Andrews et al., 2014). To further ensure measurement

accuracy through time, ancillary parameters, which include temperature, relative humidity, and atmospheric pressure, must also be measured to correct for differences between calibration, and field environmental conditions (e.g., pressure and temperature corrections; Fietzek et al., 2014). To protect against instrument damage in soil environments, protective membranes, such as silicone or polytetrafluoroethylene, are used to cover the sensor, but still allowing for gas exchange (Tang et al., 2003; Jassal et al., 2005).

## Obtaining Dissolved $CO_2$ Concentrations

Dissolved $CO_2$ concentrations are most often obtained through three common methods: (1) Estimation of $CO_2$ concentrations from alkalinity titration and carbonate species equilibria calculations (Stumm and Morgan, 1996; Abril et al., 2015; Jarvie et al., 2017); (2) Manual gas extraction from water sample collection in air-tight containers (e.g., copper tubing, manual headspace analysis; Sanford et al., 1996 and references therein); and (3) Directly measured through gas equilibration (Takahashi, 1961; Frankignoulle et al., 2001; Johnson et al., 2009; Yoon et al., 2016). The majority of dissolved $CO_2$ values reported for natural waters have been, to date, through carbonate equilibria calculations from measured pH, and total alkalinity (Abril et al., 2015; Liu and Raymond, 2018). However, reported partial pressures of dissolved $CO_2$, and corresponding dissolved $CO_2$ concentrations in organic-rich, low pH inland freshwaters are likely overestimated due to a combination of: (1) Greater total alkalinity derived from organic acid anions (e.g., greater dissolved organic carbon) and (2) Greater sensitivity of calculated dissolved $CO_2$ for low pH, low alkalinity waters vs. relatively higher pH, higher alkalinity waters (Abril et al., 2015). Importantly, Abril et al. (2015) highlight the critical need for direct measurements of $CO_2$ given the large uncertainty that may arise from carbonate equilibria estimations.

## Direct Measurement of Dissolved $CO_2$ Principles

Direct dissolved $CO_2$ measurement systems have been previously described by Yoon et al. (2016) and are separated into two categories of active-equilibration and passive-equilibration. The active-equilibration methods being: manual gas extraction; a spray-type equilibrator (Takahashi, 1961); and a marble-type equilibrator (Frankignoulle et al., 2001). In active-equilibration systems, an external power-source facilitates water-air equilibration by pumping external water through sprayers or marble media. Enclosed, internal air volumes are circulated through an IRGA. The passive method is referred to as a "membrane-enclosed sensor." Passive membrane-enclosed sensors work via diffusion and equilibration of gases across a liquid impermeable, but gas permeable, membrane (Sanford et al., 1996; Johnson et al., 2009).

Compared to spray-, and marble-type equilibrators, membrane-enclosed sensors are practical in harsher environments such as soil, and surface waters, which can be variably saturated or highly turbid, and prone to tubing clogging or instrument fouling. This method is also more useful in situations where power delivery is limited (e.g., caves). However, membrane-enclosed sensors have the drawback of longer equilibration times (>10 min), and therefore they may not fully capture short-term, large magnitude variation in surface waters (e.g., rapid mixing during storm events; Yoon et al., 2016).

Hybrid systems also exist, which interface with surrounding water through membrane mediated gas exchange (i.e., a membrane-enclosed equilibrator) but also internally circulate air for heating and thermal equilibrium (De Gregorio et al., 2011; Fietzek et al., 2014). To decrease equilibration time in membrane-enclosed systems, external pumps near the membrane move adjacent water to the membrane interface which limit expansion of a static-boundary layer (Manning et al., 2003; Fietzek et al., 2014).

For all direct-measurement systems, $CO_2$ measured by an IRGA or GC is the equivalent partial pressure of $CO_2$, $PCO_{2(aq)}$, in equilibrium with the dissolved $CO_2$ of the water in accordance with Henry's Law:

$$PCO_2 = KCO_{2(T,S,P)}C_i, \tag{1}$$

where $KCO_2$ is the Henry's Law constant for $CO_2$ at a given temperature, $T$, salinity, $S$, and pressure, $P$, and $C_i$ is the concentration of dissolved $CO_2$ in water (Colt, 2012). Dalton's Law states that the sum of partial pressures for all dissolved gas species are equal to the total dissolved gas pressure in the water, $P_{TDG}$:

$$P_{TDG} = PN_2 + PO_2 + PCO_2 + P_{other\ gases}. \tag{2}$$

For most shallow surface waters and unconfined groundwater systems $P_{TDG}$ is approximately equal to ambient atmospheric pressures (Manning et al., 2003; Gardner and Solomon, 2009). However, some notable exceptions include: (1) dam tailwaters (D'Aoust and Clark, 1980; Urban et al., 2008) and similar surface water conditions that promote entrainment of bubbles at greater depths where $P_{TDG}$ may be upwards of 1.3 times atmospheric pressure; (2) deep, confined groundwater systems (Gardner and Solomon, 2009; Ryan et al., 2015); and (3) deep, crater lake systems containing submarine gas vents at depth, such as Lakes Monoun, and Nyos in Cameroon (Kling et al., 1987; Kusakabe and Sano, 1992). In both confined groundwater and deep, lake gas vent systems, increased hydrostatic pressure allows for greater gas saturation (i.e., increased concentration). As such, $P_{TDG}$ values may be several times that of atmospheric pressure if waters are gas saturated at these greater hydrostatic pressures. In practice, dual measurement of total dissolved gas pressure and dissolved $CO_2$ are recommended in environments where $P_{TDG}$ is suspected to be higher than atmospheric pressure to account for greater dissolved concentrations (Ryan et al., 2015).

At abyssopelagic depths (> ~4,000 m) in marine systems, changes in Henry's constant due to hydrostatic pressure must also be taken in account when calculating expected $PCO_2$ for a given dissolved $CO_2$ concentration or vice versa (Enns et al., 1965; Hamme et al., 2015). Inland freshwater systems, however, do not encounter such depths. For example, Henry's Law constants for dissolved gas measurements at Lake Baikal (i.e., Earth's deepest lake at ~1600 m) would only be offset ~2.2% (Enns et al.,

1965; Hamme et al., 2015). Therefore, changes in Henry's Law constants with respect to hydrostatic pressure are negligible for relatively shallow water bodies.

## Membrane-Enclosed Equilibration Principles

Fluid movement across membranes occurs through convective mass transfer comprising diffusive and advective transport (Bergman et al., 2011; Kruczek, 2015). Diffusive gas exchange between an external environment (i.e., atmosphere, soil, or water) and a membrane-enclosed volume, or headspace containing an IRGA, has been previously described using a Solution-Diffusion model. In this Solution Diffusion model, gas exchange is driven by differences between the partial pressures of the external environment, $P_{env}$, and within the headspace, $P_{IRGA}$ (Bareer, 1939; Sanford et al., 1996; De Gregorio et al., 2005; Gardner and Solomon, 2009). From De Gregorio et al. (2005), assuming $P_{env}$ to be constant, the partial pressure of $CO_2$ in the headspace at some time, $t$, may be estimated by

$$P_{IRGA}(t) = P_{env} + (P_i - P_{env})\, e^{-\frac{K_p A}{Vh} t}, \tag{3}$$

where $P_i$ is the initial partial pressure of $CO_2$ in the headspace, $K_p$ is equal to the effective diffusivity of the gas through the environment-membrane boundary and the membrane material (Gardner and Solomon, 2009), $A$ is membrane surface area, and $h$ is membrane thickness.

Empirically, the exponential term can be calculated from experimental data using a modified form of Equation (3) whereby generalizing the exponential term, $K_pA/Vh$, as a constant $q$, and subsequently solving for $q$:

$$P_{IRGA}(t) = P_{env} + (P_i - P_{env})\, e^{-qt}. \tag{4}$$

If $K_p$ is unknown, but $A$, $V$, and $h$ are well-constrained, $K_p$ can be solved by rearranging the obtained $q$ constant:

$$K_p = \frac{qVh}{A}. \tag{5}$$

In the case of membrane submersion within water, diffusion of the gas within the water may have an important impact on transfer rates, rather than mass transfer being controlled by diffusion through the membrane alone. In this case, using a slight modification of Equation (4) to calculate the mass transfer coefficient, $k$, where $k = K_p/h$ may be more meaningful.

For description of percent equilibration of $CO_2$ to a reference gas, an exponential, or $e$-folding, timescale can be used to describe the amount of time over which changes in concentration or percent equilibration associated with an exponential process (i.e., gas equilibration in this case) occur by factors of $e \sim 2.718$. From measured $PCO_2$ using a waterproofed IRGA, $e$-folding time units, $T_f$ in seconds, can be expressed as

$$T_f = \frac{t}{\ln\left(\frac{P_{IRGA}(t)}{P_i}\right)} \tag{6}$$

where $t$ is equal to the time elapsed from the beginning of the observation period. To determine $n$ $e$-folding time, where $n$ is the folding time interval (e.g., three-folding times), $n$ is divided by the $q$ constant value, redefined here as $\tau$, obtained from the exponential function term (see Equation 4): e-folding time $= n/\tau$. For example, at three e-folding time (or $3/\tau$), equilibration of a mixture from the initial to final concentration is at ∼95%, i.e., $1–(1/e^3)$. In turn, solving for $3/\tau$ determines the specific $T_f$ equivalent to a measured value and actual time, t, where the partial pressure or concentration of $CO_2$ is 95% equilibrated.

# MATERIALS AND METHODS

## CO2-LAMP Fabrication for Humid and Aqueous Environments

Fabrication of CO2-LAMP consisted of waterproofing a relatively low-cost IRGA using a semi-permeable membrane (**Figure 1**) and interfacing the IRGA with an Arduino-based platform to read and record instrument values. The IRGAs used in this study were the K30 1 and 10% analyzers manufactured by Senseair AB (Delsbo, Sweden). Analyzer accuracies are reported by the manufacturer as ±30 parts per million by volume (ppmv) ± 3% for the K30 1% model and ±300 ppmv ±3% for the 10% model, respectively. The resolution of $CO_2$ concentrations reported by the K30 1 and 10% are 1.0 and 10.0 ppmv, respectively.

The membranes used were an expanded polytetrafluoroethylene ePTFE sleeve (Product number 200-07; International Polymer Engineering, Tempe, AZ, USA) and ePTFE gasket disc (Product number 1084N86, McMaster-Carr, Douglasville, GA, USA). Before enclosing the sensor, a serial cable was soldered to the K30 printed circuit boards (PCB) for interfacing the sensor with either Arduino microcontroller or desktop-PC. Then, the ePTFE membrane was placed over the K30 hydrophobic filter and attached to the K30 PCB by applying a small amount of Plasti Dip rubber compound (Plasti Dip International, Blaine, MN, USA). Subsequent coats of Plasti Dip were applied to create an effective seal at the contact of the membrane and the printed circuit board.

During coating steps, small holes in the rubber compound can form from degassing of the curing agent requiring multiple rubber compound coats. Small openings on the underside of the K30 PCB were then also filled with Plasti Dip. Importantly, a 1 h curing period was allowed between applying coats of the rubber compound. After application, to ensure a complete seal, a 24 h wait period was allotted allowing for a full cure of the rubber compound. A hole large enough for the serial cable was then drilled into a small plastic case and the K30 was placed inside the plastic case with the serial cable extending through the hole in the plastic case.

A small amount of Sugru silicone adhesive (FormFormForm Ltd., London, United Kingdom) was also used to horizontally level the K30. Subsequently, the K30 was then "potted" in Hysol 9460 epoxy (Henkel Corporation, Rocky Hill, CT, USA) just up to the point of covering the membrane. Lastly, a final rubber coating was applied at the contact between the epoxy and membrane and

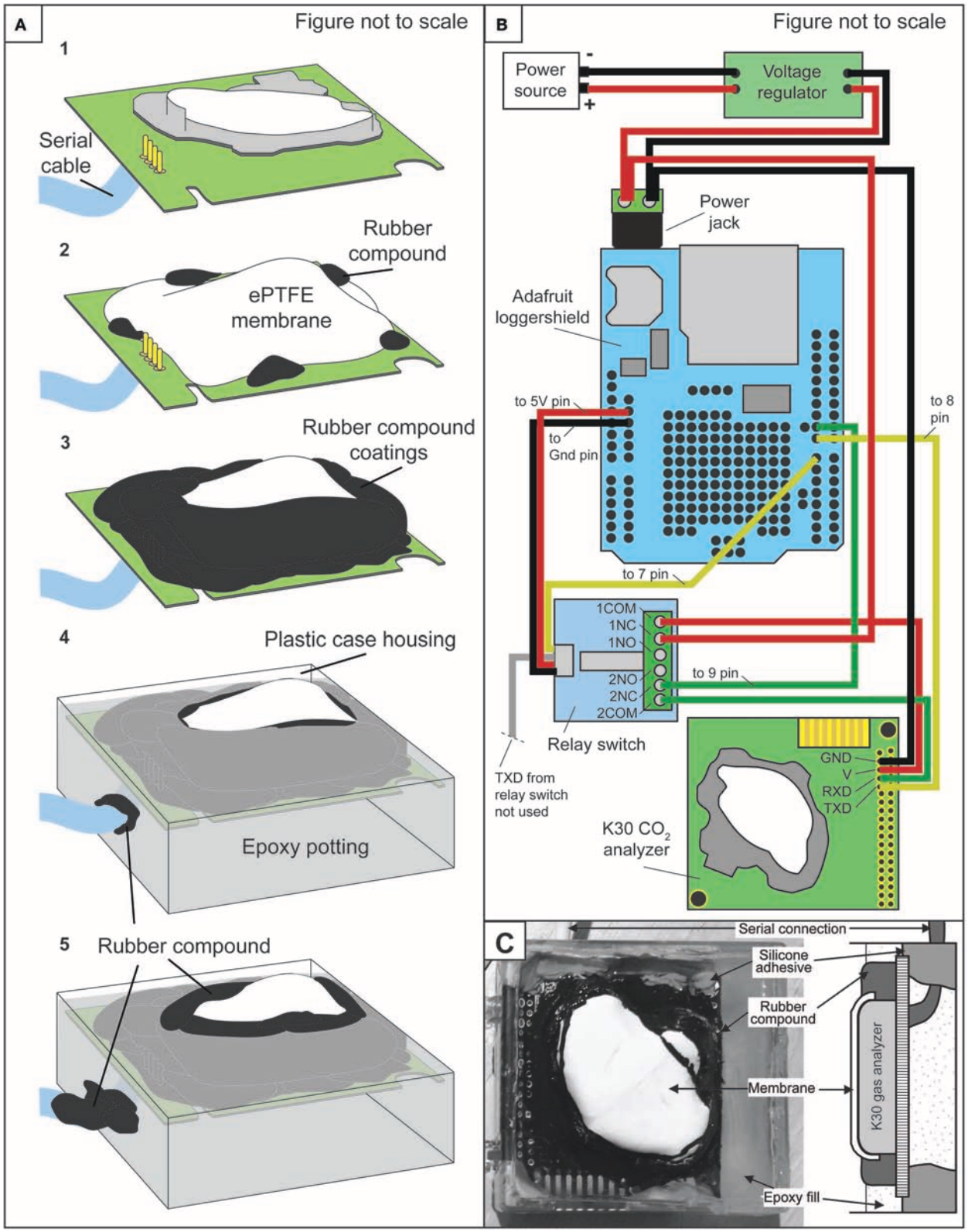



**FIGURE 1 | (A)** Simplified schematic of step-wise waterproofing of K30 sensor. **(B)** Schematic wiring diagram among the power source, voltage regulator, Arduino Uno, and Adafruit loggershield, relay switch, and K30 IRGA. **(C)** Labeled photograph of waterproofed K30 and corresponding IRGA components in cross section.

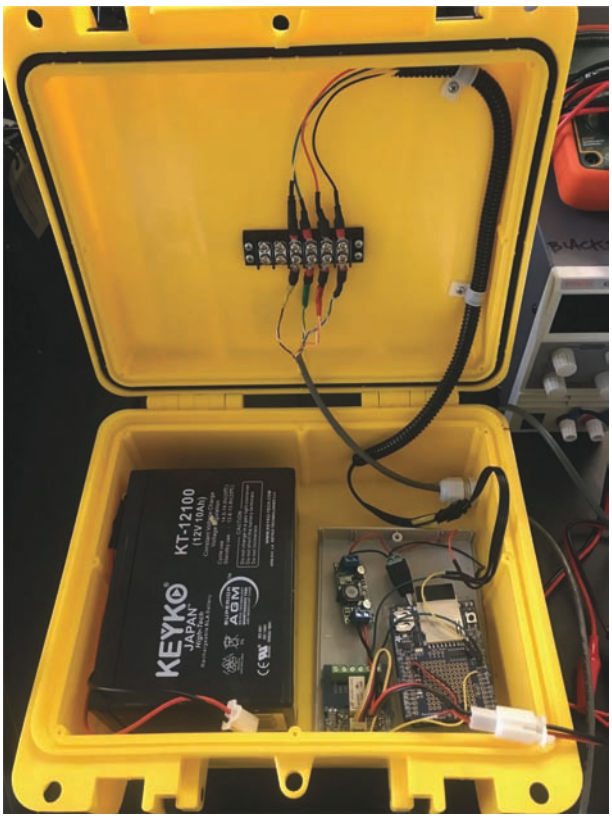



**FIGURE 2 |** Arrangement of 12 V battery power source, voltage regulator, Arduino-based data logger, relay switch, and terminal block connections that lead to a waterproofed K30. Terminal block was used to reduce physical strain and potential disruption to interior wired connections in the event the connecting cable to the waterproofed K30 is disturbed (e.g., external force pulling cable out of the box).

at the serial cable-epoxy contact (**Figure 1**). Membrane thickness and estimated area were ∼1 mm and 8 $cm^2$.

For the majority of lab experiments, respectively, and all field trials, the K30 was interfaced to an Arduino Uno (https://www.arduino.cc) with a connected Adafruit (New York, NY, USA) Data Logging shield using a universal asynchronous receiver/transmitter (UART) serial connection. During some laboratory trials, the K30s were instead interfaced via USB to a desktop computer where readings were read and logged using CO2Meter GasLab software (CO2Meter.com, Ormond Beach, FL, USA). Two Arduino sketches (i.e., programs) were written to interface the Uno and a power relay switch (Seeed Studio, Shenzhen, China) to control power delivery to the K30 in two modes: (1) a semi-continuous mode, where values were logged every 10 s for 60 s and then the sensor was powered off for 1 min before another measurement period began; and (2) a lower-power mode where values were logged every 10 s for 20 min, followed by a 45 min sleep period. Between measurement cycles the Uno was in "sleep" mode to reduce power consumption. In general, the low-power mode is advantageous in environments where direct power and battery recharge (e.g., solar panels) are not possible (e.g., caves).

Power was delivered to the K30 and Arduino Uno using regulated power supplies in the laboratory and 12 V batteries in the field (**Figure 2**). Between the power source and CO2-LAMP components, a step-down regulator was used to ensure a 6.5 V delivery to the Arduino and K30. While the K30 required only 5.5 V for operation, the additional voltage was applied to supplement for transmission loss given the length of the cable to the K30 (∼8 m). Measured K30 values were recorded on an SD Card using an Adafruit Assembled Data Logging shield for Arduino (Product 1141, Adafruit, New York, NY, USA).

## Zero and Span Reference Measurements

To initially verify K30 accuracy, span gas measurements were made using certified $CO_2$-Nitrogen balanced gas mixtures of 2,000 and 10,000 ppmv $CO_2$ (±2% analytical uncertainty) both in a dry, gas-filled chamber (**Figure 3A**) and partially water-filled chamber where the sensor was submerged (**Figure 3B**). For the dry reference measurements, a waterproofed K30 1 and 10% were placed in a dry, vented chamber while the reference gas mixture was continuously delivered to the chamber until equilibration with the reference gas was obtained. For submerged reference measurements, waterproofed K30 1 and 10% sensors were placed in a vented, partially water-filled chamber where reference gas mixtures were delivered to the chamber via a diffuser stone at the base of the chamber.

The water in the chamber was considered equilibrated to 95% once the CO2-LAMP readings reached the three e-folding time. The waterproofed K30 was then removed, allowed to re-equilibrate with the ambient laboratory air, and then re-submerged and allowed to reach the three e-folding time over three different submerged trials. Importantly, intervals for e-folding times were separately calculated for the individual submerged trials. Using Equations 4 and 5, values for $K_p$ were then calculated using an estimated volume of 5.6 $cm^3$. Hereafter, $PCO_{2(aq)}$ refers to laboratory measurement of the partial pressure of dissolved $CO_2$.

## Submerged IRGA Operation and Validation

Seminal work by Johnson et al. (2009) on the construction of a passive, permeable membrane equilibrator suggested a depth-correction for IRGA output to account for increased hydrostatic pressure acting on a submerged gas analyzer. However, gas exchange will occur across a membrane until such time that $PCO_2$ is equal between the water and the membrane enclosed volume, irrespective of changes in the enclosed headspace volume brought on by increased hydrostatic pressure, suggesting that such a depth correction is not needed. To address this discrepancy related to potential effects of increasing hydrostatic pressure on membrane-enclosed IRGA operation, we explore the theory behind $PCO_2$ calculation for a membrane-enclosed submerged IRGA and describe laboratory experiments that we use to test the derived principles.

