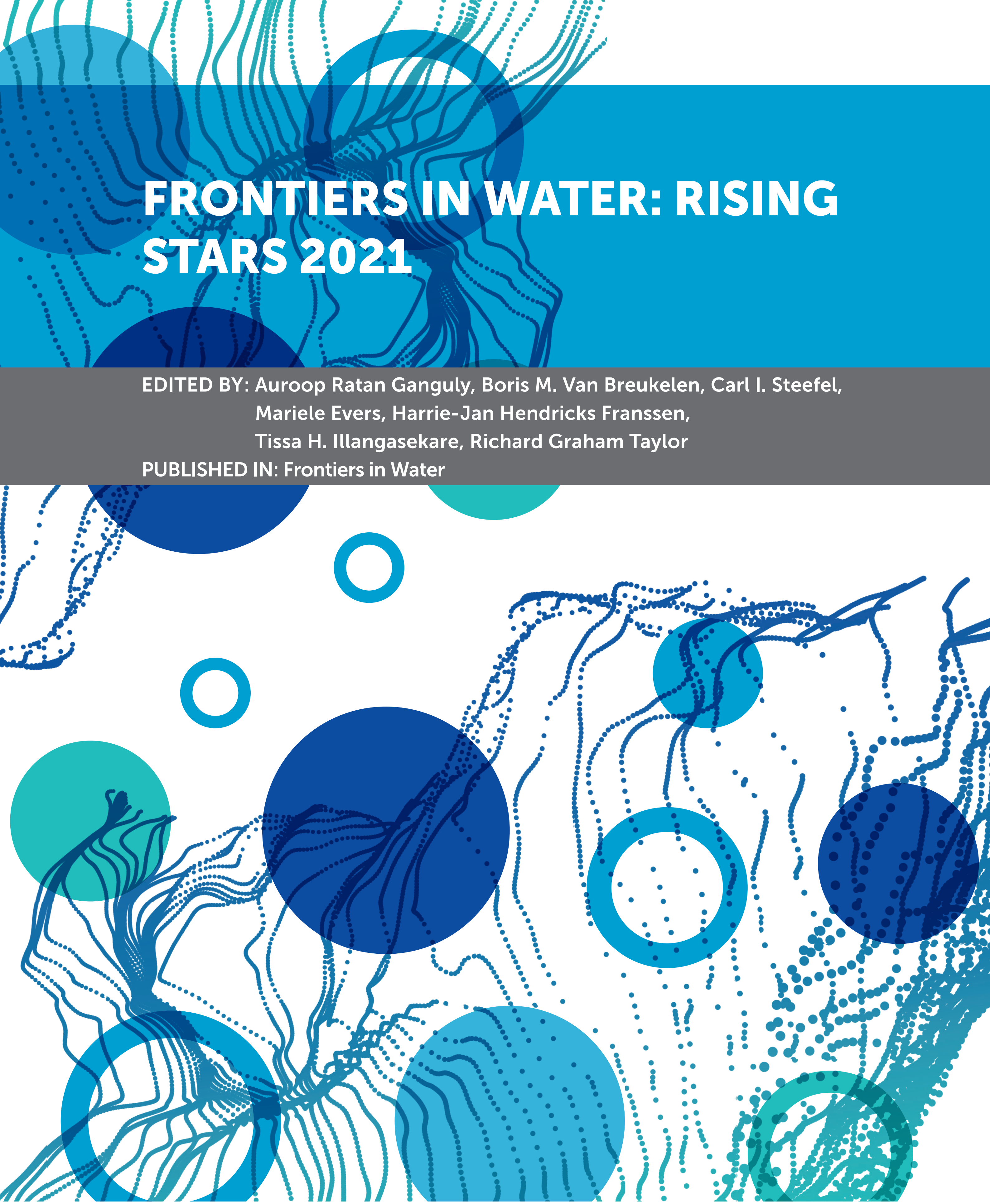


FRONTIERS IN WATER: RISING STARS 2021

EDITED BY: Auroop Ratan Ganguly, Boris M. Van Breukelen, Carl I. Steefel,
Mariele Evers, Harrie-Jan Hendricks Franssen,
Tissa H. Illangasekare, Richard Graham Taylor

PUBLISHED IN: Frontiers in Water

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-83250-544-1
DOI 10.3389/978-2-83250-544-1

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

FRONTIERS IN WATER: RISING STARS 2021

Topic Editors:
Auroop Ratan Ganguly, Northeastern University, United States
Boris M. Van Breukelen, Delft University of Technology, Netherlands
Carl I. Steefel, Berkeley Lab (DOE), United States
Mariele Evers, University of Bonn, Germany
Harrie-Jan Hendricks Franssen, Julich Research Center, Helmholtz Association of German Research Centres (HZ), Germany
Tissa H. Illangasekare, Colorado School of Mines, United States
Richard Graham Taylor, University College London, United Kingdom

Citation: Ganguly, A. R., Van Breukelen, B. M., Steefel, C. I., Evers, M., Franssen, H-J. H., Illangasekare, T. H., Taylor, R. G., eds. (2022). Frontiers in Water: Rising Stars 2021. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-544-1

Table of Contents

frontiers | Frontiers in Water

TYPE Editorial
PUBLISHED 04 October 2022
DOI 10.3389/frwa.2022.1033848

EDITED AND REVIEWED BY
Uwe Schröder,
University of Greifswald, Germany

*CORRESPONDENCE
Auroop R. Ganguly
a.ganguly@northeastern.edu

SPECIALTY SECTION
This article was submitted to
Water and Built Environment,
a section of the journal
Frontiers in Water

RECEIVED 01 September 2022
ACCEPTED 15 September 2022
PUBLISHED 04 October 2022

CITATION
Ganguly AR, van Breukelen BM,
Evers M, Hendricks Franssen H-J,
Illangasekare T, Kumar P, Steefel C and
Taylor RG (2022) Editorial: Frontiers in
water: Rising stars 2021.
Front. Water 4:1033848.
doi: 10.3389/frwa.2022.1033848

Editorial: Frontiers in water: Rising stars 2021

Auroop R. Ganguly[1,2]*, Boris M. van Breukelen[3], Mariele Evers[4], Harrie-Jan Hendricks Franssen[5], Tissa Illangasekare[6], Praveen Kumar[7], Carl Steefel[8] and Richard G. Taylor[9]

[1]Northeastern University, Boston, MA, United States, [2]Pacific Northwest National Laboratory, Richland, WA, United States, [3]Delft University of Technology, Delft, Netherlands, [4]University of Bonn, Bonn, Germany, [5]Agrosphere (IBG-3), Forschungszentrum Julich GmbH, Jülich, Germany, [6]Colorado School of Mines, Golden, CO, United States, [7]University of Illinois, Urbana-Champaign, IL, United States, [8]Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [9]University College London, London, United Kingdom

KEYWORDS

rising stars, water, wastewater, floods and droughts, ecology, energy-water nexus, climate, data science

Editorial on the Research Topic
Frontiers in water: Rising stars 2021

Providing access to clean water is one of the 14 grand challenges listed by The National Academy of Engineering in the United States. A report of the 2021 Davos agenda of the World Economic Forum, while referring to the 17 Sustainable Development Goals (SDGs) set forth by the United Nations, is provocatively entitled: *"If you want to make progress on all the major global challenges, start with water."*

Water resources including quantity and quality, water-related extremes such as floods and droughts, and the nexus of water with food, energy, infrastructures, lifelines, health, ecosystems, socioeconomics, and diplomacy, remain among the most urgent societal challenges. From climate change (Figure 1) and population movement to sociopolitical upheavals and economic progress, water remains central to our shared future. However, our knowledge of processes impacting and impacted by water remains incomplete, which leads to cascading uncertainties in scientific understanding of global challenges and becomes a barrier to generating actionable insights. Notwithstanding this urgency and significance, the description of best practices in water in the research literature remains fragmented and unwieldy, with limited integration across disciplinary fields and limited understanding or use of modern technologies. Yet there is hope. Our aspiration as the editors of Frontiers in Water has been to develop an open-access strategy to develop knowledge across water-related disciplines. To explore where we need to be as a field, we have invited a few rising stars who are boldly and creatively leading the way into the future by contributing articles that may provide a glimpse of the future. This brings us to a discussion of the 10 articles—from authors based in South and North America, Asia, and Europe—which together comprise this *Frontiers in Water Rising Stars 2021* collection.

FIGURE 1
Saint Mary Lake with a view of the Wild Goose Island in Glacier National Park, Montana, USA. Climate variability and change are beginning to have noticeable impacts on the glaciers with implications for regional freshwater and ecosystems. Photo credit: Debashree Bagchi Ganguly. Date: 18th August, 2022 CE.

Bänziger et al. from Switzerland and Nepal have developed risk-informed interventions to improve microbial water quality in rural Nepal in their original research article. Their insights and methods are likely to inform future studies and practices in low-income countries. Navato and Mueller from Boston describe novel adaptations of data-driven approaches for automated anomaly detection in wastewater influents. The methods and insights reported in their brief research report were validated based on changes during the SARS-CoV-2 pandemic. The perspective article by van Emmerik et al. from Wageningen in the Netherlands discusses how river systems act as reservoirs of plastic. Insights on retention mechanisms and the fate of plastics in riverine systems can be directional for our understanding of plastic pollution, which has been described by the United Nations Environmental Programme (UNEP) as a global problem. The original research article by Li et al. from Berkeley develops a framework to understand the pore-scale dynamics of two-phase flow and its implications for mineral reaction rates. Their pore-scale simulations and results suggest how mineral reactions are controlled by interfacial dynamics involving multiphase flow, transport, and surface reactions. The original research article by Nair et al. from India examines the ability of Surface Water and Ocean Topography (SWOT) data products to monitor water volumes in lake levels. A case study in a lake in Kerala, India suggests the possibility of new SWOT-based insights on anthropogenic impacts on and monitoring levels of lakes. A review article by Ghorpade et al. from Mumbai evaluates current literature on Intermittent Water Supply (IWS) in India. They discuss India's water supply system, as well as the drivers and the mechanisms of IWS, including what they call the vicious cycle of IWS in India. Ceola et al. from Italy, Bristol the United Kingdom, and Luxembourg, study flood risks and flood mitigation strategies in their original research article. Based on a hydrodynamic model and satellite-derived data, they investigate the dynamics of the components of flood risks with a case study on the Murray-Darling basin in Australia. Samadi from Clemson in the USA reports how what the article calls the Artificial Intelligence of Things can advance research on floods in an original research article. The vision was demonstrated through a prototype and discussion on the workflow of a flood analytics information system. The original research article by Ricardo et al. from the USA, Costa Rica, and Argentina, characterizes dissolved organic carbon variations in an elevated peatland of the Talamanca range of Costa Rica. Their insights and data on understudied high-elevation peatlands in a tropical coastal lowland point to new directions for further studies and provide a baseline, with implications for associated ecosystems. Iloejesi and Beckingham from Auburn in the USA examine the geochemical implications on subsurface porous aquifers during carbon dioxide injection, storage, and extraction. Given that such aquifers act as reservoirs for compressed energy storage of renewable energy, their original research article provides insights into an important aspect of the water-energy nexus.