### Submerged IRGA Output: Theoretical Principles

In air, concentrations of $CO_2$ are typically reported by IRGAs as volumetric fractions, $x_c$, of $CO_2$ in dimensionless units either as parts per million volume (ppmv) or percent values for

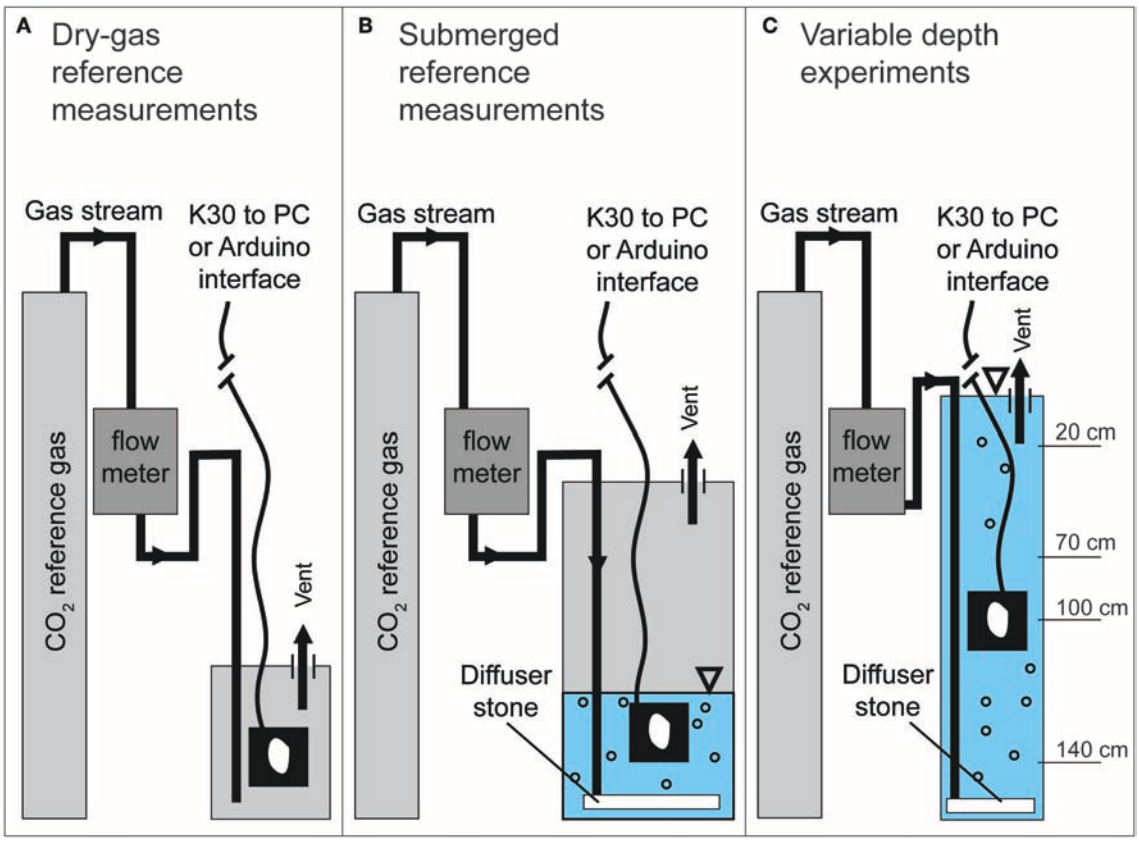



**FIGURE 3 |** Schematic of reference measurement configurations for **(A)** dry-gas, **(B)** submerged, $PCO_2$, and **(C)** variable depth experiments. Not to scale.

greater concentrations (>10,000 ppm or 1%) where $x_c$ may be expressed as

$$x_c = \frac{V_i}{V_{total}}, \tag{7}$$

where $V_i$ equals the volume of $CO_2$ per total volume of gas, $V_{total}$. Alternatively, $CO_2$ in air may also be expressed as a partial pressure, $PCO_2$, from the product of $x_i$ and total pressure (or sum of partial pressures, i.e., Dalton's Law), $P_{total}$:

$$PCO_2 = x_c P_{total}. \tag{8}$$

While $P_{total}$ can be directly measured, or assumed to be near standard pressure, IRGAs do not directly measure $x_c$.

Principally, an IRGA measures the molecular density of $CO_2$ using the Beer-Lambert Law through the measured absorbance of $CO_2$ for a given wavelength (Fietzek et al., 2014). Molecular density, $\rho$, is expressed as $\rho = NCO_2/V_{total}$ where $NCO_2$ is the number of $CO_2$ molecules.

The $x_c$ value from an IRGA is obtained using the ideal gas law, with

$$PCO_2 V_{total} = \frac{NCO_2}{N_A} RT, \tag{9}$$

$$PCO_2 = \frac{NCO_2}{V_{total}} \frac{RT}{N_A} = \rho \frac{RT}{N_A}, \text{and} \tag{10}$$

$$x_c = \frac{PCO_2}{P_{total}} = \rho \frac{RT}{N_A P_{total}}, \tag{11}$$

where $R$ is the universal gas constant, $T$ is temperature in Kelvin, and $N_A$ is Avogadro's number. From Equation 11, $x_c$ values depend on $\rho$, $T$, and $P_{total}$. If $T$ and $P_{total}$ are not measured, factory calibrated values for temperature, $T_0$, and pressure, $P_0$, are used to calculate a "reported" volume fraction, $x_r$, which is expressed as

$$x_r = \rho \frac{RT_0}{N_A P_0}. \tag{12}$$

For the majority of low-cost $CO_2$ gas analyzers, where $T$ and $P_{total}$ are not measured simultaneously, IRGA output will generally follow Equation (12), where $T_0$ and $P_0$ are at or near 25°C and 1 atm, respectively. If $T$ and $P_{total}$ are measured a corrected volume fraction $x_c$, can be calculated, with

$$x_c = x_r \frac{T}{T_0} \frac{P_0}{P_{total}}. \tag{13}$$

While the correction in Equation (13) is routinely employed for measurements of $CO_2$ concentrations in ambient air and soil,

dissolved $CO_2$ concentrations are most commonly calculated from $PCO_2$, not a volumetric fraction. From Equation (10), $PCO_2$ can be calculated directly from molecular density, $\rho$, temperature, and known constants. However, IRGA output using factory calibrated temperature, and pressure is $x_r$. To determine $PCO_2$ from $x_r$, Equation (12) is solved for $\rho$, and substituted into Equation (13), giving

$$PCO_2 = \frac{x_r N_A P_0}{RT_0}\frac{RT}{N_A} = \frac{x_r P_0 T}{T_0}. \quad (14)$$

Note that calculation of $PCO_2$ from the reported volumetric fraction only requires the calibration pressure (typically ∼1 atm), not the pressure during measurement. On the other hand, a temperature correction is needed if temperature during measurement is substantially different calibration conditions.

Equation (14) demonstrates with introduction of sensor operating principles, total pressure factors out of the calculation of $PCO_2$. Therefore, for a well-mixed, relatively shallow water body of equal temperature, salinity, and dissolved gas concentrations, the partial pressure of $CO_2$ measured by an IRGA at equilibrium (i.e., no gas exchange across the membrane) should be equal at all depths irrespective of hydrostatic pressure. Combining Equations (1) and (14), the concentration of dissolved $CO_2$ determined from direct, membrane equilibration methods using an IRGA can be expressed as

$$C_i = \frac{PCO_2}{KCO_{2(T,S,P)}} = \frac{x_r P_0 T}{KCO_{2(T,S,P)} T_0}. \quad (15)$$

While the theoretical results suggest that no depth correction is needed for calculation of $PCO_2$, if a sensor is suddenly lowered to greater depths, compression of the membrane or sensor housing may introduce increases in total gas pressure within the IRGA. This will produce a short-term spike in the pressures of all gases, including $CO_2$. However, this produces disequilibrium between the gas pressures within the water, and the IRGA which will drive exchange across the membrane until dissolved gas pressure in the water, and gas pressure in the IRGA are back in equilibrium.

### Variable Water Depth Experiments: Laboratory Simulation

An accompanying depth compensation experiment measuring $CO_2$ at multiple depth intervals (**Figure 3C**) was conducted to observe if $PCO_{2(aq)}$ values varied with submerged depth. A 7.62 cm PVC pipe, 152.5 cm in length was filled with water (i.e., synthetic well) to accommodate varying depth interval measurements. The gas mixture was delivered via a porous stone at the bottom of the well. Initially, the submerged K30 10% recorded $PCO_{2(aq)}$ values as the reference gas was delivered to the water in to confirm the $PCO_{2(aq)}$ of the water in the PVC tube had equilibrated with the reference gas (same method described in section Zero and Span Reference Measurements). Once the water in the PVC tube had equilibrated to the reference gas, the K30 10% was removed from the well and allowed to re-equilibrate with laboratory atmospheric $CO_2$ concentrations, which was assumed to be ∼500–600 ppmv.

The K30 was then quickly submerged to an initial depth of 20 cm and allowed to re-equilibrate with the $PCO_2(aq)$ imposed with the reference gas. Once equilibrated, the K30 was then dropped quickly from the 20 to 70 cm depth and allowed to re-equilibrate. This process was further repeated for depth intervals of 100 and 140 cm. During the experiment, equilibration was assumed to be reached once values were both within the analytical uncertainty of the reference gas, and reading variability was equal to the K30 10% reading resolution of 10 ppm for at least 10 min. Values for three e-folding time were estimated, however, after the experiments.

## Field Trials

Field trials were carried out at Blowing Springs Cave and the Savoy Experimental Watershed located in Northwest Arkansas, USA (**Figure 4**). The two sites represent karst environments within the Springfield Plateau physiographic province overlying the Springfield Plateau aquifer (Kresse et al., 2014). The Springfield Plateau province can be characterized as a mantled karst terrain consisting of a cherty regolith overlying the Boone Formation, a cave forming Paleozoic carbonate unit (Brahana et al., 1999; Knierim et al., 2013; Al-Qinna et al., 2014; Jarvie et al., 2014).

### Blowing Springs Cave

At Blowing Springs Cave, both cave air, $CO_{2(air)}$, and dissolved $CO_2$ within the cave stream, $PCO_{2(stream)}$ were measured independently by: (1) the CO2-LAMP, and (2) an enclosed membrane-equilibrator similar to Johnson et al. (2009), hereafter referred to as the "Vaisala system." Sensors were located ∼100 m within the cave. In the CO2-LAMP platform, concentrations of $CO_2$ for $CO_{2(air)}$, and $PCO_{2(stream)}$ were measured by a waterproofed K30 1, and 10%, respectively. For the Vaisala system, $CO_{2(air)}$, and $PCO_{2(stream)}$ were measured using a waterproofed (see Johnson et al., 2009), Vaisala GMT220 (Helsinki, Finland), and logged using a Campbell Scientific (Logan, UT) CR850. Cave air temperature and cave air pressure were measured using a Campbell Scientific HC2S3 and CS106, respectively. Cave stream temperature was recorded using Cave air direction and speed were recorded using a Campbell Scientific WINDSONIC1-L sonic wind sensor. Cave stream temperature was measured using a Campbell Scientific CS547A-L. Cave air temperature, cave air pressure, cave air flow direction and speed, and cave stream temperature were logged using the Campbell Scientific CR850. For $CO_{2(air)}$ and $PCO_{2(stream)}$ monitoring locations, waterproofed Vaisala CO2 IRGAs, and K30 IRGA sensors were placed alongside each other. Monitoring using the CO2-LAMP lasted from 25 February to 9 March, 2017.

Percent differences between the Vaisala and CO2-LAMP were calculated for measurements of $CO_{2(air)}$ and $PCO_{2(stream)}$, respectively:

$$\% = 100 \times \frac{\left|CO_{2,\ CO2LAMP} - CO_{2,\ Vaisala}\right|}{\frac{\left(CO_{2,\ CO2LAMP} + CO_{2,\ Vaisala}\right)}{2}}. \quad (16)$$

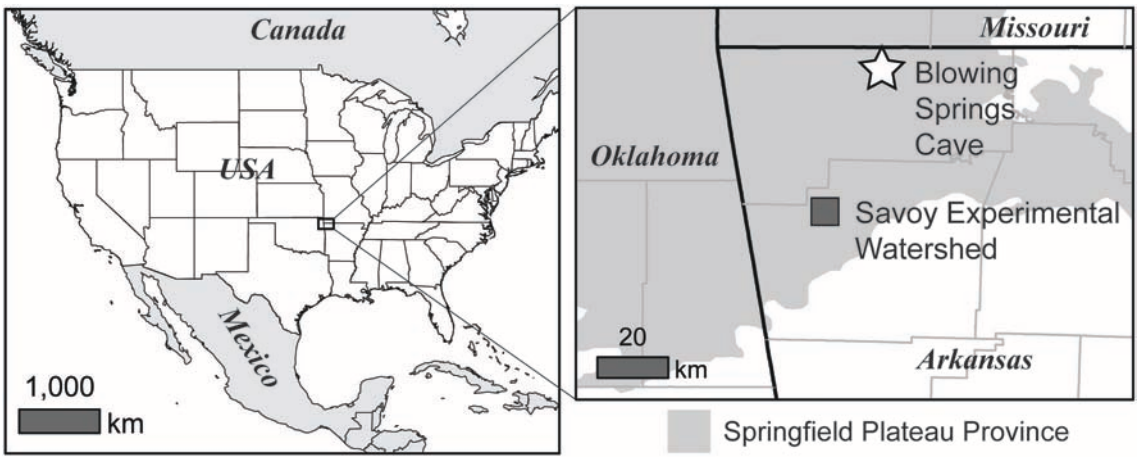



**FIGURE 4 |** Location of Blowing Springs Cave and the Savoy Experimental Watershed located within the Springfield Plateau physiographic province in northwestern Arkansas, USA.

## Savoy Experimental Watershed

The Savoy Experimental Watershed (SEW) is a long-term experimental research station owned by the University of Arkansas encompassing numerous karst features including sinking streams, caves, cave springs, and epikarst springs (Brahana et al., 1999; Al-Qinna et al., 2014; Covington and Vaughn, 2018). Soil series at SEW have been previously classified as Clarksville (Loamy-skeletal, siliceous, semiactive, mesic Typic Paleudults), Nixa (Loamy-skeletal, siliceous, active, mesic Glossic Fragiudults), Razort (Fine-loamy, mixed, active, mesic Mollic Hapludalfs), and Pickwick (Fine-silty, mixed, semiactive, thermic Typic Paleudults; Soil Survey Staff, 2019). Soils consist of very deep, moderately to excessively drained, slow to moderately permeable soils with clay contents ranging from 20 to 50% (Soil Survey Staff, 2019).

Soil $CO_2$ concentrations at the SEW are reported for the period of 9–22 July, 2017 and were measured ∼2 m from a centrally located weather station. Concentrations of $CO_{2(soil)}$ were measured using a waterproofed K30 10% at ∼10 cm depth within a soil cavity with the dimensions of ∼10 cm depth and 4 cm diameter. A small opening was dug into the wall of the soil cavity where the sensor was placed laterally in the base of the cavity wall. The soil cavity was back-filled as to minimize soil disturbance. Unlike at Blowing Springs, a greater accuracy $CO_2$ gas analyzer system was not co-deployed while the CO2-LAMP was deployed. Considering the Vaisala system (or similar) as a field reference measurement, assessment of absolute accuracies were not possible. However, relative magnitudes of daily $CO_2$ variability were compared to previous studies in a humid-temperate environment (Hirano et al., 2003). At the weather station, measurements of air temperature, soil moisture, and rainfall were recorded every 5 min.

## Post-processing Field Data

During field deployments, the low-power mode Arduino sketch was used to record measurements. As mentioned previously, $CO_2$ concentrations were logged every 10 s for 20 min, followed by a 40 min sleep period. Post-processing consisted of removing data during warm-up and stabilization periods and then extracting the final, stabilized values (**Figures 5A–D**). Final values measured during measurement cycles at Blowing Springs for cave air and dissolved $CO_2$ and soil $CO_2$ at SEW are reported here.

At Blowing Springs, stabilization periods for the sensor during warm-up changed through the monitoring period (**Figure 6**). Using a heuristic approach, the Hill-equation (Hill, 1910)—a non-linear, four-parameter equation—was fit to data collected during the monitoring after 100 s to evaluate changes in stabilization times over the monitoring period. Fitting to the data after 100 s minimized influence of the initial $CO_2$ peak (**Figure 5A**). In general, the Hill equation is useful in describing experimental data that are sigmoid in shape where multiple non-linear processes may be present (Goutelle et al., 2008; Gadagkar and Call, 2015).

The formulation of the Hill equation used in this study was

$$y = d + \left( \frac{a - d}{1 + \left(\frac{b}{t}\right)^c} \right), \tag{17}$$

where the coefficients calculated for this study were: $d$, the initial $CO_2$ value; $a$ is the final $CO_2$ value; $b$ is the time at which the $PCO_2$ value has changed halfway between $a$ and $d$; $c$, the "Hill Slope" or "steepness" value (Gadagkar and Call, 2015); and $t$ is the time elapsed during the measurement period. Calculated coefficients for curve steepness, $c$, were analyzed.

At Blowing Springs, measurement timestamps between the CO2-LAMP, and Vaisala system (which included cave air temperature, cave air pressure, and cave stream temperature) were variably offset because of different logging intervals. For the CO2-LAMP, the sum duration of time spanning the two-cycle operation (i.e., the "sleep" mode and measurement period) was 65 min with the two-cycle operation beginning as soon as

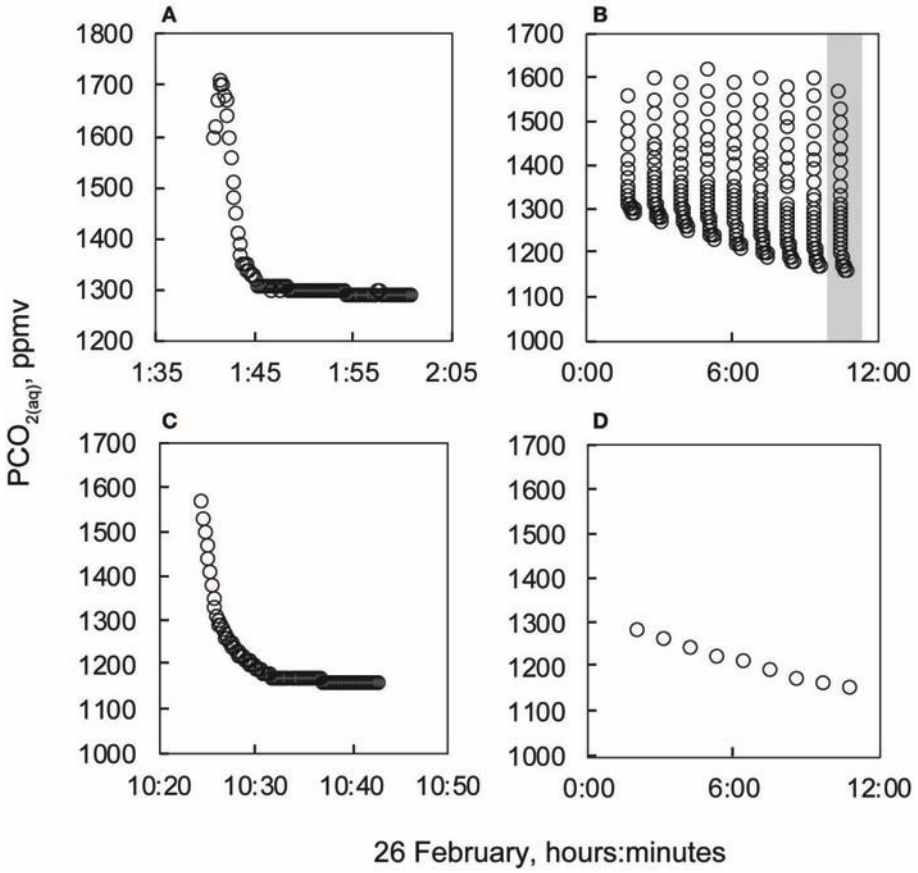



**FIGURE 5 | (A–D)** Post-processing steps for $PCO_{2(stream)}$ data. **(A)** Example of full measurement cycle data, which includes initial peak in first 100 s. **(B)** Example of initial filtering removing first 100 s of measurement cycles for nine measurement cycles on 26 February. **(C)** The last measurement value during each measurement cycle was selected as the "stabilized" value for comparison with the interpolated Vaisala data. The measurement cycle data comprise the gray-shaded area in **(B)**. **(D)** Example of filtered from data set in **(B)**.

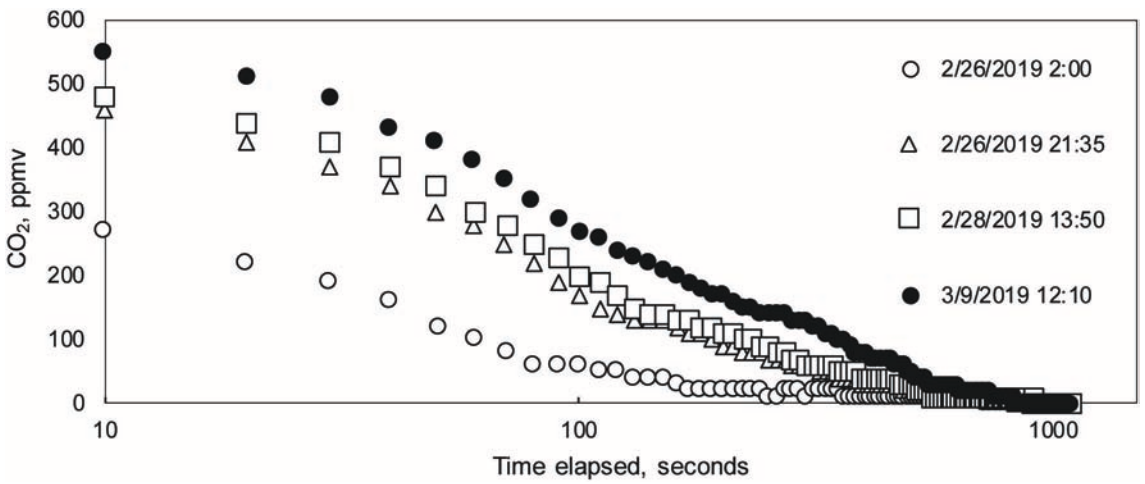



**FIGURE 6 |** Progressive change in $PCO_{2(stream)}$ stabilization curves through monitoring period.

the platform is powered. The Vaisala system was programmed to also include a two-cycle operation, however, the total time duration was 60 min. To address the variable temporal offset, values of cave air temperature, cave air pressure, cave stream temperature, Vaisala $CO_{2(air)}$, and Vaisala $PCO_{2(stream)}$ were linearly interpolated to match CO2-LAMP time stamps to

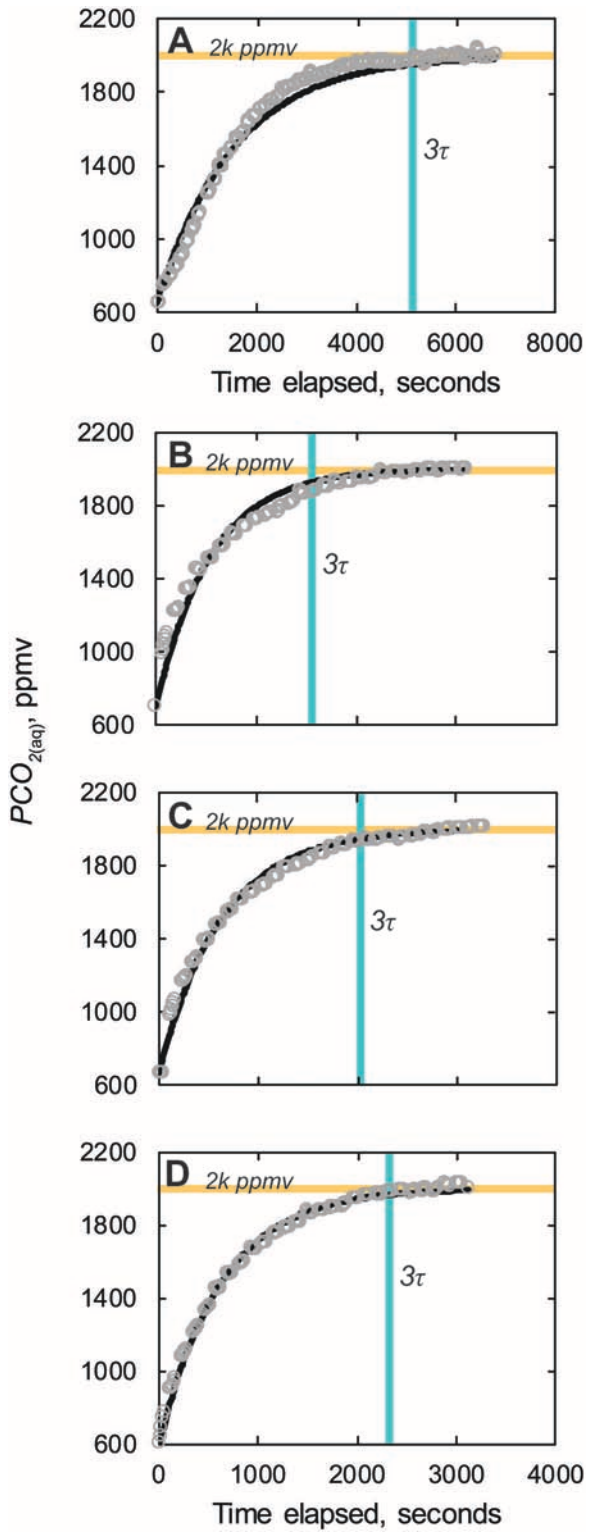



**FIGURE 7 | (A)** $PCO_{2(aq)}$ measured during injection of 2,000 ppmv $CO_2$ reference gas. **(B–D)** $PCO_{2(aq)}$ measured re-equilibration after equilibration of water to reference gas. Modeled exponential fits are solid black lines. Horizontal orange line, 2k ppmv, denotes 2,000 reference $PCO_{2(aq)}$ value. Vertical blue line indicated three e-folding time, $3\tau$, or ~95% equilibration. Measurements made using the K30 1% gas analyzer.

the nearest second. As the inter-hourly variability of $CO_{2(air)}$, $PCO_{2(stream)}$, cave air temperature, cave air pressure, and cave stream temperature were relatively low at Blowing Springs during the monitoring period, differences between true, and interpolated Vaisala values are likely small. In turn, CO2-LAMP data are directly compared to linearly interpolated Vaisala values of $CO_{2(air)}$, $PCO_{2(stream)}$, cave air temperature, cave air pressure, and cave stream temperature. For the following sections, reference to values of cave air temperature, cave air pressure, cave stream temperature, Vaisala $CO_{2(air)}$, and Vaisala $PCO_{2(stream)}$ refer to the linearly interpolated values. As cave air flow direction and speed were not directly compared to CO2-LAMP data, these values were not linearly interpolated. When cave air flow reversals were present, cave air flow was from the interior of the the cave toward the south entrance (or exiting the cave). Cave air flow reversals was defined when cave air flow direction $>100°$ (Young, 2018; Covington et al., in prep.).