The 10 articles cited above touch upon water quality and availability, flood hazards, as well as the nexus of water with ecosystems and energy. The rising stars and their co-authors are drawn from South and North America, Europe, and Asia, reflecting water-related opportunities and challenges in the global south and the global north. The depth and breadth of the articles are expected to provide new directions to water researchers and practitioners around the world.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

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

Publisher's note

frontiers
in Water

ORIGINAL RESEARCH
published: 26 August 2021
doi: 10.3389/frwa.2021.689404

Influence of Storage Period on the Geochemical Evolution of a Compressed Energy Storage System

*Chidera O. Iloejesi and Lauren E. Beckingham**

Department of Civil and Environmental Engineering, Auburn University, Auburn, AL, United States

Subsurface porous aquifers are being considered for use as reservoirs for compressed energy storage of renewable energy. In these systems, a gas is injected during times in which production exceeds demand and extracted for energy generation during periods of peak demand or scarcity in production. Current operational subsurface energy facilities use salt caverns for storage and air as the working gas. CO_2 is potentially a more favorable choice of working gas where under storage conditions CO_2 has high compressibility which can improve operational efficiency. However, the interaction of CO_2 and brine at the boundary of the storage zone can produce a chemically active fluid which can result in mineral dissolution and precipitation reactions and alter the properties of the storage zone. This study seeks to understand the geochemical implications of utilization of CO_2 as a working gas during injection, storage and extraction flow cycles. Here, reactive transport simulations are developed based on 7 h of injection, 11 h of withdrawal and 6 h of reservoir closure, corresponding to the schedule of the Pittsfield field test, for 15 years of operational life span to assess the geochemical evolution of the reservoir. The evolution in the storage system is compared to a continuously cyclic system of 12 h injection and extraction. The result of the study on operational schedule show that mineral reactivity occurs at the inlet of the domain. Furthermore, the porosity of the inner domain is preserved during the cycling of CO_2 acidified brine for both systems.

Keywords: energy storage, CO_2 sequestration, porous saline aquifer, reactive transport simulation, geochemical reactions

OPEN ACCESS

Edited by:
Carl I. Steefel,
Lawrence Berkeley National Laboratory, United States

Reviewed by:
Sumit Purohit,
Pacific Northwest National Laboratory (DOE), United States
Qingyun Li,
Stanford University, United States

***Correspondence:**
Lauren E. Beckingham
leb0071@auburn.edu

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *31 March 2021*
Accepted: *30 July 2021*
Published: *26 August 2021*

Citation:
Iloejesi CO and Beckingham LE (2021) Influence of Storage Period on the Geochemical Evolution of a Compressed Energy Storage System. Front. Water 3:689404. doi: 10.3389/frwa.2021.689404

HIGHLIGHTS

- The viability of use of a reactive cushion gas in compressed energy storage system needs to be considered
- Reactivity at cushion gas boundary for operational schedules with and without storage are compared
- Geochemical reactions increase porosity close to the injection well but are otherwise limited
- Operational schedule does not largely impact geochemical reactions in a low carbonate mineral assemblage reservoir.

INTRODUCTION

The global renewable energy share in terms of final energy consumption has increased over the past decade to 17.3% in 2017 (IEA, 2019). Increased utilization of renewables resulted from government policy incentives (Huang et al., 2007; Chandler, 2009; Lyon and Yin, 2010; Jenner et al., 2012), technical advancements (Álvarez-Herránz et al., 2017; Mensah et al., 2018; Lin and Zhu, 2019), and environmental benefits that promote social acceptance (Haar and Theyel, 2006; Aslani et al., 2012; Wang et al., 2018). Moreover, adopting renewable energy has been positively linked to economic development (Sadorsky, 2009a,b; Apergis and Payne, 2010) and reductions in anthropogenic greenhouse gas emissions while diversifying energy production (Lund, 2007; Chien and Hu, 2008; Marques et al., 2011; Chen et al., 2014). However, the intermittency of renewable energy production requires a reliable means of long term, large capacity energy storage to achieve energy security through renewable energy (van der Linden, 2006). Efforts to increase and improve energy storage have included fast discharging, low capacity options like lead-acid batteries to slow discharging, high capacity options like pumped hydro and compressed energy storage (Dunn et al., 2011). The high-capacity options store bulk energy in megawatts for hours to months, which offers increased reliability in grid-scale applications of renewable energy.

Compressed energy storage is a promising means of long-term, grid-scale energy storage that has the potential to be widely deployed across the globe in subsurface reservoir formations including salt caverns or porous saline aquifers (Aghahosseini and Breyer, 2018). In these systems, a working gas is injected into the storage formation during periods of excess energy production and extracted to power a turbine during periods of excess energy demand (Succar and Williams, 2008). Salt caverns, mined hard rock caverns and porous saline aquifers can be utilized as storage reservoirs (Allen, 1985). Energy storage has only been carried out in salt caverns to date (van der Linden, 2006). Salt caverns, however, are geographically limited. Porous saline aquifers, on the other hand, are ubiquitous which makes them a potentially more favorable option (Eckroad and Gyuk, 2003; Succar and Williams, 2008). The idea of extending the compressed energy storage medium beyond the currently used salt caverns to porous aquifers will facilitate more widespread possible utilization of this technology for municipal energy storage.

To initialize the system in a porous aquifer, a cushion gas may first be permanently injected into the porous formation to raise background pressure to help increase working gas recovery (Carden and Paterson, 1979). The cushion gas is injected into the formation to buffer the pressure fluctuation during the cycling of a working gas during energy storage. Hence, the compressibility of the cushion gas is thus an important consideration in helping to improve the economics and efficiency of compressed energy storage. The injection of these storage gases stratifies the porous saline aquifer into zones with varying compositions of gas to brine ratio (Cui et al., 2018). Although, these zones may not be distinctly classified and there could be mixing of working gas and cushion gas at the interface of the two gases (Kim et al., 2015). Furthest from the injection well is the zone where the cushion gas and the brine interface. At this zone, gas dissolution into the brine is controlled by mutual solubilities (Beckingham and Winningham, 2019).