CO2-LAMP $CO_{2(air)}$ and Vaisala $CO_{2(air)}$ data were corrected using ancillary pressure and temperature measurements made of cave air temperature using Equation (13). Values for CO2-LAMP $PCO2_{(stream)}$ and Vaisala $PCO_{2(stream)}$ were corrected using only water temperature data (Equation 14).

## Parameter Estimation and Regression Analysis

The constants $q$ (see Equation 4) and $c$ (see Equation 17) were estimated using EXCEL Solver (Microsoft, Redmond, WA, USA) applying a least-sum-square error procedure, which uses the Generalized Reduced Gradient method (Gadagkar and Call, 2015). Bivariate relationships were assessed by ordinary least squares linear regression using PAST version 3.25 (Hammer et al., 2001; Hammer, 2019).

# RESULTS

## Reference Measurements to Known Gas Mixtures

Gas equilibrated reference measurements of $CO_2$ and $PCO_2$ using the CO2-LAMP were within the accuracy stated by the manufacturer for the K30 1 and 10% IRGAS, respectively, in both dry, and aqueous environments. To begin the aqueous (or submerged) reference gas mixture experiments, tap water from the laboratory was equilibrated with the reference by delivering the gas mixture to the water using the diffuser stone. Considering the initial starting time as when the gas flow from the cylinder to the water began, the time needed for the water to reach three e-folding intervals (or 95% equilibration) was ~86 min for a volume of ~2.5 L (**Figure 7A**). This duration of time encompasses both diffusion of $CO_2$ into the water and the subsequent exchange of $CO_2$ across the membrane of the submerged waterproofed K30. Once the measurements read by waterproofed K30 reached the three e-folding time for the given reference gas mixture, the waterproofed K30 was removed from the water and allowed to re-equilibrate with laboratory ambient air. At this stage, the dissolved $PCO_{2(aq)}$ of the water in the

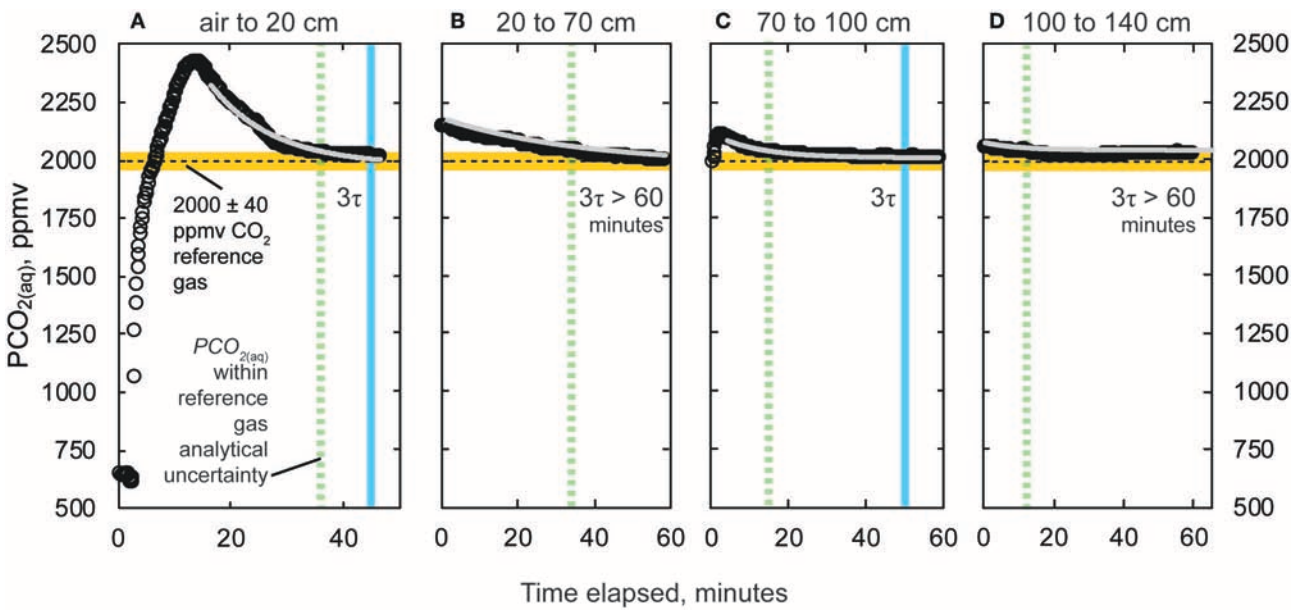



**FIGURE 8 |** Change in $PCO_{2(aq)}$ and time elapsed lowering from **(A)** air to 20 cm, **(B)** 20 to 70 cm, **(C)** 70 to 100 cm, and **(D)** 100 to 140 cm depth intervals. Recorded values missed initial peaks when submerged for **(B)** (20 to 70 cm) and **(D)** (100 to 140 cm) due to lagged start time on data logger. Modeled equilibration and three e-folding times are shown for **(A)** (20 cm) and **(C)** (100 cm). Estimated three e-folding times were >60 min on **(B)** (70 cm) and **(D)** (140 cm). However, the time is denoted in all panels when values are within analytical uncertainty of the reference gas mixture (hatched green line). Measurements made using the K30 10% gas analyzer.

wet chamber was considered equilibrated with the reference gas mixture.

The K30s were then re-submerged three separate times for a minimum period to reach three e-folding times in the reference gas equilibrated water volume (**Figures 7B–D**). The times needed to reach 95% equilibration were 27, 33, and 38 min for three reference experiments, respectively. The average effective $K_p$ value calculated was $1.2 \times 10^{-4}$ $cm^2/s$, which while nearly two orders of magnitude lower than $CO_2$ diffusivity through ePTFE from air-to-air environments (0.01 $cm^2$ $s^{-1}$; Johnson et al., 2009), was nearly an order of magnitude greater than the diffusivity of $CO_2$ in water ($1.77 \times 10^{-5}$ $cm^2$ $s^{-1}$ at 20°C; Scott, 2000). Final $PCO_{2(aq)}$ values were all within the analytical uncertainty of the reference gas composition 2,000 ppm ± 2% ppmv $CO_2$ (or 2,000 ± 40 ppmv $CO_2$).

## Variable Depth Trials

At all depths intervals, $PCO_{2(aq)}$ values during the final 10 min of data logging were: (1) within the analytical uncertainty of the 2,000 ± 2% ppmv $CO_2$ reference gas (or 2,000 ± 40 ppmv $CO_2$); and (2) did not vary more than K30 10% resolution of 10 ppm (**Figure 8**). At depths 20, 70, 100, and 140 cm, final stabilized $PCO_{2(aq)}$ values and the percent difference (%) with respect to the reference gas value of 2,000 ppm $CO_2$ were 2,020 (1%), 2,000 (0%), 2,010 (0.5%), and 2,030 (1.5%), respectively. As predicted, there were repeated patterns of an initial sharp increase in $PCO_2$ followed by a decline to imposed $PCO_2$ values upon rapid lowering of the K30 10% to greater depth. At depths 20, 70, 100, and 140, $PCO_{2(aq)}$ values were within analytical uncertainty of the reference gas after 36.7, 36.3, 16.8, and 12.8 min. Three e-folding times, $3\tau$, calculated after the experiments were 30, 72, 50, and 124 min for the respective 20, 70, 100, and 140 cm depths.

## Blowing Springs Cave $CO_{2(air)}$ and $CO_{2(stream)}$

During the field test, multiple periods occurred when cave air flow reversed whereby cave air exited through the southern entrance (**Figure 9A**). Increases in $CO_{2(air)}$ concentrations up to 749 ppm $CO_2$ (as recorded by the Vaisala system) were observed when cave air flowed toward the southern entrance (**Figure 9B**). Increases in $CO_{2(air)}$ were generally followed by periods of increased $PCO_{2(stream)}$ values (**Figure 9C**). However, broader peaks of $PCO_{2(stream)}$ (i.e., 2–3 and 6–9 March) lagged behind peaks in $CO_{2(air)}$ associated with the cave air flow reversals. Excluding $CO_{2(air)}$ during cave air reversals, $CO_{2(air)}$ concentrations ($n = 220$ measurements) were relatively constant with a mean of 472 ± 2 ppm (mean ± standard error). However, $PCO_{2(stream)}$ increased, overall, during the monitoring period from an initial value of 1,276–1,318 ppm $CO_2$ (as recorded by the Vaisala system for both $CO_{2(air)}$ and $PCO_{2(stream)}$).

Percent and ppmv differences between the Vaisala and CO2-LAMP for $CO_{2(air)}$ ranged from 2.1 to 20.9% and 13 to 147 ppmv, respectively (**Figure 9C**). Percent and ppmv differences between the Vaisala and CO2-LAMP for $PCO_{2(stream)}$ ranged from 1.3 to 11.9% and 16 to 147 ppmv, respectively, and exhibited a slight overall increase in percent difference during deployment (**Figure 9D**). Median percent and ppmv differences between

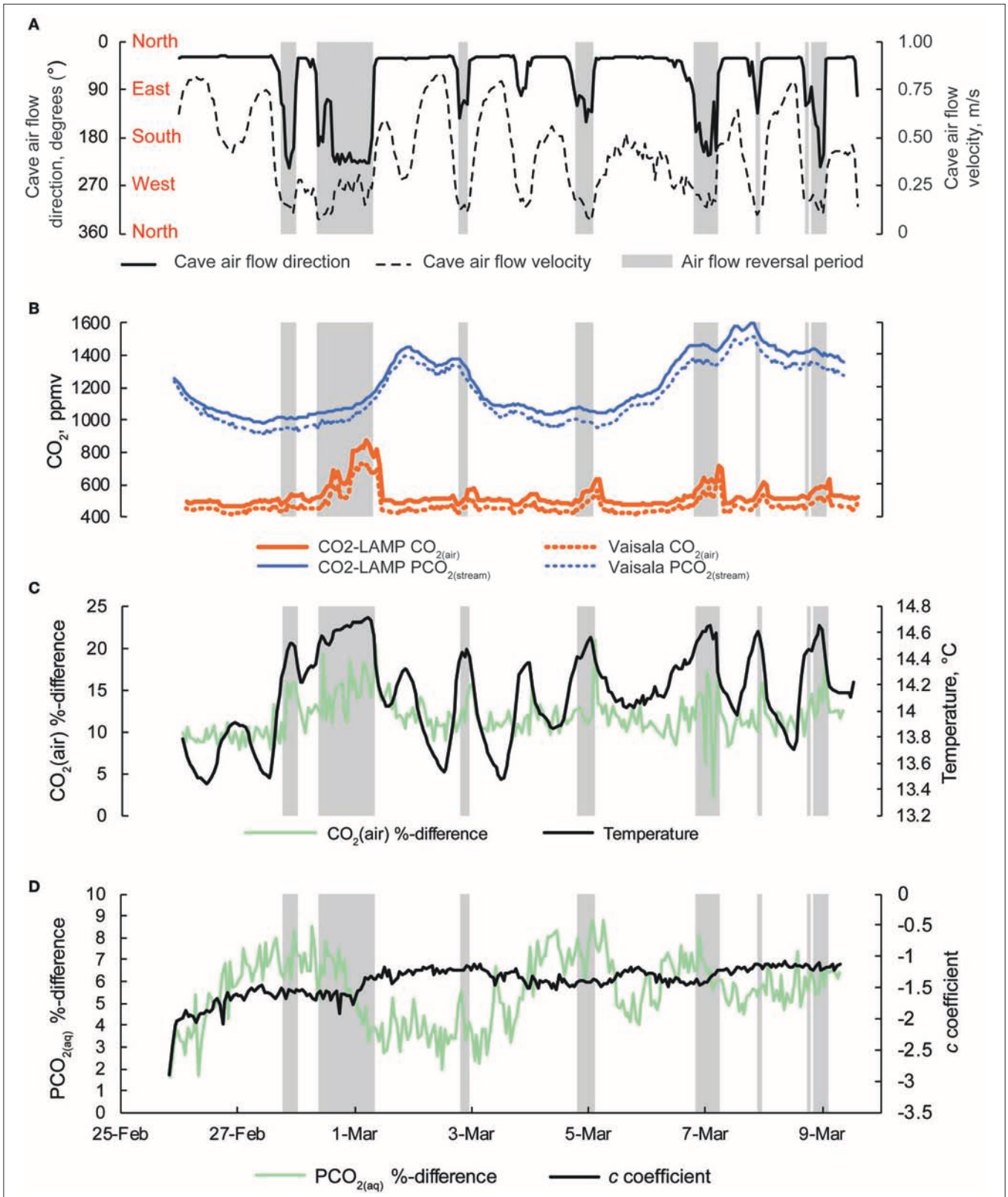



**FIGURE 9 | (A)** Cave air flow direction and cave air flow velocity. Cardinal directions are shown. Periods when cave air flow direction are >100° are shaded gray in all panels. **(B)** Concurrent measurements of $CO_{2(air)}$ and $PCO_{2(stream)}$ collected by the CO2-LAMP and Vaisala platforms from 26 February to 9 March, 2017. **(C)** Percent difference for $CO_{2(air)}$ between the CO2-LAMP and Vaisala and changes in cave air temperature over the monitoring period. **(D)** Percent differences for $PCO_{2(stream)}$ between the CO2-LAMP and Vaisala and changes in curvature (i.e., *c* coefficient) derived from the Hill-equation.

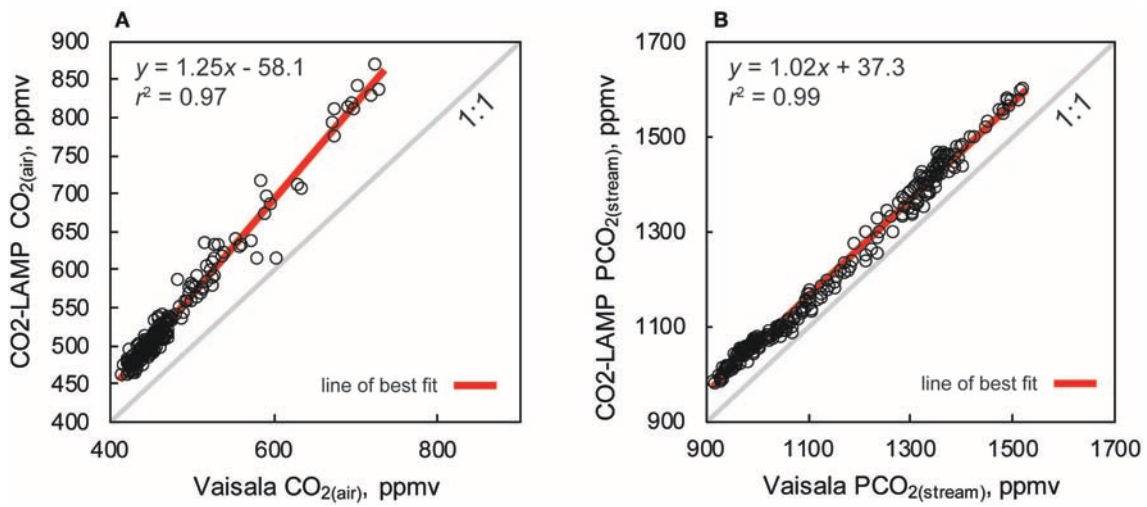



**FIGURE 10 |** CO2-LAMP vs. Vaisala measurements for **(A)** $CO_{2(air)}$ and **(B)** Vaisala $PCO_{2(stream)}$.

$CO_{2(air)}$ and $PCO_{2(stream)}$ were 11.6% and 56 ppmv and 8.1% and 92 ppmv, respectively. Values for $CO_{2(air)}$ measured using a K30 1% were often outside the manufacturer absolute accuracy ±30 ppmv ±3% stated for the K30. Values for $PCO_{2(stream)}$ measured using the K30 10%, however, were within the stated absolute accuracy of ±300 ppmv ±3%.

Measurements of $CO_{2(air)}$ and $PCO_{2(stream)}$ between the Vaisala and CO2-LAMP measurements did not appear to vary randomly during the monitoring period. The largest differences between $CO_{2(air)}$ values for the two instruments were observed during temperature peaks and coincided with cave air flow reversals. Differences in $PCO_{2(stream)}$ between the Vaisala and CO2-LAMP appeared to exhibit a quasi-oscillatory behavior and some covariation was observed between measurement differences and curvature (or $c$ coefficient) values calculated from the Hill-equation fits to the equilibration curves for the CO2-LAMP. Overall, measurements of $CO_{2(air)}$ ($r^2 = 0.97$, $p < 0.01$) and $PCO_{2(stream)}$ ($r^2 = 0.99$, $p < 0.01$) between the Vaisala and CO2-LAMP platform were well-correlated during the monitoring period (**Figure 10**).

## Savoy Experimental Watershed $CO_{2(soil)}$

Measurements of $CO_{2(soil)}$ at SEW exhibited both diurnal variation and an overall decline during the monitoring period (**Figure 11**). The daily amplitude of $CO_2$ variation ranged from 1,170 to 5,460 ppm with daily minimum and maximum values of $CO_{2(soil)}$ observed at approximately mid-night and mid-day (local time), respectively. Similar timing of minimum and maximum $CO_{2(soil)}$ values were also reported by Hirano et al. (2003). During the monitoring period, a light rain event occurred on 14 July evident from small rainfall totals and reduced daily temperatures, but no change in soil moisture was observed. However, $CO_{2(soil)}$ values decreased over 7,000 ppm from 14 to 15 July, increasing into 16 July, and subsequently decreasing over the remainder of the monitoring period. Overall, daily median $CO_{2(soil)}$ values were well-correlated with daily median soil moisture values ($r^2 = 0.84$; $p < 0.01$).

# DISCUSSION

## Measurement Accuracy and Assessment

Laboratory reference experiments using known $CO_2$ concentrations and imposing $PCO_{2(aq)}$ values in a volume of water demonstrated the viability of a K30 sensor for accurate, direct measurement of $PCO_{2(aq)}$ with equilibration times of 27–38 min. Compared to other commercial and non-commercial membrane-equilibration systems similar to Johnson et al. (2009) (<30 min), observed equilibration times in this study were slower most likely due to smaller membrane surface area to enclosed membrane volume ratios.

From both the submerged reference experiments (**Figure 7**) and variable depth trials (**Figure 8**), K30 1 and K30 10%, respective final measured values were all within the analytical error of the reference gas mixture. Initial offsets and drift that might have occurred during and post-laboratory measurements were not assessed; however, accounting for any drift over the laboratory experiment period would have had negligible difference for the reported $PCO_{2(aq)}$ values and the outcome of the reference experiments.

## IRGA Principle Operation and $PCO_2$ Depth Independence

Based on both theoretical principles and empirical evidence, the measurement of partial pressure of $CO_2$ using a submerged IRGA in equilibrium with surrounding water is independent of hydrostatic pressure (Equation 17; **Figure 8**). However, $CO_2$ concentration spikes occur with sudden increases in hydrostatic pressure (i.e., submerging to deeper depths) before the submerged IRGA returns to the reference $CO_2$ value. This temporary increase in $CO_2$ is interpreted to indicate

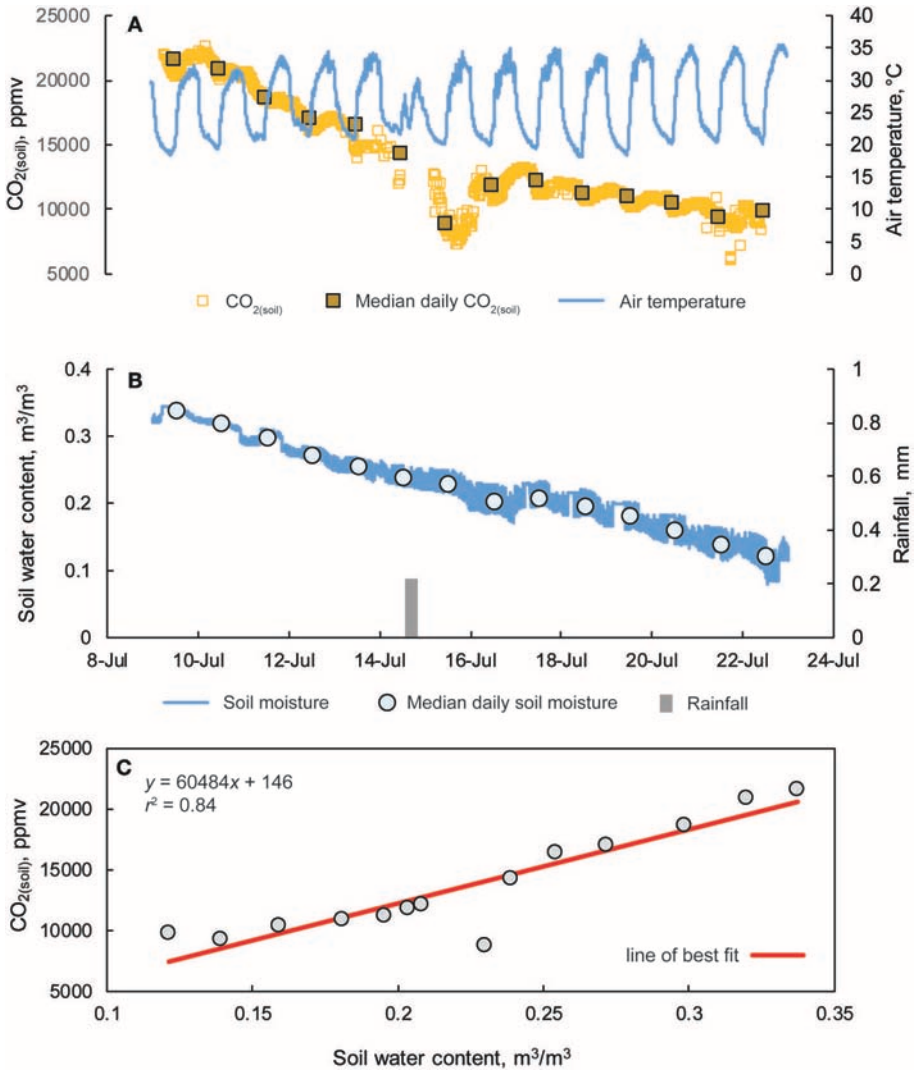



**FIGURE 11 | (A)** Hourly $CO_2$ (soil) concentrations and median daily values from a 10 cm soil cavity and overlying air temperature from 9 to 24 July at the Savoy Experimental Watershed. **(B)** Fifteen-minute interval measurement of soil moisture, median daily soil moisture values, and rainfall totals measured from a weather station located ~2 m from the soil cavity. **(C)** $CO_{2(soil)}$ vs. soil water content.

compression of the enclosed membrane volume, which leads to a decrease in the gas volume, $V_{total}$, whereby: (1) there is an increased molecular density of $CO_2$ without adding more $CO_2$ molecules; which (2) yields a greater $CO_2$ concentration measured by the IRGA; and (3) creates a situation where the total gas pressure inside the enclosed membrane volume was greater than the total dissolved gas pressure of the external water and drives re-equilibration by both diffusion (i.e., partial pressure differences) and advective (i.e., total pressure differences). As $N_2$ was the predominant species present in the reference gas mixtures (i.e., 99.8% nitrogen balance for reference gas mixture of 2,000 ppm $CO_2$), total pressure equilibration was likely driven by $N_2$ exchange. As total pressure within the enclosed membrane re-equilibrates with the total dissolved gas pressure of the water, remaining gas exchange was driven by re-equilibration of partial pressures of the dissolved $CO_2$.

Assuming an initial $V_{total}$ of 5 $cm^3$ and rearranging Equation (14), a volume change of 6.7% would produce the observed increase in $PCO_2$ of ~150 ppmv during the 20–70 cm variable depth experiment from 0 to 20 min of elapsed time (**Figure 8C**). Given the K30 10% materials and waterproofing components being partially flexible, this percent change was within reason.

Accounting for increased hydrostatic pressure acting on the sensor (Johnson et al., 2009) with depth gives rise to overestimates of $PCO_2$, and these overestimates are proportional to the submerged depth. Assuming a water density of 1,000 $kg/m^3$, every 10 cm imparts an increase in hydrostatic pressure equivalent to 9.81 hPa, which would equal an ~8.77% overestimation per meter. Considering the comparative accuracy

of dissolved $CO_2$ measurement between various equilibration methods to be ∼15% (Abril et al., 2015; Yoon et al., 2016), an equal value of overestimation because of the hydrostatic pressure correction is incurred at only 1.68 m depth.

## Field Instrument Comparison

Measured $CO_2$ relations between the Vaisala and CO2-LAMP for $CO_{2(air)}$ and $PCO_{2(stream)}$ covaried linearly and were statistically significant ($r^2 > 0.97$, $p < 0.01$). As previously mentioned, inter-comparison assessments of manual, active, and passive equilibration methods for direct $PCO_2$ measurement exhibited average differences of ∼15% between measurement methods from field sampling (Abril et al., 2015; Yoon et al., 2016). At Blowing Springs, the observed median differences for $PCO_{2(stream)}$ between the Vaisala and CO2-LAMP in this study was only 8.6%. For both $CO_{2(air)}$ and $PCO_{2(stream)}$, differences between the Vaisala and CO2-LAMP likely arose from the varying ability to drive off moisture build up inside the IRGA.

At Blowing Springs, the Vaisala IR source generates more heat than the K30 IR source. In turn, the Vaisala heating element potentially allows for faster removal of any moisture within the IRGA given 100% humidity conditions in the enclosed membrane volume, which can interfere with measurement magnitude and stability. Greater initial $PCO_{2(aq)}$ concentrations for CO2-LAMP data during warm-up periods (**Figures 5A–C**) could be resultant from liquid water condensate decreasing light intensity at the infrared detector (i.e., resulting in artificially large $CO_2$ values; Fietzek et al., 2014). This may explain greater differences among $CO_{2(air)}$ measurements vs. $PCO_{2(aq)}$ between the CO2-LAMP and Vaisala system. As greater temperature variations occurred in the cave air vs. the cave stream, the likelihood for condensation development and overestimation would have been greater for the K30 measuring cave air. Measurement stability over time was likely better sustained in the Vaisala given the ability to remove excess moisture over the deployment period.

Specific factors and correction coefficients for the aforementioned factors vary not only between manufacturers, but also among individual IRGAs of the same manufacturer (McDermitt et al., 1993; Martin et al., 2017). Fully explaining observed differences between $CO_{2(air)}$ and $PCO_{2(stream)}$ were outside of the scope of this study, but work toward accounting for humidity, temperature, and pressure within the membrane-enclosed headspace should, in theory, allow for increased measurement accuracy. Related effects from moisture interference, such as band broadening, effective pressure, and particularly, water dilution effects (McDermitt et al., 1993; Welles and McDermitt, 2005), will affect IRGA accuracy, but were also not fully assessed in this study.

## Capturing $CO_2$ Variability in Natural Settings

Carbon dioxide variability at both sites may be generally described as arising from complex carbon exchange pathways and biogeochemical cycling, which vary down to hourly time-scales. At Blowing Springs Cave, large changes in $CO_{2(air)}$, and $PCO_{2(stream)}$ are linked to cave ventilation, and air flow reversals in the cave system; when $CO_{2(air)}$ increases, the flux of $CO_2$ from the stream decreases, subsequently increasing $PCO_{2(stream)}$. At SEW, $CO_{2(soil)}$ decreases over the monitoring period are likely related to changes in soil moisture (i.e., drying), and coupled reduced soil respiration (Hirano et al., 2003).

Ultimately, the IRGA selection for capturing $CO_{2(air)}$, $PCO_{2(aq)}$, or $CO_{2(soil)}$ variability within environmental systems should be determined based on needed accuracy, priori knowledge of $CO_2$ variability (i.e., temporal and absolute magnitude), and site conditions. With respect to the K30 IRGAs, small variations in $CO_{2(air)}$, and $PCO_{2(stream)}$ <1% (or 10,000 ppm) $CO_2$ like those at Blowing Springs are better suited for the K30 1%. While no reference measurement system was in place (e.g., Vaisala or similar accuracy IRGA) at SEW, $CO_{2(soil)}$ exhibited similar ranges, and environmental response observed in previous studies (Hirano et al., 2003; Jassal et al., 2005). As such, if $CO_{2(soil)}$ is known to be >1% at times when soil respiration is more active, monitoring large changes in $CO_{2(soil)}$ present in most soil systems is better suited for the K30 10%.

When not submerged in water or a fully saturated soil, equilibration of $CO_2$ between the enclosed membrane volume and the environment will be relatively fast, and is likely to fully capture the temporal, and absolute magnitude of $CO_2$ variability. In aquatic environments, such as surface waters, the temporal, and absolute magnitude of $PCO_{2(aq)}$, however, may not be fully captured due to slower equilibration time of the membrane-equilibration method (Yoon et al., 2016). Given site conditions, however, the membrane-equilibration method may still be the only viable method. As the CO2-LAMP equilibration time for $PCO_{2(aq)}$ was measured to be up to 37 min, collection of discrete, direct measurements using faster equilibration methods (see section Direct Measurement of Dissolved $CO_2$ Principles) during varying flow regimes would, at the least, aid in elucidating the magnitude of $CO_2$ variability not captured.

In all cases, field deployments should include: (1) accounting for environmental factors (i.e., humidity, pressure, temperature); (2) performing zero-gas (i.e., no $CO_2$ gas present) measurements; and (3) span gas measurements before, during, and after deployment. Incorporation of these field checks should increase measurement accuracy for $CO_2$ measurements (Fietzek et al., 2014; Martin et al., 2017) without use of an accompanying greater-cost system (e.g., Vaisala system) and yield assessment of both the K30 1% and 10% performance over longer deployment periods (>2 weeks).

## Instrument Fouling, Fabrication Considerations, and Future Field Deployment

From initial deployments of the CO2-LAMP system, environmental factors have been noted which may have solely or in part caused temporary and permanent K30 instrument fouling. First, suspended sediments and other materials (e.g., branches, shells, etc.) can abrade the membrane surface causing microtears. Microtears, while not always visible, allow for liquid water to seep through and damage the K30 instrument's components. Second, upon epoxy application, and

waterproofing of the K30, careful attention is needed to ensure the rubber compound seals the contact between the serial cable, epoxy, and plastic case to prevent water intrusion to the K30 from openings that, similar to microtears, are not always visually apparent. Moreover, application of the rubber compound greatly aids strain relief for the serial cable exiting the plastic case. Third, silt and smaller clay size particles can accumulate on the membrane surface particularly if oriented "face up" relative to the stream surface. If left unprotected, a mud layer or biofilm can accumulate. In both cases, dissolved $CO_2$ concentrations would be more influenced by dissolved $CO_2$ changes within the mud or algal mass rather than the surrounding water. For protection against sediment and biofilm buildup on the membrane surface, it is recommended to orient the sensor vertically in the water column or "face down" relative to the stream surface. For biofilms specifically, use of a bronze mesh has been found to be successful in preventing biofilm accumulation in other freshwater and marine environments (Steven et al., 2014).

Recommendations for future, long-term field deployments using a design similar presented here should consider three modifications. During fabrication, a conformal coating was not applied; however, previous studies employing the K30 for use in floating chambers noted the utility in application of a protective coating on the electronic components for both assembly and field operation (Bastviken et al., 2015). A conformal coating would serve as a protective layer with no disturbance to the K30 printed circuit board. The conformal coating would also provide additional structural support to the initial UART serial connection made to the circuit board before attaching the membrane and rubber compound coating and may help limit any effects from either contraction or expansion of the epoxy-resin during curing. Second, increasing the surface area of the membrane relative to the enclosed membrane volume will increase equilibration time. Lastly, inability to remove excess condensation that results from membrane saturation (Manning et al., 2003) or in-stream temperature changes will greatly diminish instrument accuracy and potentially cause permanent instrument fouling over time (Fietzek et al., 2014). While condensation buildup was not directly investigated, removal of excess moisture from condensation is warranted for long-term CO2-LAMP deployment and $CO_2$ accuracy.

## CONCLUSIONS

Expanding the variety of sites and frequency of $CO_2$ measurements in ambient, soil, and aqueous environments are critical in constraining local carbon dynamics and addressing gaps in efforts to quantify the planetary-scale carbon cycle. Reduction of instrument costs provides a pathway to expand $CO_2$ monitoring across Earth, particularly in research programs where relatively greater-cost platforms are cost-prohibitive.

As part of the CO2-LAMP development, a theoretical presentation of IRGA output, and accompanying experimentation demonstrate that, for $PCO_2$ measurements, temperature is the only correcting variable; however, for measurements in ambient air, total pressure is needed for calculating $x_c$ (i.e., pressure- and temperature-corrected values). Importantly, these findings hold significant implications for past, current, and future implementation of IRGA analyzers for dissolved $PCO_2$ measurement, and, where applicable, recalculation of reported values from previous studies should be considered, particularly for probes at deeper water depths.

Recorded observations in both the laboratory and field demonstrate the CO2-LAMP to be a viable, low-cost alternative to monitoring $CO_2$ in field settings. In the case of $PCO_{2(aq)}$, reported values were within reported uncertainties between different methods. Future work will modify the gas analyzer-water interface to minimize potential fouling due to moisture intrusion and/or long-term condensation buildup.

## DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found at the CO2-LAMP GitHub repository https://github.com/CovingtonResearchGroup/CO2-LAMP.

## AUTHOR CONTRIBUTIONS

JB fabricated, designed, carried out laboratory experiments, and conducted CO2-LAMP field trial in collaboration with MC. Initial formulation of theoretical principles were done by MP. Initial selection, interfacing, and coding was carried out by JM. JB prepared the manuscript with valuable contributions from all co-authors.

## FUNDING

JB acknowledges support Van Brahana Hydrogeology from the University of Arkansas, Department of Geosciences Scholarship. MP acknowledges financial support from the Slovenian Research Agency (research core funding No. P2-0001, bilateral collaboration funding BI-US/17-18-062). JM acknowledges funding from NSF grant EAR-PF 1249895.

## ACKNOWLEDGMENTS

The authors would especially like to thank Sarah Williams, Holly Young, Josue Rodriguez, Hannah Gnoza, and Max Cooper for their assistance during field work. We would also like to thank Max Cooper for troubleshooting Arduino connectivity and power related issues in the OZ27 office. An important set of thanks go to Jerry Fairley, Megan Aunan, and the Hydrologic Computational Group at the University of Idaho for hosting, organizing, and providing invaluable feedback through the wonderful workshop at the University of Idaho MILL which introduced students and faculty to building an earlier version of the CO2-LAMP. Lastly, we would very much like to thank the reviewers. Their valuable comments and suggestions have greatly improved this manuscript.

## REFERENCES


Abril, G., Bouillon, S., Darchambeau, F., Teodoru, C. R., Marwick, T. R., Tamooh, F., et al. (2015). Technical note: large overestimation of $pCO_2$ calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences* 12, 67–78. doi: 10.5194/bg-12-67-2015

Al-Qinna, M., Scott, H. D., Brye, K. R., Van Brahana, J., Sauer, T. J., and Sharpley, A. (2014). Coarse fragments affect soil properties in a mantled-karst landscape of the ozark highlands. *Soil Sci.* 179, 42–50. doi: 10.1097/SS.0000000000000034

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., et al. (2014). $CO_2$, CO, and $CH_4$ measurements from tall towers in the NOAA Earth system research laboratory's global greenhouse gas reference network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas. *Atmos. Meas. Tech.* 7, 647–687. doi: 10.5194/amt-7-647-2014

Bareer, R. M. (1939). Permeation, diffusion and solution of gases. *Trans. Faraday Soc.* 35, 628–643. doi: 10.1039/tf9393500628

Bastviken, D., Sundgren, I., Natchimuthu, S., Reyier, H., and Gålfalk, M. (2015). Technical note: cost-efficient approaches to measure carbon dioxide ($CO_2$) fluxes and concentrations in terrestrial and aquatic environments using mini loggers. *Biogeosciences* 12, 3849–3859. doi: 10.5194/bg-12-3849-2015

Beddows, P. A., and Mallon, E. K. (2018). Cave pearl data logger: a flexible arduino-based logging platform for long-term monitoring in harsh environments. *Sensors* 18:530. doi: 10.3390/s18020530

Bergman, T. L., Lavine, A. S., Incropera, F. P., and DeWitt, D. P. (2011). *Fundamentals of Heat and Mass Transfer, 7th Edn*. New York, NY: John Wiley and Sons.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A., and Crowther, T. W. (2016). Managing uncertainty in soil carbon feedbacks to climate change. *Nat. Clim. Chang.* 6, 751–758. doi: 10.1038/nclimate3071

Brahana, J. V., Hays, P. D., Kresse, T. M., Sauer, T. J., and Stanton, G. P. (1999). "The Savoy Experimental Watershed–early lessons for hydrogeologic modeling from a well-characterized karst research site," in *Karst Waters Institute Special Publication 5*, eds A. N. Palmer, M. V. Palmer, and I. D. Sasowsky (Charles Town, WV: Karst Waters Institute), 247–254.

Brantley, S. L., Dibiase, R. A., Russo, T. A., Shi, Y., Lin, H., Davis, K. J., et al. (2016). Designing a suite of measurements to understand the critical zone. *Earth Surf. Dyn.* 4, 211–235. doi: 10.5194/esurf-4-211-2016

Brantley, S. L., Goldhaber, M. B., and Vala Ragnarsdottir, K. (2007). Crossing disciplines and scales to understand the critical zone. *Elements* 3, 307–314. doi: 10.2113/gselements.3.5.307

Breecker, D., and Sharp, Z. D. (2008). A field and laboratory method for monitoring the concentration and isotopic composition of soil $CO_2$. *Rapid Commun. Mass Spectrom.* 22, 449–454. doi: 10.1002/rcm.3382

Broecker, W. S., and Sanyal, A. (1998). Does atmospheric $CO_2$ police the rate of chemical weathering? *Global Biogeochem. Cycles* 12, 403–408. doi: 10.1029/98GB01927

Burton, M. R., Sawyer, G. M., and Granieri, D. (2013). Deep carbon emissions from volcanoes. *Rev. Mineral. Geochem.* 75, 323–354. doi: 10.2138/rmg.2013.75.11

Colt, J. (2012). *Computation of Dissolved Gas Concentration in Water as Functions of Temperature, Salinity and Pressure*, *2nd Edn*. Oxford: Elsevier.

Covington, M. D., and Vaughn, K. A. (2018). Carbon dioxide and dissolution rate dynamics within a karst underflow-overflow system, Savoy Experimental Watershed, Arkansas, USA. *Chem. Geol.* doi: 10.1016/j.chemgeo.2018.03.009. [Epub ahead of print].

Cressey, D. (2017). Toolbox: age of the Arduino. *Nature* 544, 125–126. doi: 10.1038/544125a

D'Aoust, B. G., and Clark, M. J. R. (1980). Analysis of supersaturated air in natural waters and reservoirs. *Trans. Am. Fish. Soc.* 109, 708–724. doi: 10.1577/1548-8659(1980)109<708:AOSAIN>2.0.CO;2

Davidson, E. A., Figueiredo, R. O., Markewitz, D., and Aufdenkampe, A. K. (2010). Dissolved $CO_2$ in small catchment streams of eastern Amazonia: a minor pathway of terrestrial carbon loss. *J. Geophys. Res. Biogeosci.* 115, 1–6. doi: 10.1029/2009JG001202

De Gregorio, S., Camarda, M., Longo, M., Cappuzzo, S., Giudice, G., and Gurrieri, S. (2011). Long-term continuous monitoring of the dissolved $CO_2$ performed by using a new device in groundwater of the Mt. Etna (southern Italy). *Water Res.* 45, 3005–3011. doi: 10.1016/j.watres.2011.03.028

De Gregorio, S., Gurrieri, S., and Valenza, M. (2005). A PTFE membrane for the *in situ* extraction of dissolved gases in natural waters: theory and applications. *Geochem. Geophys. Geosyst.* 6, 1–13. doi: 10.1029/2005GC000947

Decina, S. M., Hutyra, L. R., Gately, C. K., Getson, J. M., Reinmann, A. B., Short Gianotti, A. G., et al. (2016). Soil respiration contributes substantially to urban carbon fluxes in the greater Boston area. *Environ. Pollut.* 212, 433–439. doi: 10.1016/j.envpol.2016.01.012

Demars, B. O. L., Thompson, J., and Manson, J. R. (2015). Stream metabolism and the open diel oxygen method: principles, practice, and perspectives. *Limnol. Oceanogr. Methods* 13, 356–374. doi: 10.1002/lom3.10030

Enns, T., Scholander, P. F., and Bradstreet, E. D. (1965). Effect of hydrostatic pressure on gases dissolved in water. *J. Phys. Chem.* 69, 389–391. doi: 10.1021/j100886a005

Fietzek, P., Fiedler, B., Steinhoff, T., and Körtzinger, A. (2014). *In situ* quality assessment of a novel underwater $pCO_2$ sensor based on membrane equilibration and NDIR spectrometry. *J. Atmos. Ocean. Technol.* 31, 181–196. doi: 10.1175/JTECH-D-13-00083.1

Fisher, D. K., and Gould, P. J. (2012). Open-source hardware is a low-cost alternative for scientific instrumentation and research. *Mod. Instrum.* 1, 8–20. doi: 10.4236/mi.2012.12002

Florea, L. J. (2015). Carbon flux and landscape evolution in epigenic karst aquifers modeled from geochemical mass balance. *Earth Surf. Process. Landforms* 40, 1072–1087. doi: 10.1002/esp.3709

Frankignoulle, M., Borges, A., and Biondo, R. (2001). A new design of equilibrator to monitor carbon dioxide in highly dynamic and turbid environments. *Water Res.* 35, 1344–1347. doi: 10.1016/S0043-1354(00)00369-9

Gadagkar, S. R., and Call, G. B. (2015). Computational tools for fitting the Hill equation to dose–response curves. *J. Pharmacol. Toxicol. Methods* 71, 68–76. doi: 10.1016/j.vascn.2014.08.006

Gardner, P., and Solomon, D. K. (2009). An advanced passive diffusion sampler for the determination of dissolved gas concentrations. *Water Resour. Res.* 45, 1–12. doi: 10.1029/2008WR007399

Goutelle, S., Maurin, M., Rougier, F., Barbaut, X., Bourguignon, L., Ducher, M., et al. (2008). The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam. Clin. Pharmacol.* 22, 633–648. doi: 10.1111/j.1472-8206.2008.00633.x

Hamme, R. C., Berry, J. E., Klymak, J. M., and Denman, K. L. (2015). *In situ* $O_2$ and $N_2$ measurements detect deep-water renewal dynamics in seasonally-anoxic Saanich inlet. *Cont. Shelf Res.* 106, 107–117. doi: 10.1016/j.csr.2015.06.012

Hammer, Ø. (2019). *Reference Manual: PAST PAleontological STatistics*. Version 3.25. System. Olso: Natural History Museum University of Oslo.