For over eight decades that subsurface storage has been prevalent (Carden and Paterson, 1979; Allen, 1985), a variety of gases have been considered as cushion and working gases including nitrogen (Bauer et al., 2015; Pfeiffer and Bauer, 2015), native methane (Oldenburg, 2003), and air (Succar and Williams, 2008). It was in the wake of geologic CO_2 sequestration that studies began to investigate the potential of using CO_2 as cushion gas. At the typical depth of saline aquifers, CO_2 becomes a supercritical fluid with a high density and compressibility. This makes it a particularly favorable choice of cushion gas where its high compressibility may improve available storage capacity, recovery and efficiency (Oldenburg, 2003). Further consideration of the compressibility of CO_2 shows that using CO_2 as cushion gas can help reduce significant pressure variation during working gas cycling (Oldenburg and Pan, 2013). An additional positive benefit of using CO_2, a major greenhouse gas, is that it can remain permanently sequestered in the storage formation once injected. This is because the compressed energy system is designed to recycle only the working gas while the cushion gas remains permanently in the formation. Hence, utilizing CO_2 for this purpose would provide potential environmental benefits and economic benefits in the form of 45Q tax credits (Mai et al., 2016). However, careful system design needs to be considered to not produce CO_2 during the extraction of the working gas (Oldenburg and Pan, 2013). This can lead to pressure losses in the formation that affect working gas recovery (Ma et al., 2019). Also, failure to sequester CO_2 while being used as cushion gas could offset potential tax credits gained by using CO_2 as a cushion gas (Rul Internal Revenue Service, 2009). It should also be noted that CO_2 can also be used as the working gas but this would require using a closed-loop system to exploit the beneficial physical properties of CO_2 for energy production while ensuring that none escapes to the atmosphere (Alami et al., 2019).

Injecting CO_2 in porous saline aquifers, however, introduces additional technical complexities that need to be considered before adoption. This includes a need to understand possible geochemical limitations that can result from interactions between the injected gas, formation brine, and formation minerals. One of the key geochemical considerations is the interaction of the aquifer formation with the cushion gas which occupies one-third of the total storage volume (Walters, 1976). While CO_2 is a favorable choice of cushion gas from its physical property point of view, there could be potential geochemical implications of reactions between CO_2, formation brine, and formation minerals that could pose to be a challenge to system operation (Zhang and Huisingh, 2017). These reactions, however, have not largely been considered. Injecting reactive fluids like CO_2 in the subsurface can acidify the formation brine and result in complex mineral dissolution and precipitation reactions within the formation rock matrix, as has been observed in the context of geologic CO_2 sequestration (Gunter et al., 1993; Fischer et al., 2010; Kharaka and Cole, 2011; DePaolo and Cole, 2013).

Flow conditions in geologic CO_2 sequestration systems are distinctly different than those in energy storage systems. In geologic CO_2 sequestration systems, flow is predominately unidirectional away from the injection well while energy storage systems utilize cyclic, bi-directional flow conditions as a result of the injection and extraction of the working gas and may contain intermittent storage periods. Previous reactive transport simulations have considered the impact of cyclic flow conditions on geochemical reactions in energy storage systems in a porous saline aquifer. These simulations have shown that the rate and extent of potential dissolution and precipitation reactions at three locations in the formation are significantly reduced in the energy storage system in comparison to what is anticipated in geologic CO_2 sequestration system (Iloejesi and Beckingham, 2021). While this result is promising, the study only considered a 4 month operational period with constant injection and extraction cycling. These flow conditions can be referred to as a continuous operational schedule. To fully assess reactivity, longer time durations and variations in operational regime need to be considered. Operational schedules in most subsurface storage aquifers integrate periods of storage or shut-in. This is when the injected gas is allowed to sit in the aquifer with little to no flow before it is extracted to meet energy demand. This operational schedule can be referred to as a periodic schedule. The resulting geochemical reactions at the cushion gas boundary and implications for these types of operating conditions have not been considered.

This study aims to enhance understanding of the rate and extent of potential geochemical reactions in a porous saline aquifer utilized for energy storage which operates using a periodic schedule. This is done by comparing the difference between the geochemical evolution of a porous aquifer compressed energy storage system operating with a continuous schedule with the geochemical evolution when the system operates using the periodic operational schedule. Here, reactive transport simulations are developed considering daily cyclic interactions between the cushion gas (CO_2), formation brine, and formation minerals over a 15 years study period. The evolution of major ions in the formation brine, formation minerals, and porosity for each system is tracked and compared to aid in the understanding of the use of CO_2 as a cushion gas for compressed energy storage systems in porous saline aquifers operating with periodic schedule.

METHOD

The impact of using CO_2 as a cushion gas for energy storage in porous saline aquifer formations is considered here through reactive transport simulations. Simulations consider a case study of energy storage in the Paluxy formation for two operational regimes, one with a storage period and one without.

Sample

The Paluxy formation is considered here as a potential storage reservoir. This formation has been considered for CO_2 enhanced oil recovery (Robinson and Davis, 2012) and geologic CO_2 storage projects (Petrusak et al., 2010; Qin and Beckingham, 2019; Bensinger and Beckingham, 2020). Previous investigation of this formation has found it as high porosity, high permeability sandstone (Pashin et al., 2018; Bensinger and Beckingham, 2020). This formation is predominantly quartz (76.45%) with 9.64% calcite, 8.23% smectite, and the remainder minor phases (<5%), as determined from SEM imaging analysis in Qin and Beckingham (2019), where the mineral abundances are given in **Table 1**. The porosity of the sample obtained from image segmentation is 24.84% and the calculated permeability, as estimated using pore network modeling in Bensinger and Beckingham (2020), is 1414.3 mD.

Reactive Transport Simulations

Reactive transport simulations to understand the geochemical evolution of the reservoir in the brine-saturated region surrounding the gas plume were developed here for energy storage systems using CO_2 as a cushion gas. Here, the initial simulations developed in Iloejesi and Beckingham (2021) are extended to consider two storage operational schedules for 15 years. Simulations were developed in CrunchFlow, a multi-species reactive transport simulation code (Steefel and Molins, 2009). CrunchFlow has been used extensively used to understand subsurface geochemical reactions (Zhang et al., 2015; Dávila et al., 2016).

To consider the most reactive zone in the storage system, the acidified-brine zone adjacent to the two-phase zone, where brine and supercritical CO_2 exist, was considered here (**Figure 1**). In this region, CO_2 will dissolve into the brine phase, creating a region of acidified-brine favorable for geochemical reactions in the system (Huq et al., 2015). Here, a simplified system that ignores the advective mixing of the brine with CO_2 is considered to maximize brine acidification and resulting reactivity. Here, a 15-cm domain in the acidified-brine zone was selected for simulations. The model domain was discretized into 45 equally spaced grid cells. The model domain was bounded by upstream and downstream "ghost" cells that serve as the boundary condition in the one-dimensional flow through simulation. The upstream "ghost cell" is the closest grid cell to the injection well. For ease of comparison of the concentration of major ion species, mineral volume fractions, and porosity evolutions, the model system evolution is monitored across the domain at select time intervals.

Two daily operational schedules were used in this study, one consisting of a periodic operational schedule that includes a storage period and the other a continuous operational schedule without the storage period. We refer to the flow condition for the continuous schedule as the injection–extraction flow regime and the flow condition for the periodic schedule as the injection–storage–extraction flow regime. One cycle of the continuous schedule consisted of a 12 h injection period followed by a 12 h extraction period. One cycle of the periodic schedule incorporated a 6 h storage period between 7 h of injection and 11 h of extraction. The periodic schedule is similar to the schedule used during the Pittsburg, Illinois field test (Allen, 1981; Allen et al., 1983). In both simulations, the duration of injection corresponded to flow of CO_2 saturated brine away from the injection well supplied by a "ghost" boundary cell. The effluent brine was captured in the downstream boundary "ghost" cell and recycled during the extraction period. The upstream boundary or "ghost" cell served as a catchment for the effluent during the extraction period. The captured brine was then recycled through

TABLE 1 | Mineral abundance as percent volume, volume fraction, and surface area of the Paluxy formation obtained from multi-scale imaging of the sample in Qin and Beckingham (2019) and rate constants for the respective mineral phases at reservoir condition as obtained from the literature for quartz (Knauss and Wolery, 1988; Brady and Walther, 1990), calcite (Alkattan et al., 1998), K-Feldspar (Bevan and Savage, 1989), Smectite (Amram and Ganor, 2005), Muscovite (Oelkers et al., 2008), and Siderite (Golubev et al., 2009).

Mineral	Mineral reaction	Mineral abundance (v%)	Volume fraction	Surface area (m^2g^{-1})	Log rate constant (mol s^{-1} m^{-2})
Primary mineral					
Quartz	Quartz = $SiO_{2(aq)}$	76.45	0.5740	2.59E−2	−11.60
Calcite	$CaCO_3 + H^+ = Ca^{2+} + HCO_3^-$	9.63	0.0724	1.42E−3	−4.21
K-Feldspar	K-Feldspar + $4H^+ = K^+ + Al^{3+} + SiO_{2(aq)} + 2H_2O$	3.50	0.0263	1.15E−3	−11.65
Smectite	Smectite + $7.8H^+ = 0.2K^+ + 1.25Al^{3+} + 3.5SiO_{2(aq)} + 4.9H_2O + 0.7Fe^{2+} + 0.1Na^+ + 0.025Ca^{2+} + 1.15Mg^{2+} + 0.05O_{2(aq)}$	8.23	0.0619	1.63E+1	−13.35
Muscovite	Muscovite + $10H^+ = 3SiO_{2(aq)} + 6H_2O + 3Al^{3+} + K^+$	0.31	0.0023	1.10E+0	−12.67
Siderite	Siderite + $H^+ = HCO_3^- + Fe^{2+}$	1.98	0.0141	6.49E−4	−5.69
Secondary mineral					
Kaolinite	Kaolinite + $6H^+ = 5H_2O + 2Al^{3+} + 2SiO_{2(aq)}$				−12.43
Gibbsite	Gibbsite + $3H^+ = 3H_2O + Al^{3+}$				−10.00
Albite	Albite + $4H^+ = 2H_2O + Al^{3+} + Na^+ + 3SiO_{2(aq)}$				−11.11
Ankerite	Ankerite + $2H^+ = Ca^{2+} + Mg^{2+} + 2HCO_3^-$				−7.70
Chlorite	Chlorite + $10H^+ = 2Fe^{2+} + SiO_{2(aq)} + 2Al^{3+} + 7H_2O$				−11.15
Chalcedony	Chalcedony = SiO2(aq)				−11.58