Hammer, Ø., Harper, D. A. T., Ryan, D. D., and Ryan, P. D. (2001). PAST : paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4:9. doi: 10.1016/j.bcp.2008.05.025

Hari, P., Pumpanen, J., Huotari, J., Kolari, P., Grace, J., Vesala, T., et al. (2008). High-frequency measurements of productivity of planktonic algae using rugged nondispersive infrared carbon dioxide probes. *Limnol. Oceanogr. Methods* 6, 347–354. doi: 10.4319/lom.2008.6.347

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *Proc. Physiol. Soc.* 53, 1689–1699. doi: 10.1017/CBO9781107415324.004

Hirano, T., Kim, H., and Tanaka, Y. (2003). Long-term half-hourly measurement of soil $CO_2$ concentration and soil respiration in a temperate deciduous forest. *J. Geophys. Res. Atmos.* 108, 1–13. doi: 10.1029/2003JD003766

Jarvie, H. P., King, S. M., and Neal, C. (2017). Inorganic carbon dominates total dissolved carbon concentrations and fluxes in British rivers: application of the THINCARB model – thermodynamic modelling of inorganic carbon in freshwaters. *Sci. Total Environ.* 575, 496–512. doi: 10.1016/j.scitotenv.2016.08.201

Jarvie, H. P., Sharpley, A. N., Brahana, V., Simmons, T., Price, A., Neal, C., et al. (2014). Phosphorus retention and remobilization along hydrological pathways in karst terrain. *Environ. Sci. Technol.* 48, 4860–4868. doi: 10.1021/es405585b

Jassal, R., Black, A., Novak, M., Morgenstern, K., Nesic, Z., and Gaumont-Guay, D. (2005). Relationship between soil $CO_2$ concentrations and forest-floor $CO_2$ effluxes. *Agric. For. Meteorol.* 130, 176–192. doi: 10.1016/j.agrformet.2005.03.005

Jochheim, H., Wirth, S., and von Unold, G. (2018). A multi-layer, closed-loop system for continuous measurement of soil $CO_2$ concentration. *J. Plant Nutr. Soil Sci.* 181, 61–68. doi: 10.1002/jpln.201700259

Johnson, M. S., Billett, M. F., Dinsmore, K. J., Wallin, M., Dyson, K. E., and Jassal, R. S. (2009). Direct and continuous measurement of dissolved carbon dioxide in freshwater aquatic systems-method and applications. *Ecohydrology* 3, 68–78. doi: 10.1002/eco.95

Joos, O., Saurer, M., Heim, A., Hagedorn, F., Schmidt, M. W. I., and Siegwolf, R. T. W. (2008). Can we use the $CO_2$ concentrations determined by continuous-flow isotope ratio mass spectrometry from small samples for the Keeling plot approach? *Rapid Commun. Mass Spectrom.* 22, 4029–4034. doi: 10.1002/rcm.3827

Kling, G. W., Clark, M. A., Compton, H. R., Devine, J. D., Evans, W. C., Humphrey, A. M., et al. (1987). The 1986 Lake Nyos gas disaster in Cameroon, West Africa. *Science* 236, 169–175. doi: 10.1126/science.236.4798.169

Knierim, K. J., Pollock, E., and Hays, P. D. (2013). Using isotopes of dissolved inorganic carbon species and water to separate sources of recharge in a cave spring, Northwestern Arkansas, USA. *Acta Carsologica* 42, 261–276. doi: 10.3986/ac.v42i2-3.667

Kresse, T. M., Hays, P. D., Merriman, K. R., Gillip, J. A., Fugitt, D. T., Spellman, J. L., et al. (2014). *Aquifers of Arkansas-Protection, Management, and Hydrologic and Geochemical Characteristics of Groundwater Resources in Arkansas: U.S. Geological Survey Scientific Investigations Report*, U.S. Geological Survey, 2014–5149. doi: 10.3133/sir20145149

Kruczek, B. (2015). "Convective transport," in *Encyclopedia of Membranes*, eds E. Drioli and L. Giorno (Berlin; Heidelberg: Springer), 1–3. doi: 10.1007/978-3-642-40872-4_1994-1

Kusakabe, M., and Sano, Y. (1992). "The origin of gases in Lake Nyos, Cameroon," in *Natural Hazards in West and Central Africa* eds S. J. Freeth, C. O. Ofoegbu, and K. M. Onuoha (Wiesbaden: Vieweg+Teubner Verlag), 83−95. doi: 10.1007/978-3-663-05239-5_9

Liu, S., and Raymond, P. A. (2018). Hydrologic controls on $pCO_2$ and $CO_2$ efflux in US streams and rivers. *Limnol. Oceanogr. Lett.* 3, 428–435. doi: 10.1002/lol2.10095

Lombardozzi, D. L., Bonan, G. B., Smith, N. G., Dukes, J. S., and Fisher, R. A. (2015). Temperature acclimation of photosynthesis and respiration: a key uncertainty in the carbon cycle-climate feedback. *Geophys. Res. Lett.* 42, 8624–8631. doi: 10.1002/2015GL065934

Lowenstern, J. B. (2001). Carbon dioxide in magmas and implications for hydrothermal systems. *Miner. Depos.* 36, 490–502. doi: 10.1007/s001260100185

Manning, A. H., Solomon, D., and Sheldon, A. L. (2003). Applications of a total dissolved gas pressure probe in ground water studies. *Ground Water* 41, 440–448. doi: 10.1111/j.1745-6584.2003.tb02378.x

Martin, C. R., Zeng, N., Karion, A., Dickerson, R. R., Ren, X., Turpie, B. N., et al. (2017). Evaluation and environmental correction of ambient $CO_2$ measurements from a low-cost NDIR sensor. *Atmos. Meas. Tech.* 10, 2383–2395. doi: 10.5194/amt-10-2383-2017

McDermitt, D. K., Welles, J. M., and Eckles, R. D. (1993). *Effects of Temperature, Pressure and Water Vapor on Gas Phase Infrared Absorption by $CO_2$*. Lincoln, NE: LI-COR Tech. Publ.

McDowell, N., Baldocchi, D., Barbour, M., Bickford, C., Cuntz, M., Hanson, D., et al. (2008). Understanding the stable isotope composition of biosphere-atmosphere $CO_2$ exchange. *Eos Trans. Am. Geophys. Union* 89, 94–95. doi: 10.1029/2008EO100002

Moran, D., Tirsgård, B., and Steffensen, J. F. (2010). The accuracy and limitations of a new meter used to measure aqueous carbon dioxide. *Aquac. Eng.* 43, 101–107. doi: 10.1016/j.aquaeng.2010.07.003

Olah, G. A., Prakash, G. K. S., and Goeppert, A. (2011). Anthropogenic chemical carbon cycle for a sustainable future. *J. Am. Chem. Soc.* 133, 12881–12898. doi: 10.1021/ja202642y

Pearce, J. M. (2012). Building research equipment with free, open-source hardware. *Science* 337, 1303–1304. doi: 10.1126/science.1228183

Queiβer, M., Granieri, D., and Burton, M. (2016). A new frontier in $CO_2$ flux measurements using a highly portable DIAL laser system. *Sci. Rep.* 6:33834. doi: 10.1038/srep33834

Richter, D. deB., and Mobley, M. L. (2009). Monitoring earth's critical zone. *Science* 326, 1067–1068. doi: 10.1126/science.1179117

Ryan, M. C., Roy, J. W., and Heagle, D. J. (2015). Dissolved gas "concentrations" or "concentration estimates" - a comment on 'Origin, distribution and hydrogeochemical controls on methane occurrences in shallow aquifers, southwestern Ontario, Canada' by Jennifer C. McIntosh, Stephen E. Grasby, Stewart M. *Appl. Geochem.* 63, 218–221. doi: 10.1016/j.apgeochem.2015.08.015

Sánchez-Cañete, E. P., Scott, R. L., van Haren, J., and Barron-Gafford, G. A. (2017). Improving the accuracy of the gradient method for determining soil carbon dioxide efflux. *J. Geophys. Res. Biogeosci.* 122, 50–64. doi: 10.1002/2016JG003530

Sanford, W. E., Shropshire, R. G., and Solomon, D. K. (1996). Dissolved gas tracers in groundwater: Simplified injection, sampling, and analysis. *Water Resour. Res.* 32, 1635–1642. doi: 10.1029/96WR00599

Schimel, D., Stephens, B. B., and Fisher, J. B. (2015). Effect of increasing $CO_2$ on the terrestrial carbon cycle. *Proc. Natl. Acad. Sci. U.S.A.* 112, 436–441. doi: 10.1073/pnas.1407302112

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., et al. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* 414, 169–172. doi: 10.1038/35102500

Scott, H. D. (2000). *Soil Physics: Agricultural and Environmental Applications*. Ames, IA: Iowa State University Press.

Serrano-Ortiz, P., Roland, M., Sanchez-Moral, S., Janssens, I. A., Domingo, F., Goddéris, Y., et al. (2010). Hidden, abiotic $CO_2$ flows and gaseous reservoirs in the terrestrial carbon cycle: review and perspectives. *Agric. For. Meteorol.* 150, 321–329. doi: 10.1016/j.agrformet.2010.01.002

Soil Survey Staff (2019). *National Cooperative Soil Characterization Database*. Soil Series Classification Database. Natural Resources Conservation Service, United States Department of Agriculture.

Steven, A., Hodge, J., Cannard, T., Carlin, G., Franklin, H., McJannet, D., et al. (2014). *ContinuousWater Quality Monitoring on the Great Barrier Reef*. Brisbane, QLD: CSIRO. Available online at: https://publications.csiro.au/rpr/pub?pid=csiro:EP15170

Stumm, W., and Morgan, J. J. (1996). *Aquatic Chemistry, 3rd Edn*. New York, NY: John Wiley & Sons.

Takahashi, T. (1961). Carbon dioxide in the atmosphere and in Atlantic Ocean water. *J. Geophys. Res.* 66:477. doi: 10.1029/JZ066i002p00477

Tang, J., Baldocchi, D. D., Qi, Y., and Xu, L. (2003). Assessing soil $CO_2$ efflux using continuous measurements of $CO_2$ profiles in soils with small solid-state sensors. *Agric. For. Meteorol.* 118, 207–220. doi: 10.1016/S0168-1923(03)00112-6

Urban, A. L., Gulliver, J. S., and Johnson, D. W. (2008). Modeling total dissolved gas concentration downstream of spillways. *J. Hydraul. Eng.* 134, 550–561. doi: 10.1061/(ASCE)0733-9429(2008)134:5(550)

Ward, H. C., Kotthaus, S., Grimmond, C. S. B., Bjorkegren, A., Wilkinson, M., Morrison, W. T. J., et al. (2015). Effects of urban density on carbon dioxide exchanges: observations of dense urban, suburban and woodland areas of southern England. *Environ. Pollut.* 198, 186–200. doi: 10.1016/j.envpol.2014.12.031

Welles, J. M., and McDermitt, D. K. (2005). "Measuring carbon dioxide in the atmosphere," in *Micrometerology in Agricultural Systems*, eds M. K. Viney, J. L. Hatfield and J. M. Baker (Madison, WI: American Society of Agronomy, Inc.; Crop Science Society of America, Inc; Soil Science Society of America, Inc.), 287–320. Available online at: https://dl.sciencesocieties.org/publications/books/tocs/agronomymonogra/micrometeorolog

Werner, C., and Cardellini, C. (2006). Comparison of carbon dioxide emissions with fluid upflow, chemistry, and geologic structures at the Rotorua geothermal system, New Zealand. *Geothermics* 35, 221–238. doi: 10.1016/j.geothermics.2006.02.006

Yoon, T. K., Jin, H., Oh, N. H., and Park, J. H. (2016). Technical note: assessing gas equilibration systems for continuous $pCO_2$ measurements in inland waters. *Biogeosciences* 13, 3915–3930. doi: 10.5194/bg-13-3915-2016

Young, H. (2018). *Quantifying carbon dioxide fluxes in the air and water in Blowing Springs Cave, Arkansas* (Theses and Dissertations), 2844. Available online at: https://scholarworks.uark.edu/etd/2844

**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.

*Copyright © 2019 Blackstock, Covington, Perne and Myre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

METHODS
published: 10 December 2019
doi: 10.3389/feart.2019.00330

# High-Resolution Bathymetry Mapping of Water Bodies: Development and Implementation

*Liah X. Coggins* and Anas Ghadouani**

*Department of Civil, Environmental and Mining Engineering, The University of Western Australia, Crawley, WA, Australia*

Traditionally, bathymetry mapping of ponds, lakes, and rivers have used techniques which are low in spatial resolution. Waste stabilization ponds (WSPs) are utilized worldwide for wastewater treatment, and throughout their operation require periodic sludge surveys. Sludge accumulation in WSPs can impact performance by reducing the effective volume of the pond, and altering the pond hydraulics and wastewater treatment efficiency. Traditionally, sludge heights, and thus sludge volume, have been measured using techniques such as the "sludge judge" and the "white towel" test. Both of these methods have low spatial resolution, are subjective in terms of precision and accuracy, are labor intensive, and require a high level of safety precautions. A sonar device fitted to a remotely operated vehicle (ROV) can improve the resolution and accuracy of sludge height measurements, as well as reduce labor and safety requirements. This technology is readily available; however, despite its applicability, it has not been previously assessed for use on WSPs. This study aimed to design, build, and assess the performance of an ROV to measure sludge height in WSPs. Profiling of several WSPs has shown that the ROV with autonomous sonar device is capable of providing bathymetry with greatly increased spatial resolution in a greatly reduced profiling time. To date, the ROV has been applied on in excess of 400 WSPs across Australia, several large lakes, stormwater retention ponds, river beds, and drinking water reservoirs. ROVs, such as the one built in this study, will be useful for not only determining sludge profiles, but also in calculating sludge accumulation rates and in evaluating pond hydraulic efficiency. As demonstrated, this technology is not limited to application in wastewater management, with the potential for wider application in the monitoring of other small to medium-sized water bodies, including reservoirs, lakes, channels, recreational water bodies, river beds, mine tailing dams and commercial ports.

**Keywords: bathymetry, mapping, survey, waste stabilization ponds, lakes, ponds, remote sensing, water bodies**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology,*
*Netherlands*

***Reviewed by:***
*Ulf Mallast,*
*Helmholtz Centre for Environmental Research (UFZ), Germany*
*Julia Hopkins,*
*Delft University of Technology,*
*Netherlands*
*Hessel Winsemius,*
*Delft University of Technology,*
*Netherlands*

****Correspondence:***
*Liah X. Coggins*
*liah.coggins@uwa.edu.au*
*Anas Ghadouani*
*anas.ghadouani@uwa.edu.au*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *31 January 2019*
***Accepted:*** *27 November 2019*
***Published:*** *10 December 2019*

***Citation:***
*Coggins LX and Ghadouani A (2019) High-Resolution Bathymetry Mapping of Water Bodies: Development and Implementation. Front. Earth Sci. 7:330. doi: 10.3389/feart.2019.00330*

## INTRODUCTION

Bathymetry mapping of ponds, lakes, and rivers often use techniques which are low in spatial resolution, subjective in terms of precision and accuracy, labor intensive, and which require a high level of safety precaution. Waste stabilization ponds (WSPs) are simple, highly efficient, low-cost, low-maintenance and robust systems for treating wastewater (Mara, 2004; Nelson et al., 2004; Picot et al., 2005). In WSPs, wastewater constituents are removed by sedimentation

or transformed by biological and chemical processes, and a sludge layer forms due to the sedimentation of influent suspended solids, algae, and bacteria (Nelson et al., 2004). Sludge accumulation can affect pond performance by reducing pond effective volume and changing the bottom bathymetry, thus altering pond hydraulics (e.g., Persson, 2000; Nelson et al., 2004; Coggins et al., 2017). and compromising the discharge quality (e.g., Ghadouani and Coggins, 2011). Effective, safe and sustainable operation of WSPs therefore requires detailed information about sludge accumulation, distribution, and its effect on hydraulic characteristics. Furthermore, sludge accumulation can lead to increased methane production, thus contributing to greenhouse gas emissions (Hernandez-Paniagua et al., 2014; Glaz et al., 2016). This knowledge is essential for planning pond maintenance, in particular sludge removal and disposal, which can be highly expensive and complex (Nelson et al., 2004; Picot et al., 2005; Alvarado et al., 2012a). Traditional methods of measuring sludge height, and thus total sludge volume, in WSPs include the use of a "sludge judge" (a clear plastic pipe) (Westerman et al., 2008), or the "white towel" test (Mara, 2004). Sludge surveys using these techniques are typically conducted on a rectangular grid, with height measurements taken by an operator deploying the measuring apparatus from a boat. The number of point measurements taken in each pond is dependent on both the size of the pond and the grid spacing chosen by the operator. Such surveys are time consuming and have low spatial resolution; however, data from these surveys is vital for sludge management (Peña et al., 2000; Nelson et al., 2004; Picot et al., 2005).

Small sonar devices equipped with global positioning system (GPS) technology, also known as fishfinders, are not only readily available and widely used by people in boating, but have also previously been used to determine the depth of water and sludge height in small agricultural lagoons (Singh et al., 2008). Through the use of GPS technology in conjunction with sonar, the location and vertical distance to the top of the sludge layer (or sludge blanket) can be simultaneously recorded to a memory card; this data can then be used to develop contour maps of sludge and in the determination of total sludge volume in the pond (Singh et al., 2008). However, despite this technology being highly applicable for bathymetry mapping studies, it has so far been underutilized.

Remotely operated vehicles (ROVs) are becoming increasingly popular for research applications, with ROVs being developed for water sampling (Kaizu et al., 2011), and current profiling (Kriechbaumer et al., 2015). The coupling of sonar technology with an ROV platform has several advantages over traditional sludge measurement techniques, as they:

1. Require less manual labor,
2. Reduce safety risks and occupational health and safety requirements, as it does not require launching a boat onto the WSP,
3. Reduce the subjectivity of measurements taken with traditional techniques (i.e., removes human error in reading and the differences in having different people conduct the surveys),
4. Collect many more data points in a shorter period of time (i.e., increase measuring efficiency) with greater spatial resolution (i.e., smaller grid spacing), and
5. Are relatively cost effective, widely accessible, and low maintenance.

Additionally, the combination of an ROV fitted with sonar will be a significant advantage for bathymetric surveys of many water bodies other than WSPs. ROVs may also be applied to small to medium sized water bodies, such as lakes and stormwater retention wetlands, drinking water reservoirs, rivers, pools, channels, and recreational and commercial ports. The improvement in spatial resolution of pond bathymetry data alone will greatly improve models used to understand pond hydraulics and how sludge accumulation and geometry affect performance (Passos et al., 2016; Coggins et al., 2017, 2018); these could in turn be used to develop new WSP coupled models of hydraulics and biology. Thus, the main objective of this study was to assess the performance of a ROV with GPS-equipped sonar to measure sludge height in a WSP, with the aim to develop it to a point where it could be implemented for research and within industry.

## MATERIALS AND METHODS

### GPS Equipped Sonar Unit

For the development of the ROV, a sonar unit with GPS (model HDS-5, Lowrance Electronics, Tulsa, Oklahoma) with an 83/200 kHz transducer was selected and tested, as the built-in GPS allows for the simultaneous acquisition of water depth and map coordinate data. The unit also allows for continuous data logging, and saves files to an SD memory card; online user forums for this sonar unit report GPS accuracy between 1–6 m. This particular unit was chosen after field trials on WSPs during 2010; the unit was commercially available, and reasonable in price. At specific locations, point measurements of sludge height were taken by both the sludge judge and the sonar. As the sonar unit measures and records local water depth, sludge heights were calculated by subtracting depth measurements from the average pond depth (from pond manager asset data). There was a very strong correlation ($R^2$ = 0.98) between the two measurement techniques (**Figure 1**) with a tendency for the sonar reading to be slightly higher than the corresponding sludge judge reading (Coggins et al., 2017).

### Remotely Operated Vehicle Design

Previous sonar profiling studies have used unmanned airboats (Singh et al., 2008; Kaizu et al., 2011), and more recently unmanned aerial vehicles (e.g., Bandini et al., 2018). WSPs in Western Australia, and Australia in general, are commonly located in cleared areas, and thus do not have any shelter from the wind. For example, weather data for August 2011 from a station near a Western Australian WSP recorded wind gusts of up to 85 km $h^{-1}$ (data from Australian Bureau of Meteorology), with an average 9 am wind speed of 13 km $h^{-1}$, and observing wind ripples on the pond surface is extremely common. Considering the medium to

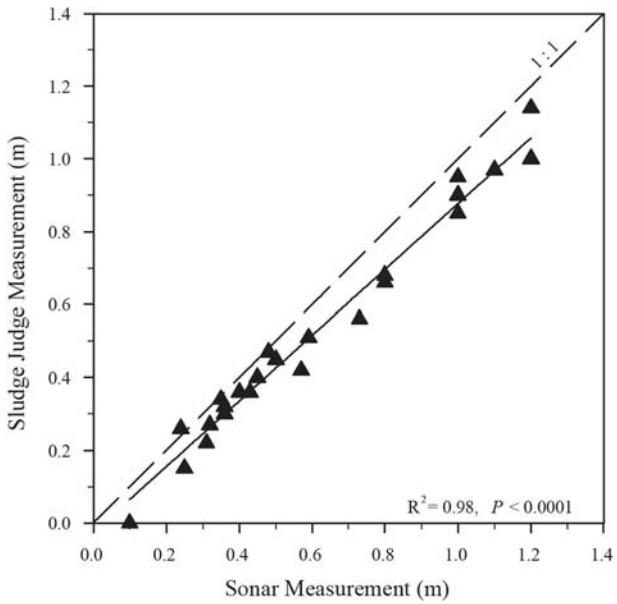



**FIGURE 1 |** Sonar vs. sludge judge measurement ($n$ = 24). The dashed line indicates the 1:1 relationship, with the relationship between sonar and sludge judge (solid line) being very close to ideal (Reprinted from Coggins et al., 2017, Page 357, Copyright (2016), with permission from Elsevier).