**CrunchFlow has a default surface area for secondary minerals.*

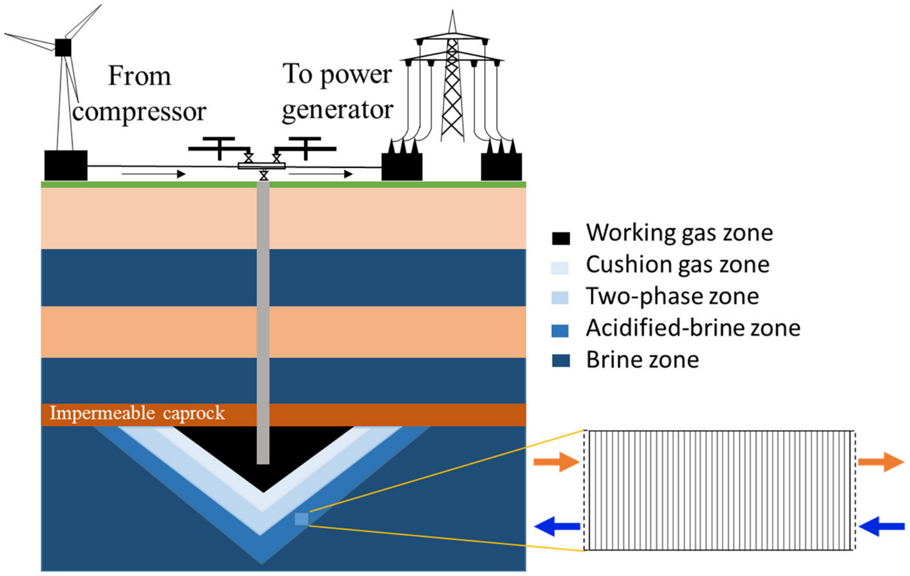


FIGURE 1 | Schematic of the energy storage system and model domain for the reactive transport simulations. The "two-phase zone" corresponds to the presence of brine and cushion gas.

TABLE 2 | Simulation Parameters for the periodic and continuous simulation model.

Simulation parameters		Periodic schedule	Continuous schedule
Flow rate (m/day)	Injection	0.838	0.489
	Shut-in	0	–
	Extraction	0.533	0.489
Operation schedule (h)	Injection duration	7	12
	Shut-in duration	6	–
	Extraction duration	11	12
Reservoir permeability		1555.5 mD	
Reservoir porosity		24.84 %	

the model domain during the next injection period. As such, the composition of the influent brine for each component of the flow cycle was based on the effluent of the preceding period of the flow cycle. The influent brine for each injection period, however, was first equilibrated with CO_2 before flowing through the model domain. For the flow regime that incorporated storage, flow was ceased during storage period and the system effectively resembled a batch system. In both operational regimes, a complete cycle took 24-h which corresponds to a compressed energy storage system used daily for power generation (Allen et al., 1983; Pfeiffer et al., 2017; Fleming et al., 2018). Simulations for each operational regime were carried out for a 15 year study period.

The model system was assumed as a homogenous and isotropic domain and initialized based on the mineral compositions, surface areas, and porosity from Qin and Beckingham (2019). The initial brine composition was calculated based on a 1 M NaCl brine in equilibrium with formation minerals for 10,000 years at reservoir temperature and pressure (Xu et al., 2007; Qin and Beckingham, 2021). The corresponding reservoir temperature and pressures were 50°C and 100 bar based on the geothermal gradient at Kemper, Mississippi and a typical pressure gradient (Nathenson and Guffanti, 1988; Bachu, 2000; Reysa, 2005). The pH of the brine was determined via charge balance. The CO_2 saturated brine composition was determined by equilibrating brine with a constant partial pressure of CO_2 at the depth of storage in the Paluxy formation. The Duan CO_2 solubility model, which factors for high pressure and temperature conditions, was used to calculate the CO_2 solubility in brine (Duan et al., 2006). A brine flowrate of 0.489 m/day was used for the continuous schedule. The flowrate is extrapolated for the Paluxy reservoir conditions from field-scale simulations of brine fluid velocities adjacent to injected CO_2 plumes in a similar reservoir condition (Zhang and DePaolo, 2017). Two different flow rates were used for the periodic operational schedule to maintain pore volumes equal to the continuous operational system during each 24-h cycle. Flowrates of 0.838 m/day was used for the 7 h injection period and 0.533 m/day used for the 11 h extraction period (**Table 2**).

Mineral reactions in CrunchFlow are simulated using a rate law based on transition state theory. The corresponding parallel rate laws are given by,

$$r_s = -Ak\left(1 - \left(\frac{Q_s}{K_s}\right)^M\right)^n \tag{1}$$

where r_s is the reaction rate, A is the reactive surface area of a constituting mineral in the rock sample, k is the equilibrium dissolution rate constant, n and M are exponents which are experimentally determined to explain nonlinear dependence of the dependence on oversaturation, K_s is the equilibrium constant, and Q_s is the ion activity product for the rock–water interaction (Steefel et al., 2015). Here, reaction rate constants from previously published experimental works were extrapolated to formation conditions, following the method in Beckingham et al. (2017). Mineral accessible surface areas determined from a multi-scale imaging analysis in Qin and Beckingham (2019) on a sample from the Paluxy formation were used here as reactive surface areas. These surface areas reflect the mineral surfaces accessible to reactive fluids in the formation. The aqueous activity coefficients of the brine solution were calculated using the extended Debye-Huckle model as given by,

$$\log \gamma_{\pm} = -\frac{|Z_+Z_-|\, A_\gamma I^{\frac{1}{2}}}{I + \mathring{a} B_\gamma I^{\frac{1}{2}}} + b_\gamma I \tag{2}$$

where $\gamma_{\pm}$ is activity coefficient of a completely dissociated binary electrolyte consisting of ions with charges Z_+Z_-, A_γ and B_γ are Helgeson's extended Debye-Huckle constant, I is the trues ionic strength and a, b_γ are empirical parameter characteristic of the electrolyte (Helgeson and Kirkham, 1974). The extended Debye-Huckel equation applies to dilute solution where concentration is $< \sim 1$ molal (Xu et al., 2017).

RESULT AND DISCUSSION

The results of the reactive transport simulations for the continuous and periodic operational schedules are presented here. For each simulation, the evolution of the mineral volume fractions and saturation index of mineral phases, including the potential for secondary mineral precipitation, is tracked. The resulting evolution of porosity is then presented.

Continuous Operation Schedule

The simulated evolution of mineral volume fractions across the domain length for the continuous operational schedule (cyclic injection and extraction) are shown in **Figure 2**. During the first cycle, the flow of acidified brine into the domain during injection results in rapid dissolution of calcite close to the injection point. This increases the calcium ion concentration (**Supplementary Figure 1**) and buffers the pH (**Supplementary Figure 1**) such that no additional calcite dissolution occurs further away from the source of injection (**Supplementary Figure 2**). Siderite also dissolves resulting in an increase in iron concentrations and further buffering of the pH. Dissolution of siderite occurs throughout the domain with greater extents of dissolution closer to the injection well where the brine is the most undersaturated with respect to siderite. Smectite dissolves equally across the simulation domain during the first injection cycle. Smectite dissolution results in a brief period of early supersaturation and precipitation of muscovite, shown by the increase in muscovite volume fraction in **Supplementary Figure 2**. This is followed by slight dissolution of muscovite throughout the domain at later times. Quartz and K-feldspar remain stable across the domain throughout the first injection period.

After 12 h, the flow reverses as the first extraction cycle begins and continues until 24 h. During this period, a constant composition brine, corresponding to the brine composition at the end of the proceeding injection cycle, enters the domain at the location furthest from the injection well and flows toward the injection well. The recycled brine has a pH of 4.93 (**Supplementary Figure 1**) and results in dissolution of siderite, producing iron ions (**Supplementary Figure 1**). The dissolution rate of siderite is the highest furthest from the injection well, where the distance from equilibrium is greatest and reduces closer to the injection well as iron concentrations in solution increase. Smectite dissolves across the simulation domain at a relatively constant dissolution rate. No additional calcite dissolution occurs during this period due to elevated calcium ion concentrations in the solution from dissolution during the proceeding injection cycle. Muscovite, quartz and K-feldspar remain stable.

The second injection period begins after 24 h. During this period, a constant composition brine saturated with CO_2 enters the simulation domain. As the brine is recycled, the ion concentrations from the proceeding extraction period are maintained with additional acid added from equilibrium with the adjacent CO_2 plume. This results in conditions favorable for additional mineral dissolution, most notably siderite and smectite which are still undersaturated in solution. No additional calcite dissolution occurs because of the high concentration of calcium in the brine.

This continuous cycling of injection and extraction continues over the 15-year study period and the resulting simulated evolution of mineral volume fractions are shown in **Figure 2**. Muscovite continuously dissolves and is completely consumed within the first 1.5 years. The dissolution of muscovite results in an increase in the concentrations of potassium, aluminum, and silicate ions. Released ions, from muscovite and other dissolving phases, create conditions favorable for formation of secondary mineral phases where saturation indices for potential secondary mineral phases are given in **Figure 3**. Chlorite precipitation is promoted by muscovite and siderite dissolution where siderite dissolves throughout the domain during the first 1.5 years. Once muscovite is consumed, siderite starts precipitating (**Supplementary Figure 3**) as conditions no longer favor chlorite precipitation as shown in equation 3 through equation 5. An increase in smectite dissolution also occurs, increasing magnesium ion concentrations. The elevated concentrations of calcium and magnesium ions in the brine result in ankerite precipitation. As calcium is consumed with ankerite precipitation, calcite dissolution is again favored starting at 1.5 years and calcite is completely depleted near the injection well after 5.5 years. The elevated silica concentrations from dissolution of muscovite and smectite also result in conditions favorable for precipitation of chalcedony. Conditions are also favorable for precipitation of gibbsite, kaolinite, and albite (**Figure 3**). K-feldspar remains stable throughout the simulation.

Operational Schedule Comparison

The simulated evolution of mineral volume fractions for the operational schedule that includes the storage period is shown in **Figure 2**. In this simulation, 7 h of injection is followed by 6 h of storage and 11 h of extraction. During injection, acidified brine flows through the domain away from the injection well at a flowrate of 0.838 m/day for a period of 7 h corresponding to a period of excess energy production. During storage, flow slows to 0 m/day for a period of 7 h, resembling a batch reactor. During extraction, flow recycles back toward the injection well at a rate of 0.533 m/day reflecting the extraction of stored energy from the formation. Overall, injection, storage, and extraction result in the dissolution of calcite at the simulation domain boundaries and siderite, smectite, and muscovite dissolution throughout the simulation domain. Quartz precipitation occurs throughout the simulation domain.

Variations in the evolution of minerals due to the two operational schedules are most easily compared by considering calcite and siderite. Noted differences in these minerals occur closest and furthest from the injection well, especially within the first two cycles (**Supplementary Figure 2**). Both simulations have equal pore volumes, or amount of fluid flow, during injection and extraction periods. Therefore, flow rates are higher in the injection–storage–extraction operational schedule to compensate for the shorter flow duration. In both systems, flow of acidified brine into the system during the first injection period results in rapid dissolution of calcite and siderite. Injection continues for 12 h in the continuous schedule and 7 h in the periodic schedule. The prolonged period of injection in the continuous schedule results in additional calcite and siderite dissolution near the injection well even though the flowrate is higher for the injection–storage–extraction schedule.