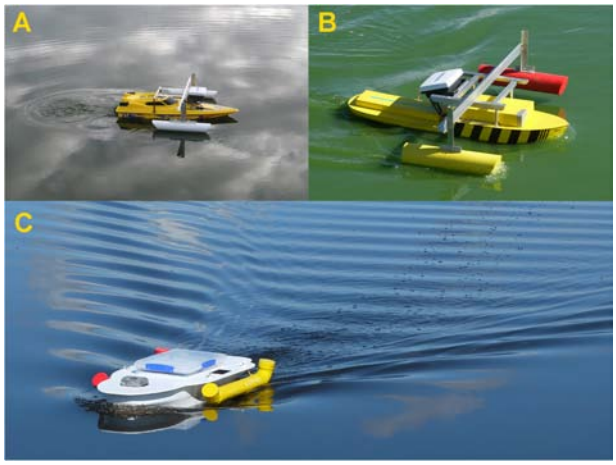



**FIGURE 2 |** Design iterations of the sonar profiling ROV (detailed specifications in **Table 1**). **(A)** The prototype was an off-the-shelf speed boat, fitted with a frame to mount the sonar unit and stabilizers to the boat. **(B)** Redesign of the boat with a more robust deep V-shaped hull, with frame for sonar and stabilizers. **(C)** Current boat design with U-shaped hull and sonar mounted inside.

strong prevailing wind conditions at WSPs across Australia, it was decided that a boat with submerged rudder and propeller would be better suited for complete and rapid sludge profiling of ponds.

A prototype ROV was built using an off-the-shelf model boat, with the sonar mounted to a frame on top of the boat, and with outriggers to stabilize the boat when turning (**Figure 2A**). The ROV was controlled using a 2.4 GHz surface radio, and driven by the operator and not by a pre-determined GPS-referenced path.

Implementation in industry was always at the forefront of the development of the boat, however, after proof of concept testing it was obvious that some improvements would be required to make the ROV more suitable and robust. In trials, the optimum speed for profiling was determined to be 2–4 km $h^{-1}$; however, the prototype boat, an off-the-shelf model with a shallow V-shape hull was built for speeds in excess of 30 km $h^{-1}$. The non-ideal hull shape resulted in shorter battery life, as the electronic components were not suited for low-speed use. Slow moving water vessels, such as tugboats and barges, have hulls with a deeper V-shape or U-shape, designed to cut through water with very little propulsion. These types of hulls are not only more suitable for low-speed applications, but also more stable in the water. An improvement to the hull shape would thus improve both boat stability and battery life. In addition to a different hull shape, efforts were invested in making the physical and electronic components of the boat more robust and reliable to increase run-time (battery life), and decrease wear and tear. Most importantly, the boat needed to be simple for operators to be able to service and replace mechanical and electrical components as required. A summary of specifications of the boat from the prototype boat (**Figure 2A**), the first redesign (**Figure 2B**), to the final design (**Figure 2C**) can be found in **Table 1**.

The final design of the boat (**Figure 2C**) has a U-shaped hull made of fiberglass. The use of a brushless, low-RPM/V motor along with Nickel Metal Hydride (NiMH) batteries has extended run-time from 20 min to 2–4 h. The U-shaped hull also allows for room for the sonar to be mounted inside the boat, removing the need for a frame. The sonar unit and batteries were positioned near the center of gravity of the boat, while the sonar transducer is fixed at the front of the hull. Due to the stability provided by the hull, the use of stabilizers is optional. Displacement hull boats require a significant amount of ballast for stability; the NiMH batteries, and lead weight provide this. The fully laden boat weighs approximately 8 kg. The boat is driven manually using a 2.4 GHz surface radio with a range of up to 200 m. This boat design is: (1) durable, easily shipped on planes and in cars, (2) consistent in operation, and (3) low maintenance. In addition, we have demonstrated in the field that this ROV design is suitable in strong wind conditions (60–70 km $h^{-1}$), with boat stability and data quality not being affected; however, windy conditions can reduce battery life and may not be ideal for operators.

## Assessment of Prototype ROV Operation

In the prototype stage, the remote control boat with sonar was tested on several ponds to ensure that it: (1) was suitable for use on WSP, (2) was accurate in its measurement, and (3) had high reproducibility of results.

The ROV was tested at two wastewater treatment plants close to the Perth metropolitan area, Western Australia. Two ponds were chosen for testing: Pond 1, a secondary maturation, and Pond 2, a primary facultative pond; dimensions of the selected ponds were 59 × 62 m and 84 × 84 m, respectively. Pond managers profiled both ponds using a sludge judge during

**TABLE 1 |** Specifications of the remote control boats built, showing their evolution from an off-the-shelf prototype (**Figure 2A**), to the redesign (**Figure 2B**), and finally to a robust and reliable shape (**Figure 2C**).

| Component | Iteration | | |
|---|---|---|---|
| | ***Prototype*** | ***Redesign*** | ***Final design*** |
| Hull shape | V-shape | Deep V-shape | U-shape |
| Hull material | Light plastic | High-density plastic | Fiberglass |
| Length | 1000 mm | 800 mm | 900–1000 mm |
| Width | 400 mm | 180 mm | 270 mm |
| Depth | 100 mm | 170 mm | 170–200 mm |
| Remote control | 2.4 GHz surface radio | | |
| Remote control range | 150–200 m | | |
| Motor | Information not available | Brushless, 710 KV (RPM/V) | Brushless, 400 KV |
| Speed controller | Information not available | Waterproof, brushless for RC car | Waterproof, brushless 50A for boat |
| Servo | Information not available | Digital, waterproof, with operating range of 5–7 V | |
| Batteries | 5 A lead acid (boat); 7.2 V lead acid (sonar) | 7.2 V 5000 mAh NiMH | |
| Number of batteries | 2 | 8 | 4 to 6 |
| Battery life per set | 20 min | | 2–4 h |
| Propeller | 30 mm 2-blade (plastic) | 35 mm 3-blade (brass) | 35 mm 3-blade (brass)/40 mm 2-blade (stainless steel) |
| Number of propellers | 2 | 1 | 1 |
| Outriggers | Yes | | Optional |
| Sonar mount | Outside | | In-hull |
| Sonar unit | Lowrance HDS-5 | | Lowrance HDS-5/Hook 5/Hook2-5 |
| Transducer | 83/200 kHz | | 83/200 kHz/Lowrance Hybrid Dual Imaging (HDI) 83/200 kHz–455/800 kHz/HDI Splitshot 200 kHz–455/800 kHz |
| Memory card | SD | | SD/microSD |
| Ballast required | None | | 2–4 kg lead weight |

June 2011. The selected ponds were profiled several times during the period of June-August 2011, with data collected using the logging function on the sonar. Data was collected along transects approximately 2 m apart in both the lateral and longitudinal directions. The boat was maintained at a constant low speed (approximately 2–4 km $h^{-1}$) while profiling, and kept in constant motion for as long as possible. In addition, some profiles also included a "run" around the pond perimeter to obtain measurements as close to the edge of the pond as possible. The sonar and transducer set for shallow water using the manufacturer specifications. Additionally, ping speed was set to the maximum resolution of 3200 bytes per ping.

## Data Processing and Analysis

Data was downloaded from the sonar SD memory card into Sonar Log Viewer (version 2.1.2, Lowrance Electronics, Tulsa, Oklahoma), and then exported to Microsoft Excel Comma Separated Value (csv) format (**Figure 3**). During processing, false depth data was removed, i.e., depth readings greater than the pond depth from pond operator asset data (1–2% of total data); these false depth readings occur due to the logging of sonar data being started prior to launching the boat onto a pond. Depth measurements were converted to meters, then depths converted to sludge heights (i.e., the average depth of the pond minus the local water depth). GPS coordinates were converted to Universal Transverse Mercator (UTM) (for more details on GPS conversion, see Singh et al., 2008). It was assumed that the pond bottom surface was uniform. Coordinates of each measurement point were then defined relative to the lowest easting and northing values. Output measurement locations and sludge heights (m) were input into 3D surface mapping software Surfer (version 9.0, Golden Software Inc), to create a graph coordinate file (i.e., *xyz* file). This file was then run through the gridding toolbox to filter the data, where points were retained according to median $z$ values (sludge height) for any given $(x,y)$. Using a simple kriging interpolation, grids were generated at a spacing of 1 m in both $x$ and $y$, then used to create a 3D surface plot of the sludge. Overall, processing the data using this method takes between 30–60 min; this proved time consuming when there were several profiles to process, and could not be easily done on site just in case another profile needed to be taken (e.g., if there was an error with the sonar). In addition, these processing steps require the user to have a level of familiarity with 3–4 independent standalone software packages, making processing not user friendly. Furthermore, some of these software packages require a license, of which the cost may be prohibitive for some users, e.g., small water utilities. To overcome this issue, we developed a software package with open source tools to make the process more user-friendly, and significantly quicker.

As described in Coggins et al. (2017), the SludgePro software performs all of the filtering, analysis, and plotting of data, and can

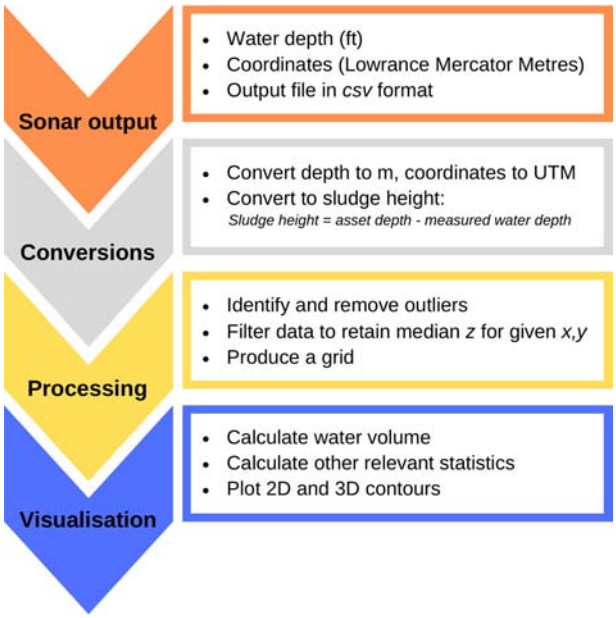



**FIGURE 3 |** Summary of the data processing required after profiles are collected with the ROV. These processing steps can be achieved manually, or automated using a script.

be used to produce a report suitable for use by pond managers. As a result of the development of this script, the analysis and plotting can be achieved in less than 30 s. The data processing involves a number of steps (as outlined above), including the conversion of geographical coordinates, the definition of pond boundaries, the removal of outliers or duplicated data, and visualization of data. Firstly, the script automates the conversion of the sonar map units to UTM. Using Google Earth, the pond boundaries are defined by drawing a path around the perimeter; this path is then saved as a Keyhole Markup Language (KML) file for input into the software. The software then uses the KML file to filter out points that are outside the pond boundary (i.e., those collected on land when the sonar logging is initiated). Due to the locational accuracy of the GPS being 1–6 m, there is also a tool for the user to input a shift of the data, so that all relevant data points can be included in the processing. Pond depth, from asset data, is also input into the software, and this value is used not only for the calculation of sludge volume, but also to remove the occasional outlier that is significantly out of the possible range of depth. These outliers typically occur at a frequency of 1 in 1000 points collected, and may be attributed to the boat rocking in the water while taking measurements. As the sonar has a high sampling rate, a significant number of duplicate data points are collected at each location. This amount of data per position is superfluous for the creation of a grid of the collected data. Therefore, the data is also processed to determine the median $z$ value for each location, which is then retained for gridding; the determination of the median also helps to remove the previously mentioned outliers that may occur. Water depth is converted to sludge height by subtracting the measured water depth (the direct measurement from the sonar) from the known asset depth, assuming that the pond bottom surface is uniform. In the case of the pond depth being unknown, SludgePro has the ability to process and visualize data based upon the measured water depth only, and thus values will not be filtered out based upon pond depth. The SludgePro frontend guides the user into providing the necessary information (i.e., KML file for pond perimeter, and csv files of collected data) to process and plot the survey data (Coggins et al., 2017). Most importantly, all the files and data can be processed, read, and modified by the user without the need for any programing knowledge.

## RESULTS AND DISCUSSION

### Initial Assessment of ROV Operation

Profiles of Ponds 1 and 2 were measured on four occasions between June and August 2011. Pond 1 was the first pond upon which the ROV was tested and more profiles were conducted here to ensure the reliability and reproducibility of profiling data. Each profile completed with the ROV took approximately 20 and 30 min for Ponds 1 and 2, respectively.

The profile of Pond 1 (**Figure 4A**) shows a reasonably uniform sludge distribution, consistent with it having been partially desludged by the pond managers in early 2011. The walls of the pond are the high points visible surrounding the edge of this plot. In comparison, the profile of Pond 2 (**Figure 4B**) shows more variability in sludge height. With the high-resolution data collected by the sonar, it is possible to spot the channel feature that has formed between the inlet and outlet; the average sludge height in this region is 0.3 m (color: dark blue). Sludge accumulation on the side of the pond adjacent to the inlet is also visible, with the change in sludge height from the channel to this region being abrupt (0.2 m higher than the channel itself). It is also evident that there is a very large accumulation of sludge in the southwest corner of the pond. Finally, highlighting the advantages of the increased spatial resolution of provided by the sonar, the profile shows pockets that have formed throughout the sludge blanket.

Reproducibility of sludge profile data collected was assessed by comparing data collected on different profiling days at Pond 1. Three profiles taken with the sonar were compared to the sludge judge measurement along the transect $y$ = 35 m (for sludge surface shown in **Figure 4A**), and all data were normalized against the sonar survey taken on the first profiling day. This comparison shows the reproducibility of the sonar profiling, with measurements on different days being within 5% range of each other over the overall depth (i.e., within 5 cm) (**Figure 5**). The small differences between the sonar surveys over the sampling period could be due to GPS positioning accuracy on the day, however, overall we see that the sonar technique is consistent in its measurement. The inconsistencies between values from the two different survey methods can be attributed to the "human factor" of sludge judge surveys; here, showing that the sludge judge survey overestimates the sludge height (contrary to **Figure 1**). Sludge judge survey accuracy relies on a number of factors including: (1) the experience of

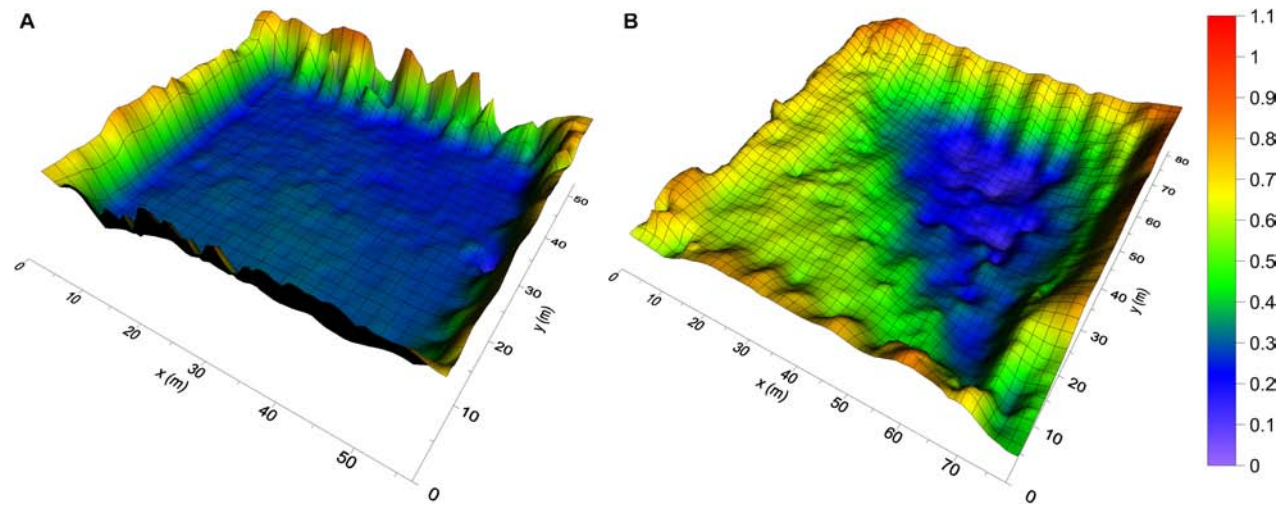



**FIGURE 4 |** Sludge profiles of test ponds. **(A)** Pond 1, inlet located at (52,48) and outlet at (1,44), has a reasonably uniform sludge distribution. **(B)** Pond 2, with inlet at (77,1) and outlet at (6,80), with a highly uneven sludge distribution, with a noticeable channel feature. The color scale indicates the height of the sludge-water interface (in meters) above the bottom of the pond.

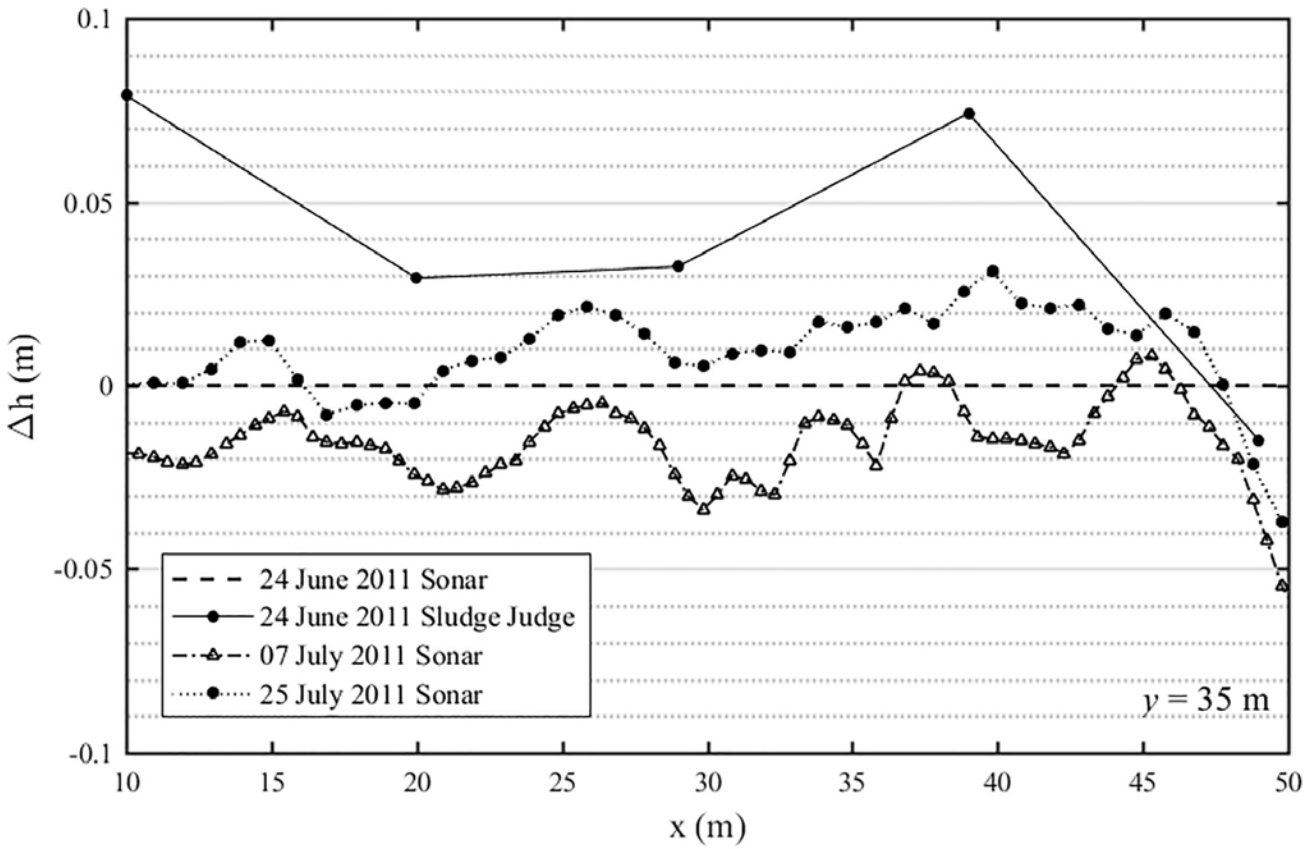



**FIGURE 5 |** A comparison of sludge height data collected at Pond 1 along transect $y$ = 35 m, on three different profiling days using both the sonar and sludge judge techniques, normalized and shown as a change in height (Δh) against the survey data collected on 24 June 2011 (Δh = 0). From this comparison, we can see that the sonar profiling technique is very comparable between surveys, with the departure only being 2–3 cm. In comparison to the sludge judge survey conducted on the same day, it can be seen that the sludge judge technique overestimated the amount of sludge in the pond. It should be noted here that these departures could be due to GPS positioning differences, however, overall, we can see that the sonar technique is consistent in measurement.

the operator; (2) the subjectivity and accuracy of their readings (and their readings compared to other operators); (3) whether or not the sludge blanket surface is more consolidated or "fluffy"; and (4) sample position relative to the marks on the side of the pond, where for example, drift of the boat could impact the accuracy of the reading along the transect used in **Figure 5**.

## Advantages of Autonomous Profiling

Compared to a sludge judge survey, the ROV not only reduced sludge profiling time, but also greatly increased the spatial resolution of the data collected. The ROV surveys of these ponds were completed in 20 to 30 min, while sludge judge surveys for Ponds 1 and 2 took between 1 to 1.5 h. Furthermore, the sludge judge surveys for these ponds yielded 25 and 42 data points, respectively, whereas the sonar surveys yielded 27192 and 30500 data points. Of the collected sonar measurements, 1513 data points for Pond 1 and 1886 for Pond 2 were mapped to generate the grids used to create the 3D surface plots; data points were filtered for gridding using the median $z$ value (i.e., depth) recorded at each GPS coordinate (i.e., the data points retained are unique, while duplicates are discarded). Using sludge judge, sludge volume estimates were calculated by using the average profiled sludge height, based on the number of data points taken. In comparison, estimates with sonar data were calculated with Simpson's 3/8 rule for numerical integration, and used all of the mapped data points rather than just the average.

Overall, the low-resolution measurement using sludge judge can only capture pond-scale features of the sludge distribution, while a sonar survey determines the pond/sludge bathymetry (**Figure 4**), showing much higher detail and resolution of the sludge blanket, including highlighting the presence of channels and pockets. The high-resolution data collected can then be used for other purposes, such as quantifying the relationship between sludge accumulation and hydraulics, and is suitable for input into computer models. Therefore, the testing of the boat satisfied the requirements of higher resolution sludge height data collection, and removed the safety risks of going out onto WSPs with a boat.

## Implementation by Australian Water Authorities and Safety

Over the past 8 years, several boat hull designs have been tested, however, all have been battery powered, and had a radio range of up to 200 m. Boat size has ranged between 800–1000 mm in length, and up to 400 mm in width, making them easily packed away into carrying cases, and easily transportable to and from site. The robustness of the boat has meant that is has now been used on >400 WSPs across Australia, ranging in size from those used in the testing phase (Ponds 1 and 2) to ponds/lagoons up to 2000 m in length. To date, four major Australian water utilities now use the ROV for sludge measurement in their WSP assets.

Overall, in addition to the operational reliability of the ROV, the boat also addressed the safety issues associated with sludge judge, and other on-pond profiling techniques. In particular, the development of the boat fits in with the "zero harm" safety policies of many Australian water utilities – many of which place high levels of safety practice when working in and around water. Due to the improvements in safety provided by the ROV, it has been shown as an exemplar for safety practice within Western Australia and nationally, having being nominated for and/or receiving several safety awards.

Part of any successfully used piece of equipment is a comprehensive, well-explained and easy to follow manual for use. Due to the safety requirements parties interested in the profiling boat, such as water utilities, government agencies, mining companies, it was necessary to invest time in writing an informative guide for safe usage. This would not only make it easier for others to use the boat, but also help to formulate the necessary risk assessments for permits to conduct ROV profiling of assets. The operation manual has been assessed and revised based on comments received from industry partners over the past 5 years. The manual includes sections on a quick start on site, information about the boat components, operation and maintenance, troubleshooting, data processing, and a manual for the use of SludgePro. Additional to that provided in the manual, step-by-step guides for adding and processing ponds have also been made for users of SludgePro. These have been designed assuming that the user has no prior knowledge of programing. These guides, along with the quick start on site, have been very successful in aiding knowledge transfer to users.

## Surveys

### WSP Sludge Management

The management of sludge is rarely considered in the pond design process, despite the inevitability of sludge accumulation in WSPs (Nelson et al., 2004). Reasons for overlooking sludge management in design include the lack of information about sludge distribution within ponds, sludge characteristics, and accumulation rates (Nelson et al., 2004). Sludge distribution within ponds is of particular interest and importance, as this can have a significant impact on pond hydraulics and treatment efficiency. More information and understanding about how sludge distribution affects pond characteristics will lead to better informed maintenance decisions by ponds managers, and could lead to design improvements.

The high-resolution data collected by the sonar is a significant improvement on traditional sludge profiling techniques, with output images clearly showing the formation of channels and areas of high sludge (e.g., **Figure 4B**). In addition, due to ability to collect the high-resolution data with the ROV rapidly, it is possible to use it for diagnostic purposes, such as in the case of operators observing abnormal hydraulic characteristics in pond after sludge removal. An example of using the ROV as a diagnostic tool is shown in **Figure 6**, which clearly indicates the areas where sludge has been removed. It was then determined that the sub-optimal sludge removal, coupled with the large decants of incoming water that this pond receives several times a day, had created a scour effect around the inlet end, resulting in the formation of a 10 m wide U-shaped channel, explaining the abnormal hydrodynamic behavior being observed by pond managers.

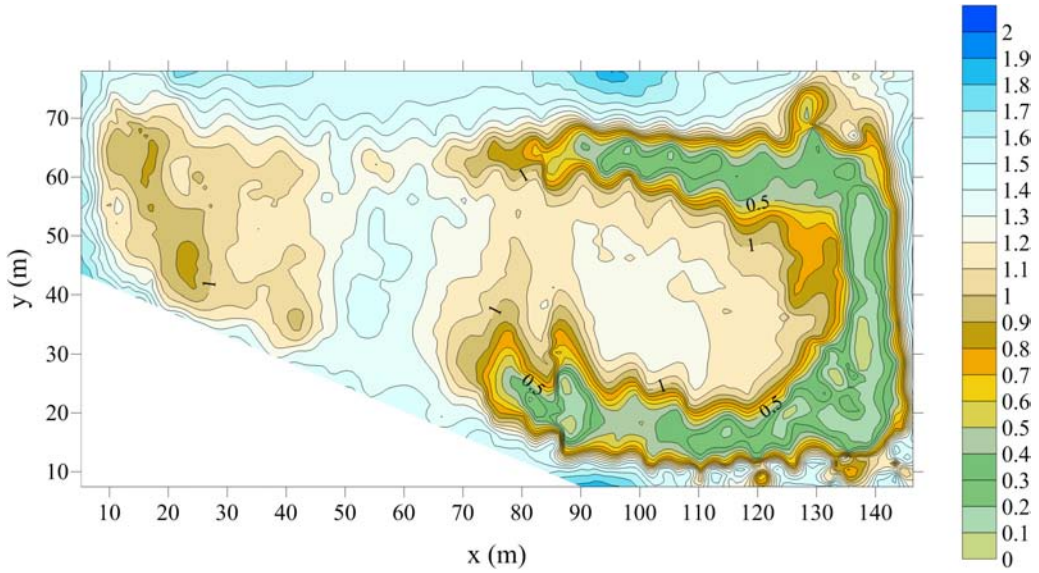



**FIGURE 6 |** Sludge bathymetry in a pond that was reported to be displaying abnormal hydraulic behavior. Color scale indicates sludge height from the bottom of the pond in meters. The most significant feature in this profile is the U-shaped channel that has formed around the eastern edge of the pond. The inlet is located at approximately (142,70).

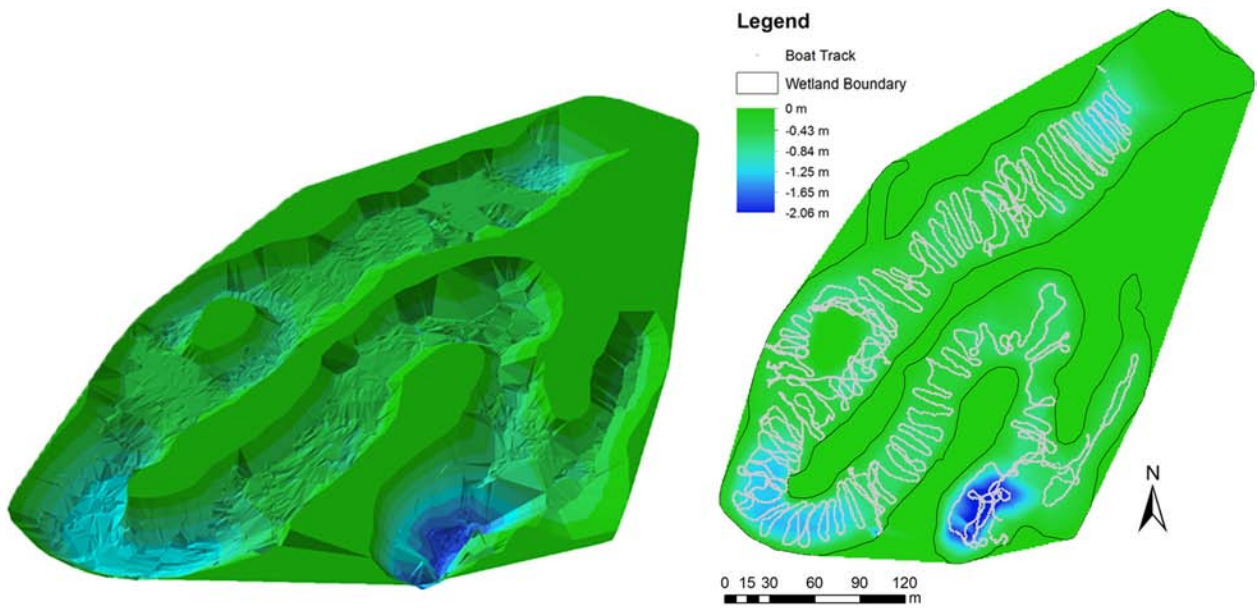



**FIGURE 7 |** Bathymetric survey of the main wetland of Troups Creek, 3D (left) and plan view with scale and boat survey track (right), both showing water depth (Image: P. M. Bach).

This example makes it clear that the high-resolution bathymetric data collected using sonar profiling is an extremely useful tool for the determination of sludge distribution in ponds. Using traditional profiling techniques, which can only capture pond-scale features, this discovery could have easily been missed, as channel features could be at smaller scales than the discreet sample spacing. Moreover, it shows that the sonar can be used high-resolution diagnostic tool to understand the hydraulics in ponds. Most importantly, this critical high-resolution data can be collected without the need to go out onto WSPs in a boat, addressing several safety considerations. The portability and convenience of this technology could be applied on ponds on a more frequent basis (e.g., monthly vs. yearly), to collect valuable data on sludge accumulation rates.

Previously bathymetric data from traditional sludge surveys in WSPs have been used to create computational fluid dynamics

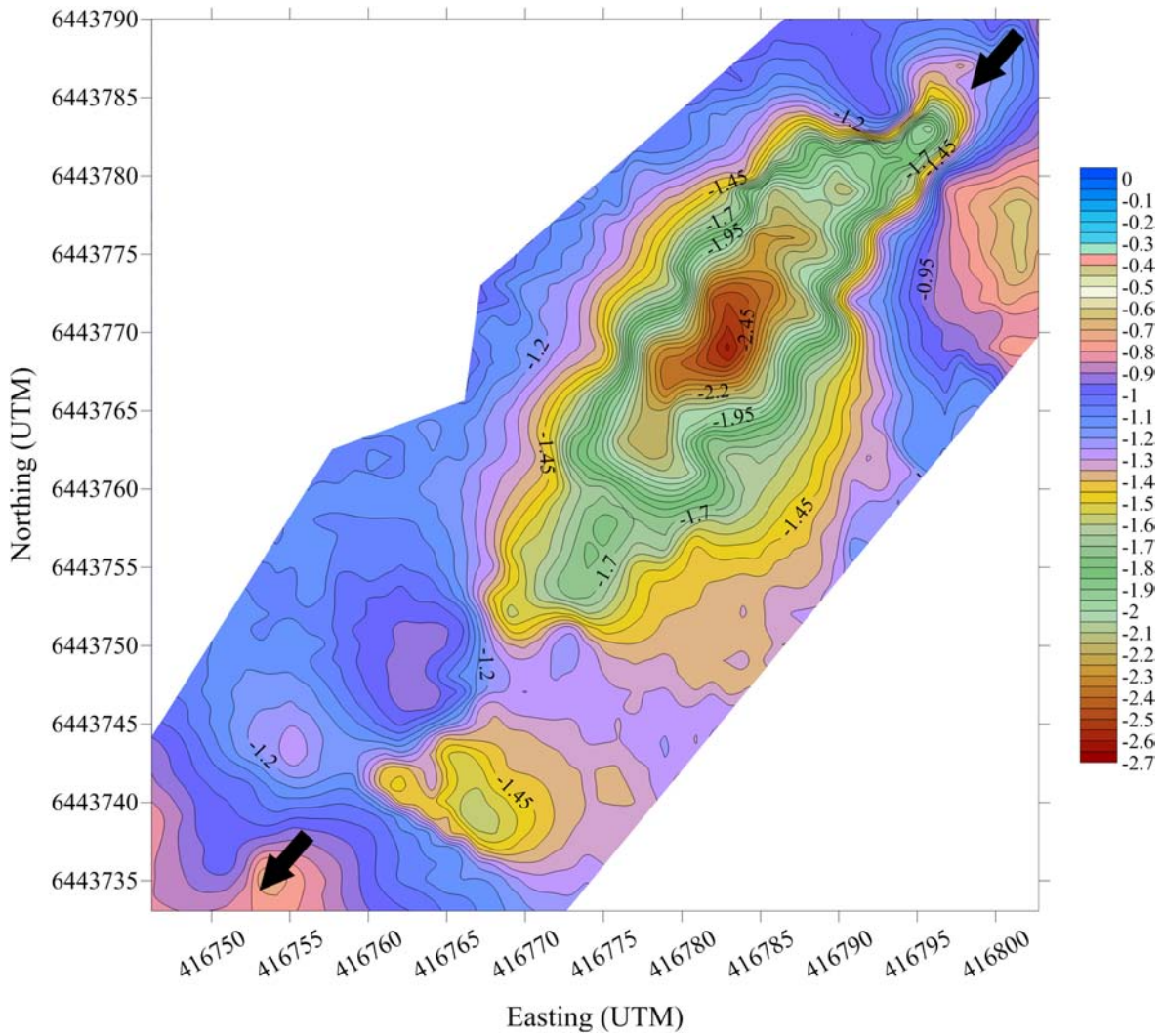



**FIGURE 8 |** Bathymetry of a pool along the Canning River. Color scale indicates water depth in meters and the arrows indicate the direction of flow.

(CFD) models with varying levels of success (Olukanni and Ducoste, 2011; Alvarado et al., 2012a, 2013; Sah et al., 2012; Passos et al., 2014). Due to low-resolution of data input as bathymetry, these models have been harder to validate against actual conditions. Several studies have found that higher resolution data would significantly improve the accuracy to CFD models of WSPs (Daigger, 2011; Alvarado et al., 2012a); with the increased spatial resolution provided by profiling with sonar being an ideal solution to this. The increase in bathymetric resolution, coupled with tracer test data, will allow for better calibration and validation of models (Sah et al., 2011; Alvarado et al., 2012b; Passos et al., 2016), and will increase the reliability of model outputs. In turn these improvements to models will increase our understanding of WSP systems, and allow for more accurate modeling of the effects of pond installations, such as baffles, as well as aid in the management of ponds, including desludging.

Applying this to sludge management, a CFD model could be used to determine the differences between different sludge infill scenarios, for example, the current operating sludge distribution in a pond versus the same pond with no sludge; this has been demonstrated in Coggins et al. (2017). Modeling with high-resolution bathymetric data in WSPs is very promising, and with the addition of wind forcing (such as in: Shilton and Harrison, 2003), will become a useful diagnostic and predictive tool for existing and new systems, respectively; this could also extent to the design of baffles for existing/new systems, as demonstrated in Coggins et al. (2018). Therefore, the addition of high-resolution bathymetry data is a step in the right direction for more spatial accuracy in the CFD modeling of WSP systems.

The idea for this project was a grassroots idea from the operational level of our water industry partner. Since the inception of the project, it has relied upon the input of a large community ranging from on-site operators to executives. Unforeseen by us, and as a pleasant surprise, this initially small community has developed into a Community of Practice

(CoP) across Australia, which has provided us continuous and ongoing support, as well as suggestions for improvement to the hardware, software, and training. One of the benefits of the growing CoP is the expansion of the application of the ROV beyond WSPs and the water industry, into a range of applications by local and state government, mining companies, and researchers. Recently, the CoP has extended into the upgrade and software architecture redesign of the SludgePro software through a university innovation initiative. This significant upgrade is expected to result in an improved product for analyzing and storing data, as well as possibly including machine learning to aid managers in decision making and forecasting.

### Stormwater Wetlands

In addition to being used extensively on WSPs, the boat has been used on a series of stormwater retention wetlands in Melbourne, Australia. Rather than using the data collected to infer sludge height in ponds, the boat was used for bathymetry surveys of three basins along Troups Creek. For these surveys SludgePro is not used to analyze data, however, the same methodology that is used in SludgePro was used (as outlined in section "Data Processing and Analysis"). Filtered data (using a Python script) was input into ArcGIS to create a Triangular Irregular Network (TIN) for visualization. This shows that there are several different ways to analyze, visualize, and interpret collected data. The high-resolution bathymetry data (**Figure 7**) was also then able to input into a flow model of the creek network, to assist with the other research being carried out.

### Lakes

Rottnest Island, approximately 20 km off the coast of Western Australia, has a series of environmentally significant salt lakes, and despite a long history of people inhabiting the island, and it being a popular tourist destination, a comprehensive bathymetric survey of the lakes had never been conducted. The Rottnest Island Authority deployed the remote control boat on this series of lakes over a 1-month period in November/December 2015. The collection of this data was vital for the island authority in order to protect the unique environment of these lakes from the impacts of tourism and development on the island. The profiling of these lakes was a significant test for the boat, as these lakes were significantly larger than any of the WSPs previously profiled, with the largest lake approximately 1900 m in length, and 650 m in width. Eight lakes on the island were surveyed, and the data processed and analyzed using the statistical tools available in the ArcGIS package.

### River Pools

The Canning River is a major tributary of the Swan River, in the southwest of Western Australia. In the upper reaches of the river, there are many small to medium sized natural pools. Anecdotal evidence from regular users of the river, and the River Guardians (State Government Department of Biodiversity, Conservation and Attractions, Western Australia), is that these pools are ecologically significant; however, the effect of natural sediment build-up in these pools is unknown. Due to its size and portability, the ROV with sonar is well suited for conducting surveys of these pools, with 10 pools having been profiled since mid-2015 (e.g., **Figure 8**).

## CONCLUSION

From this study, we can conclude that the remote control boat (or ROV) successfully measures sludge distribution with high-resolution. This then allows the construction of detailed 2D and 3D plots of the sludge blanket, showing the formation of features such as channels and pockets. Most importantly, this method allows the collection of high-resolution bathymetric data without going onto WSPs in a boat, addressing several important safety considerations. The ROV is a reliable tool that has been successfully deployed on over 400 Australian WSP of various geometries and sludge distributions, stormwater retention wetlands, lakes, and river pools. Our ability to obtain sludge distribution and accumulation data rapidly will prove invaluable in the future. This technology will help in the development of frameworks for wastewater sludge management, and could potentially have a wider application in the monitoring of other small to medium water bodies, including reservoirs, channels, recreational water bodies and commercial ports.

## AUTHOR CONTRIBUTIONS

LC and AG designed the equipment and experiments. LC collected the field data and completed the statistical analysis. LC wrote the manuscript with input from AG.

## FUNDING

This work was supported by the Water Corporation of Western Australia and an Australian Research Council Grant (LP130100856). LC was supported by a Prescott Postgraduate Scholarship and Research Impact Grant from The University of Western Australia and a TasWater Wastewater Engineering Scholarship.

## ACKNOWLEDGMENTS

We would like to thank A. Chua, D. Italiano, T. Rintoul, S. McPhee, B. Kerenyi, and K. Eade from the Water Corporation for all their help with fieldwork, suggestions, and sourcing information. At UWA we would like to thank F. Tan for carrying out modifications to the early boats, D. Stanley for electronics advice and help, J. J. Langan for repairs, maintenance, and enthusiastic assistance in the field, A. Stubbs for boat repairs, and E. S. Reichwaldt for advice and help with fieldwork. We would also like to thank Richard from Stanbridges Hobbies for early sourcing of different hulls and electrical work and repairs, S. Stratfold for building us robust fiberglass boat hulls in bulk, C. Bosserelle for writing the SludgePro script and design of the frontend, and S. A. A. Shah for modifications to the scripts.

## REFERENCES


Alvarado, A., Sanchez, E., Durazno, G., Vesvikar, M., and Nopens, I. (2012a). CFD analysis of sludge accumulation and hydraulic performance of a waste stabilization pond. *Water Sci. Technol.* 66, 2370–2377. doi: 10.2166/wst.2012.450

Alvarado, A., Vedantam, S., Goethals, P., and Nopens, I. (2012b). A compartmental model to describe hydraulics in a full-scale waste stabilization pond. *Water Res.* 46, 521–530. doi: 10.1016/j.watres.2011.11.038

Alvarado, A., Vesvikar, M., Cisneros, J. F., Maere, T., Goethals, P., and Nopens, I. (2013). CFD study to determine the optimal configuration of aerators in a full-scale waste stabilization pond. *Water Res.* 47, 4528–4537. doi: 10.1016/j.watres.2013.05.016

Bandini, F., Olesen, D., Jakobsen, J., Kittel, C. M. M., Wang, S., Garcia, M., et al. (2018). Technical note: bathymetry observations of inland water bodies using a tethered single-beam sonar controlled by an unmanned aerial vehicle. *Hydrol. Earth Syst. Sci.* 22, 4165–4181. doi: 10.5194/hess-22-4165-2018

Coggins, L., Sounness, J., Zheng, L., Ghisalberti, M., and Ghadouani, A. (2018). Impact of hydrodynamic reconfiguration with baffles on treatment performance in waste stabilisation ponds: a full-scale experiment. *Water* 10:109. doi: 10.3390/w10020109

Coggins, L. X., Ghisalberti, M., and Ghadouani, A. (2017). Sludge accumulation and distribution impact the hydraulic performance in waste stabilisation ponds. *Water Res.* 110, 354–365. doi: 10.1016/j.watres.2016.11.031

Daigger, G. T. (2011). A practitioner's perspective on the uses and future developments for wastewater treatment modelling. *Water Sci. Technol.* 63, 516–526. doi: 10.2166/wst.2011.252

Ghadouani, A., and Coggins, L. X. (2011). Science, technology and policy for water pollution control at the watershed scale: current issues and future challenges. *Phys. Chem. Earth Parts A/B/C* 36, 335–341. doi: 10.1016/j.pce.2011.05.011

Glaz, P., Bartosiewicz, M., Laurion, I., Reichwaldt, E. S., Maranger, R., and Ghadouani, A. (2016). Greenhouse gas emissions from waste stabilisation ponds in Western Australia and Quebec (Canada). *Water Res.* 101, 64–74. doi: 10.1016/j.watres.2016.05.060

Hernandez-Paniagua, I. Y., Ramirez-Vargas, R., Ramos-Gomez, M. S., Dendooven, L., Avelar-Gonzalez, F. J., and Thalasso, F. (2014). Greenhouse gas emissions from stabilization ponds in subtropical climate. *Environ. Technol.* 35, 727–734. doi: 10.1080/09593330.2013.848910

Kaizu, Y., Iio, M., Yamada, H., and Noguchi, N. (2011). Development of unmanned airboat for water-quality mapping. *Biosyst. Eng.* 109, 338–347. doi: 10.1016/j.biosystemseng.2011.04.013

Kriechbaumer, T., Blackburn, K., Everard, N., and Rivas Casado, M. (2015). Acoustic Doppler current profiler measurements near a weir with fish pass: assessing solutions to compass errors, spatial data referencing and spatial flow heterogeneity. *Hydrol. Res.* 47, 591–605.

Mara, D. (2004). *Domestic Wastewater Treatment in Developing Countries*. London: Earthscan.

Nelson, K. L., Cisneros, B. J., Tchobanoglous, G., and Darby, J. L. (2004). Sludge accumulation, characteristics, and pathogen inactivation in four primary waste stabilization ponds in central Mexico. *Water Res.* 38, 111–127. doi: 10.1016/j.watres.2003.09.013

Olukanni, D. O., and Ducoste, J. J. (2011). Optimization of waste stabilization pond design for developing nations using computational fluid dynamics. *Ecol. Eng.* 37, 1878–1888. doi: 10.1016/j.ecoleng.2011.06.003

Passos, R. G., Dias, D. F. C., and von Sperling, M. (2016). Review of practical aspects for modelling of stabilization ponds using computational fluid dynamics. *Environ. Technol. Rev.* 5, 78–102. doi: 10.1080/21622515.2016.1251499

Passos, R. G., von Sperling, M., and Ribeiro, T. B. (2014). Hydrodynamic evaluation of a full-scale facultative pond by computational fluid dynamics (CFD) and field measurements. *Water Sci. Technol.* 70, 569–575. doi: 10.2166/wst.2014.265

Peña, M. R., Mara, D. D., and Sanchez, A. (2000). Dispersion studies in anaerobic ponds: implications for design and operation. *Water Sci. Technol.* 42, 273–282. doi: 10.2166/wst.2000.0660

Persson, J. (2000). The hydraulic performance of ponds of various layouts. *Urban Water* 2, 243–250. doi: 10.1016/S1462-0758(00)00059-5

Picot, B., Sambuco, J. P., Brouillet, J. L., and Riviere, Y. (2005). Wastewater stabilisation ponds: sludge accumulation, technical and financial study on desludging and sludge disposal case studies in France. *Water Sci. Technol.* 51, 227–234. doi: 10.2166/wst.2005.0469

Sah, L., Rousseau, D. P. L., and Hooijmans, C. M. (2012). Numerical modelling of waste stabilization ponds: where do we stand? *Water Air Soil Pollut.* 223, 3155–3171. doi: 10.1007/s11270-012-1098-4

Sah, L., Rousseau, D. P. L., Hooijmans, C. M., and Lens, P. N. L. (2011). 3D model for a secondary facultative pond. *Ecol. Model.* 222, 1592–1603. doi: 10.1016/j.ecolmodel.2011.02.021

Shilton, A., and Harrison, J. (2003). Development of guidelines for improved hydraulic design of waste stabilisation ponds. *Water Sci. Technol.* 48, 173–180. doi: 10.2166/wst.2003.0114

Singh, K., Worley, J., and Risse, M. (2008). Sludge measurement using global positioning system (GPS) enabled sonar equipped airboat in a lagoon. *Appl. Eng. Agric.* 24, 603–609. doi: 10.13031/2013.25275

Westerman, P. W., Shaffer, K. A., and Rice, J. M. (2008). *Sludge Survey Methods for Anaerobic Lagoons*. Raleigh, NC: North Carolina State University.


**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.

*Copyright © 2019 Coggins and Ghadouani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers in Earth Science

BRIEF RESEARCH REPORT
published: 17 December 2019
doi: 10.3389/feart.2019.00338

# A User-Printable Three-Rate Rain Gauge Calibration System

*Jose M. Lopez Alcala[1,2], Chester J. Udell[1,3] and John S. Selker[1,3]**

[1] Openly Published Environmental Sensing Lab, Department of Biological and Ecological Engineering, College of Agricultural Sciences, Oregon State University, Corvallis, OR, United States, [2] School of Electrical Engineering and Computer Science, College of Engineering, Oregon State University, Corvallis, OR, United States, [3] Department of Biological and Ecological Engineering, College of Agricultural Sciences, Oregon State University, Corvallis, OR, United States

Our objective was to develop and validate a freely downloadable, open-source, 3D printed rain gauge calibrator that can be adjusted for a wide range of gauges. The calibrator applies constant low, medium, and high-intensity water delivery rate, and allows the user to modify the design to conform to their system based on parametric design. The design may be modified and printed using freely available computer-aided design (CAD) software. Currently available devices for calibration tend to be designed for specific rain gauges, are expensive, employ low-precision water reservoirs, are not field portable, and do not offer the flexibility needed to test the ever more popular small-aperture rain gauges (smaller surface area to catch precipitation than the classical 200 mm standard). To overcome the fact that different 3D printers yield different print qualities, we devised a simple post-printing step that controls critical dimensions to assure robust performance. Specifically, orifices of the calibrator are drilled to reach the target flow rates. Laboratory tests showed that flow rates of 25, 50, and 83 ml/min were consistent between prints (coefficient of variation of 3.9, 2.2, and 1.8%, respectively), and between trials of each part, while the total applied water was precisely controlled (0.1%) by the use of a volumetric flask as the reservoir. The entire system costs under US$10.

**Keywords: calibration, rain gauge, 3D printing, OPEnS lab, Mariotte bottle, computer-aided design, ABS plastic**

OPEN ACCESS

***Edited by:***
*Rolf Hut,*
*Delft University of Technology,*
*Netherlands*

***Reviewed by:***
*Benjamin Michael Clemens Fischer,*
*Stockholm University, Sweden*
*Witold F. Krajewski,*
*The University of Iowa, United States*
*Jan Friesen,*
*Helmholtz Centre for Environmental Research (UFZ), Germany*

***Correspondence:**
*John S. Selker*
*john.selker@oregonstate.edu*

***Specialty section:***
*This article was submitted to Hydrosphere, a section of the journal Frontiers in Earth Science*

***Received:*** *01 November 2018*
***Accepted:*** *02 December 2019*
***Published:*** *17 December 2019*

***Citation:***
*Lopez Alcala JM, Udell CJ and Selker JS (2019) A User-Printable Three-Rate Rain Gauge Calibration System. Front. Earth Sci. 7:338. doi: 10.3389/feart.2019.00338*

## INTRODUCTION

Rain gauges are essential tools for high-quality and reliable observation of precipitation. They are the most direct method for surface rainfall quantification, as utilized for hydrological, climatological, and agricultural studies (Habib et al., 2012). The following question arises concerning the validity of rain gauges: How do we know that the sampler is correctly reporting the rainfall amount. To validate the integrity of the rain gauge requires the application of a known volume of water at a known constant flow rate. It is essential to do dynamic calibration, calibration that requires a constant flow rate, because it allows for compensation of error cause by spillage (Ciach, 2003; Texas Electronics, 2019). Spillage occurs when a tipping bucket is not able to capture the rainfall due to the transition between the tipping, the rain is then spilled and

not measured. Such devices already exist but have important limitations. Some employ simple gravity fed inverted bottles wherein the rate of flow decreases with the depletion of the water supply, and thus do not provide a constant rate of application. Others are tuned to specific size rain gauges, applying rates of flow and requiring a mechanical attachment that is not adaptable to alternative sizes and shapes of rain collectors. An example of this ATMOS 41 rain gauge's (Meter group, Pullman, WA, United States) collection area smaller than those for which calibration tools had been developed, requiring that we develop a design capable of delivering lower rates of application. Custom calibrators may include a setup of programmable pumps, loggers, digital scales, and computers to be able to perform the calibration of the rain gauge (Humphrey et al., 1997). With numerous external instruments, it is time consuming to organize and reprogram all the devices required to calibrate different rain gauges.

We sought to provide a robust system with a repeatability within less than 1% in total amount and 5% of the target rate that can be modified to match a variety of rain gauge specifications and geometries. The repeatability of 5% was chosen because it is an achievable target given the technologies used to make the calibrator: 3D printing and drilling. We chose 740 mm/h (high), 450 mm/h (medium), and 220 mm/h (low); these will be referred to as 6-, 10-, and 20-min stoppers, respectively, which represent the time to drain for our setup. Three options were chosen so that the user had a range of rates to calibrate their device. These rates were selected with the objective to be able to capture data in the events of high intensity rainfall. We believe that it is in these events during which the data collection must be performed more carefully. In addition to this, it is not quite possible to design a calibrator with extremely fine resolution due to the elevated cost of the finding the equipment to manufacture it. We sought to remove the constraints of the high cost of many calibration solutions and the requirement to return a rain gauge to the lab to conduct calibration (Bergmann et al., 2001; Vasvári, 2005). Thus, we present a low-cost rain gauge calibration system which is easily customized in geometry to fit a wide range of gauges, provides for a user-definable range of constant flow rates, and most importantly is easily employed in the field. It is imperative that rain gauges are not used out-of-the box as they need proper calibration so that the collected data is correct. Our proposed solution removes the need to perform the calibration in the lab and allows the user to perform these calibrations while in the field; this reduces the time to setup and speeds up the data collection process.

A robust passive approach that delivers a constant flow rate is the Mariotte bottle (Mariotte, 1679). A typical Mariotte bottle setup has a container with a tube coming in from the top for air and another orifice for the liquid output usually on the side or through the top (**Supplementary Figure 1**). The flow rate is dictated by the following equation: $h_{air} - h_{output} = h$. $h_{air}$ is the level from the bottom of the container to the air inlet inside the container and $h_{output}$ is the level from bottom of the container to the water outlet. The difference between these two is net head which can be seen depicted in **Supplementary Figure 1**. The bottle is designed to deliver a constant flow of liquid, which is a function of the distance between the bottom of the air inlet and liquid outlet, orifice sizes, and the hydraulic resistance.

Our design employs the Mariotte bottle principle (**Figure 1A**) with the bottle inverted and both tubes formed into an O-ring sealed element we refer to as "Mariotte stoppers" that seal into the top of a volumetric flask. The materials list for seals and bottles can be found in **Supplementary Table 4**. Multiple flow rates were achieved in the same device by placing three Mariotte stoppers on a single plate that rests in the rain gauge (**Figures 1A,B**). Each Marriot stopper is cylindrical with two orifices on the top, a short-path air inlet and a long-path liquid outlet, with half of the cone suppressed toward the core; the suppressed section goes from the core to the outside in the direction of the air inlet (**Figure 1C**). The suppression is to facilitate the flow of air into the bottle and limit the path it must travel to enter. The side that is not suppressed gives the calibrator the change in height, from the air inlet to liquid outlet, required to dictate the flow rate (**Figure 1C**). The rate of outflow is dictated by the combination of the constant head, hydraulic resistance, and the outlet aperture. Except for the first few seconds of operation, the delivered water is in the form of discrete drips, and so the device could be used with disdrometers as well as other mechanical rain gauges.

## MATERIALS AND METHODS

The three-rate calibrator was designed in Fusion 360 (Fusion360, 2018) which provides a freely available, user-friendly interface which allows the design to be edited per the requirements of a particular rain gauge. The design is parametric allowing the user to change defined variables that will change the design to be adapted to a classic round rain gauge. The design also includes the individual calibrator, without the base attachment, that can be used to create a custom shape base for a different geometry rain gauge. The reservoir of water employed in the test was a 500 ml (±0.5 ml, or ±0.1%) volumetric flask made from polypropylene, as purchased online for under US$5. Volumetric flasks of 100, 250, and 1000 ml are available with the same size openings, allowing a variety of total volumes of water delivery. This specific bottle does not need to be used but it will require customization if another bottle is used.

At the OPEnS lab[1], we employed two 3D printers to confirm multi-platform production capability: a Lulzbot TAZ5 and Fusion3 F400. The Fusion3 F400 uses a 0.4 mm nozzle and used a 1.75 mm diameter filament. The slicer used was Simplify3D and used their F400_0.4_HatchABS printing profile with the auto-configure option "Standard." The Lulzbot TAZ5 uses a 0.5 mm nozzle and used a 3.0 mm diameter filament. The slicer software used was Cura Lulzbot edition and used their ABS (Village Plastics) printing profile with the auto-configure option "Standard." Important printer settings are listed in **Supplementary Tables 1, 2** for the Fusion3 and TAZ5, respectively. The layer height set by the profiles is adequate for a working calibrator. A smaller layer height will produce a smoother part, whereas a thicker layer height will have a coarse

[1]www.open-sensing.org

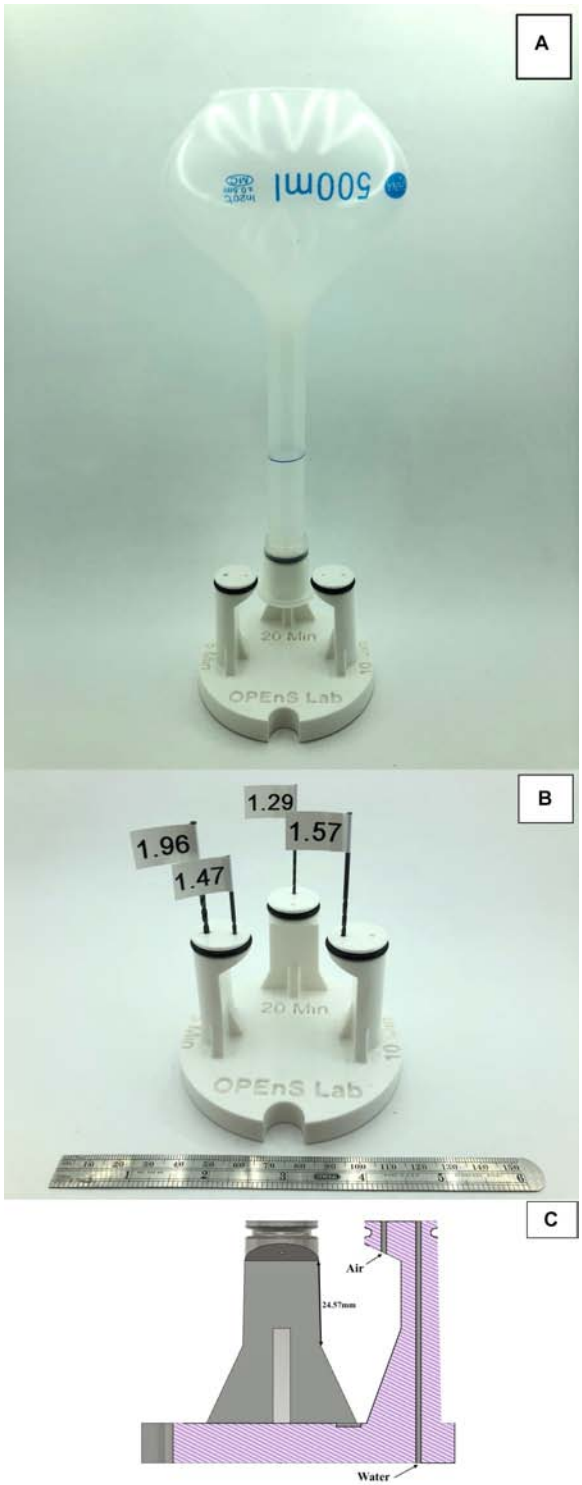



**FIGURE 1 | (A)** Assembled three-rate assembly system; the circular cutout in this version accommodates the solar radiation sensor located within the rain gauge for which this was designed. **(B)** A particular print of the design illustrates the limitations of many printers to achieve a smooth surface and presents the drill-bits sizes, in mm, employed to obtain reproducible performance. **(C)** CAD rendering of the design illustrates that each stopper has an air inlet port and water outlet port.

finish. Generally, having a smaller layer height will result in a better part, but ultimately the accuracy and precision of the calibrator is controlled by the post-processing steps. All parts were printed from Acrylonitrile Butadiene Styrene (ABS) plastic filament. ABS plastic is considered a good engineering plastic and it was chosen for its structural stability, impact resistance, and price. It is worth noting that there are other filaments that could be used for this application but ABS offers a good balance between price and material properties. After cooling, the precise diameter of the water outlet and air inlet apertures were established by drilling out the excess plastic using a drill press and drill bits of the sizes in **Supplementary Table 3**. It is recommended to use a drill press and not a hand drill for the post-processing.

Once the drill bit has been locked in on the drill press the calibrator's corresponding hole must be aligned to the drill. For example, if the 6-min setting is being drilled out, a 1.47 mm drill bit would correspond to the air hole and a 1.96 mm drill bit would correspond to the water hole. Two methods for the post-processing with the drill were tested on the 20-min stoppers. For the first method, used on stoppers 1 and 2, the drill was inserted one time and it penetrated as far as it could go. In the more careful method, for stoppers 3 and 4, the drill was allowed only to enter a small bit and then taken out for cleaning: this way the drill was only cutting into new material and not heating excessive residue inside the orifice. All of the drilling should be done at the same speed and at the recommend speed for the specific machine used. **Figure 1B** shows the drill bits' diameters, in millimeters, used to drill out the holes. The 6-min setting was drilled with a 1.47 mm drill bit for the air inlet and a 1.96 mm for the water outlet. The 10-min setting employed a 1.57 mm (1/16-in) drill bit for both orifices. The 20-min setting had the air inlet drilled with a 1.29 mm and a 1.57 mm (1/16-in) for the water outlet. These hole sizes and configurations were experimentally determined to approximate the target times and are summarized in **Supplementary Table 3**. **Supplementary Figures 2–7** illustrate the iterative design steps used to prototype the final result.

## RESULTS

**Figure 2** present the results of the flow rates of the 6-, 10-, and 20-min setting. Tests of the 6-min stoppers resulted with an average time of 6.28 min, 0.11 min standard deviation, and 0.018 coefficient of variation. The small standard deviation and coefficient of variation indicate that the post-processing is eliminating most of the variation is produced by the 3D printer's inability to reproduce the same piece. Tests of the 10-min stopper resulted in with an average of time of 10.29 min, with a standard deviation of 0.23 min and a coefficient of variation of 0.022. Tests of the 20-min stoppers resulted with an average time of 19.76 min, standard deviation of 0.78 min, and coefficient of variation of 0.04. The last two stoppers also demonstrate the same response to the post-processing. It reduces the variation between each print and gives the stoppers consistent performance. This data was created using stoppers from both printers, the Fusion3 F400 and Lulzbot 5. It is apparent that regardless of where the part came

| Time Setting (min) | Average Time (min) | Standard Deviation | Coefficient of Variability |
|---|---|---|---|
| 30 | 29.8 | 7.59 | 25.54 |
| 20 | 19.76 | 0.78 | 3.9 |
| 10 | 10.29 | 0.23 | 2.2 |
| 6 | 6.28 | 0.11 | 1.8 |

**FIGURE 2 |** Table that holds the results of 30, 20, 10, and 6 min stoppers showing the performance of each.

from, the performance is consistent after the post-processing. To test the importance of post-printing adjustment of the hole sizes, we tested an "as-printed" 30-min design finding variation both between prints and tests. The results we obtained were an average time of 29.8 min, standard deviation of 7.59, and standard error of 2.53. The finishing step of drilling the apertures for the 6-, 10-, and 20-min stoppers resulted in consistent behavior, while the undrilled 30-min stopper gave flow rates which varied greatly between prints.

## DISCUSSION

There was still variability in the rates even with careful drilling of all of the orifices, but it is significantly reduced compared to the stoppers that are not drilled after printing. The stoppers demonstrate average times that fall within 5% of the target time, but individual tests had a variability up to approximately 8%. It can be concluded that post-processing is needed to achieve the desired rates and that different post-processing methods such as the ones described above yield different accuracy. The results demonstrate that our targets can be approximated if the post-processing is done carefully. For this reason, the 30-min stoppers were not further developed as they require very specialized equipment to approximate the rate and it would have created a barrier to create such a device. The quality of a stopper can be attributed to the post-processing methods. If the drill is cleaned every time it is inserted, then the drilling will be cleaner as it will not be drilling excessive material as well as overheating the plastic. The results indicate that the variability of the physical object produced by the 3D printing process due to the 3D printer model, environmental conditions, slicer program, and print settings can be compensated for by careful post-processing. To obtain lower flow rates, and potentially more consistency, one could establish the water delivery tube diameter using sealed-in glass capillary tubes, which are available in a wide range of sizes.

## CONCLUSION

Even with high performance Fused Filament Fabrication (FFF) printers, such as the Fusion3 F400, the orifice sizes do not come out to the design specification in the CAD: the resolution is not accurate enough. Drilling is the most exact way to get the results that are needed across all 3D printers. An alternate approach to the drilling process is to use reamers; they offer better consistency across the whole length of the hole because of the structure of the tool. This offers more accuracy to the hole and creates a more precise calibrator.

When the small apertures in the device were drilled post-printing, the multi-rate rain gauge calibrator performs on average within 5% of target time and can be readily modified, printed and employed in the field. The post-processing will not always yield a stopper that is within 5%; this can be attributed to the variability of the post-processing and human error. We demonstrate that without post-printing drilling, small orifices will have enough variability between prints and between printers to yield unacceptable performance. We also note that the most critical aspect of calibration of the rain gauge, that a known total volume of water is applied, is guaranteed to less than 0.1% deviation through use of a calibrated volumetric flask, purchased for this effort for under US$5 and made from high-impact polypropylene. Use and transport of this flask over 9 months in five countries has confirmed that it is amply sturdy for field use.

## AUTHOR CONTRIBUTIONS

JS contributed to the concept and design of the project. JL prototyped and validated the designs and wrote the first draft of the manuscript. JL, JS, and CU performed design revisions and contributed to the manuscript revision and read and approved the submitted version.

## FUNDING

This work was supported by the USDA National Institute of Food and Agriculture, Hatch project NI18HFPXXXXXG055.

## ACKNOWLEDGMENTS

The authors would like to thank the OPEnS Lab for providing the space and tools necessary to develop this project and to all members of the OPEnS Lab for all their support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00338/full#supplementary-material

## REFERENCES

Bergmann, H., Breinhalter, H., Hable, O., and Krainer, R. (2001). Calibration of tipping bucket hyetographs. *Phys. Chem. Earth, Part C Solar, Terr. Planet. Sci.* 26, 731–736. doi: 10.1016/S1464-1917(01)95017-2

Ciach, G. J. (2003). Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol.* 20, 752–759. doi: 10.1175/1520-0426(2003

Fusion360 (2018). *Autodesk*. Available at: https://www.autodesk.in/products/fusion-360/overview#generative-design (accessed October 21, 2018).

Habib, E., Haile, A. T., Tian, Y., and Joyce, R. J. (2012). Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates. *J. Hydrometeorol.* 13, 1784–1798. doi: 10.1175/JHM-D-12-017.1

Humphrey, M. D., Istok, J. D., Lee, J. Y., Hevesi, J. A., and Flint, A. L. (1997). A new method for automated dynamic calibration of tipping-bucket rain gauges." *J. Atmosph. Ocean. Technol.* 14:1513. doi: 10.1175/1520-0426(1997)014<1513:anmfad>2.0.co;2

Mariotte, E. (1679). *Essais de Physique: ou Mémoires Pour Servir á la Science des Choses Naturelles*. Paris: Chez Estienne Michallet.

Texas Electronics (2019). *Field_Calibration_Device_Spec. [PDF] Dallas: Texas Electronics, p.*1. Available at: https://texaselectronics.com/media/mconnect_uploadfiles/f/i/field_calibration_device_spec.pdf (accessed Oct 2, 2019).

Vasvári, V. (2005). Calibration of tipping bucket rain gauges in the Graz urban research area. *Atmos. Res.* 77, 18–28. doi: 10.1016/j.atmosres.2004.12.012

**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.

*Copyright © 2019 Lopez Alcala, Udell and Selker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

# Advantages of publishing in Frontiers

**OPEN ACCESS**
Articles are free to read for greatest visibility and readership

**FAST PUBLICATION**
Around 90 days from submission to decision

**HIGH QUALITY PEER-REVIEW**
Rigorous, collaborative, and constructive peer-review

**TRANSPARENT PEER-REVIEW**
Editors and reviewers acknowledged by name on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** info@frontiersin.org | +41 21 510 17 00

**REPRODUCIBILITY OF RESEARCH**
Support open data and methods to enhance research reproducibility

**DIGITAL PUBLISHING**
Articles designed for optimal readership across devices

**FOLLOW US**
@frontiersin

**IMPACT METRICS**
Advanced article metrics track visibility across digital media

**EXTENSIVE PROMOTION**
Marketing and promotion of impactful research

**LOOP RESEARCH NETWORK**
Our network increases your article's readership