After the 7 h of injection, a 6-h storage period with no flow begins in the periodic schedule. During this storage time, the elimination of the transport component of the reaction slows the reactivity of calcite and siderite. The reduction in reactivity is more noted in calcite than siderite as calcite is almost stable during the following storage period. On the other hand, smectite and muscovite dissolution behavior is unaffected by the storage period. Quartz and K-Feldspar maintain their early stability through the first storage period.

The first extraction period begins after 13 h in the periodic schedule and goes on for 11 h. During this time, the recycled brine returns through the system, flowing toward the injection well. The elevated calcium concentrations in solution prevent calcite dissolution. The dissolution rate of siderite, however, increases and is highest at the location furthest from the injection well, closest to the source of brine injection during the extraction period. The extent of dissolution during extraction is higher in the injection–storage–extraction system in comparison to the continuous operational schedule system. This is because more

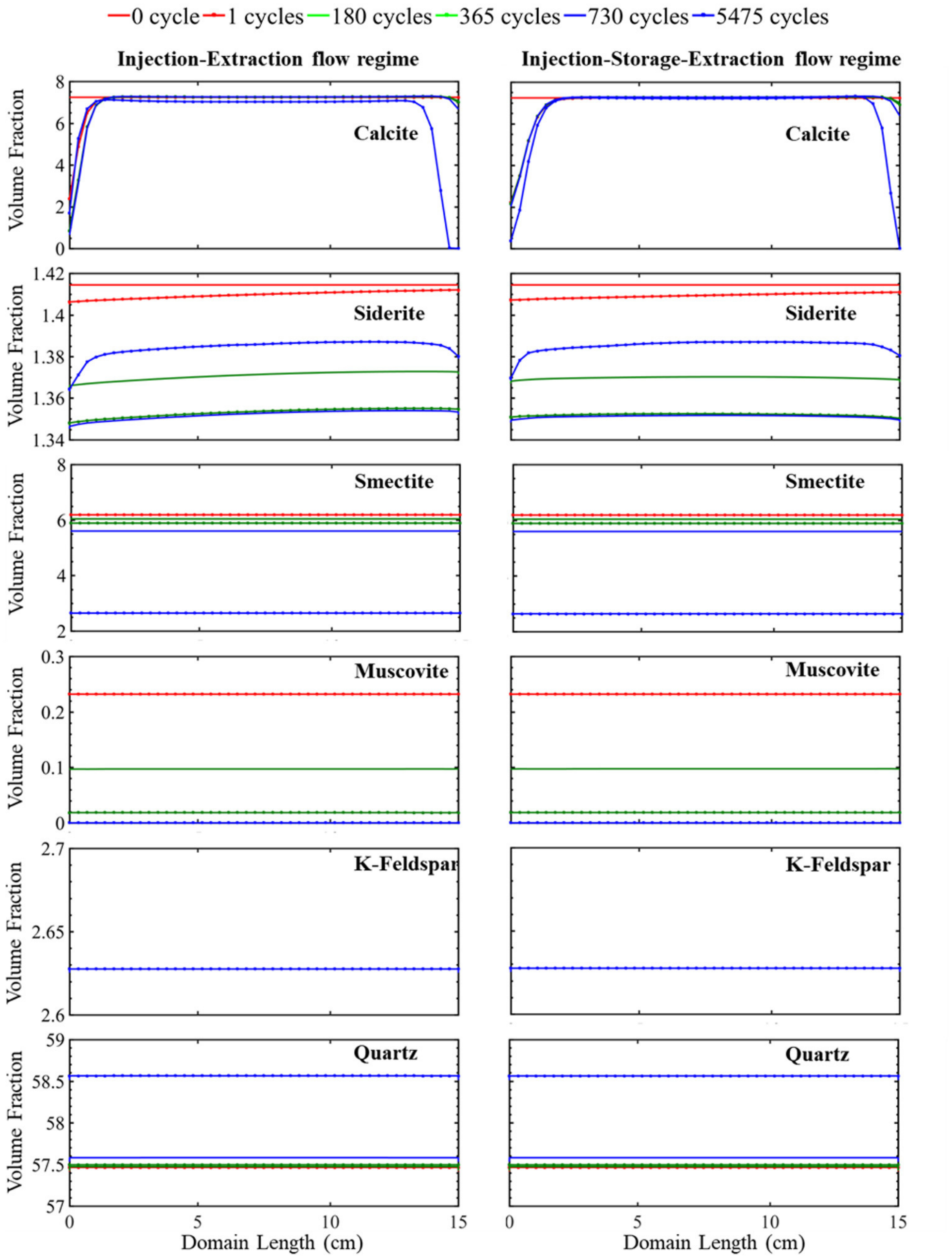


FIGURE 2 | The simulated evolution of mineral volume fractions across the simulation domain at certain times over 15 years for the injection–extraction flow regimes (left) and injection–storage–extraction flow regimes (right). 0 h is the initial condition, and 15 years is the last condition of the porous media. Red reflects 0 h, dotted red 1 day, green 0.5 years, dotted green 1 year, blue 2 years, and dotted blue 15 years.

siderite dissolves in the continuous schedule during the injection period, resulting in a higher saturation of the brine with respect to siderite that limits the extent of siderite dissolution during extraction. In comparison, lower iron concentrations in the recycling brine following storage drive additional dissolution during extraction for the periodic schedule.

Overall, there is not a significant variation in the evolution of minerals between the two operational schedules. In both systems, initial calcite and siderite dissolution buffer the pH. Muscovite dissolves throughout the domain and is completely consumed within the first 1.5 years of simulation in both operational schedules. This complete dissolution of muscovite impacts the reaction pathway of siderite and calcite where the calcite dissolution rate increases as siderite begins to precipitate. Smectite dissolves at a similar dissolution rate in both the continuous and periodic operational

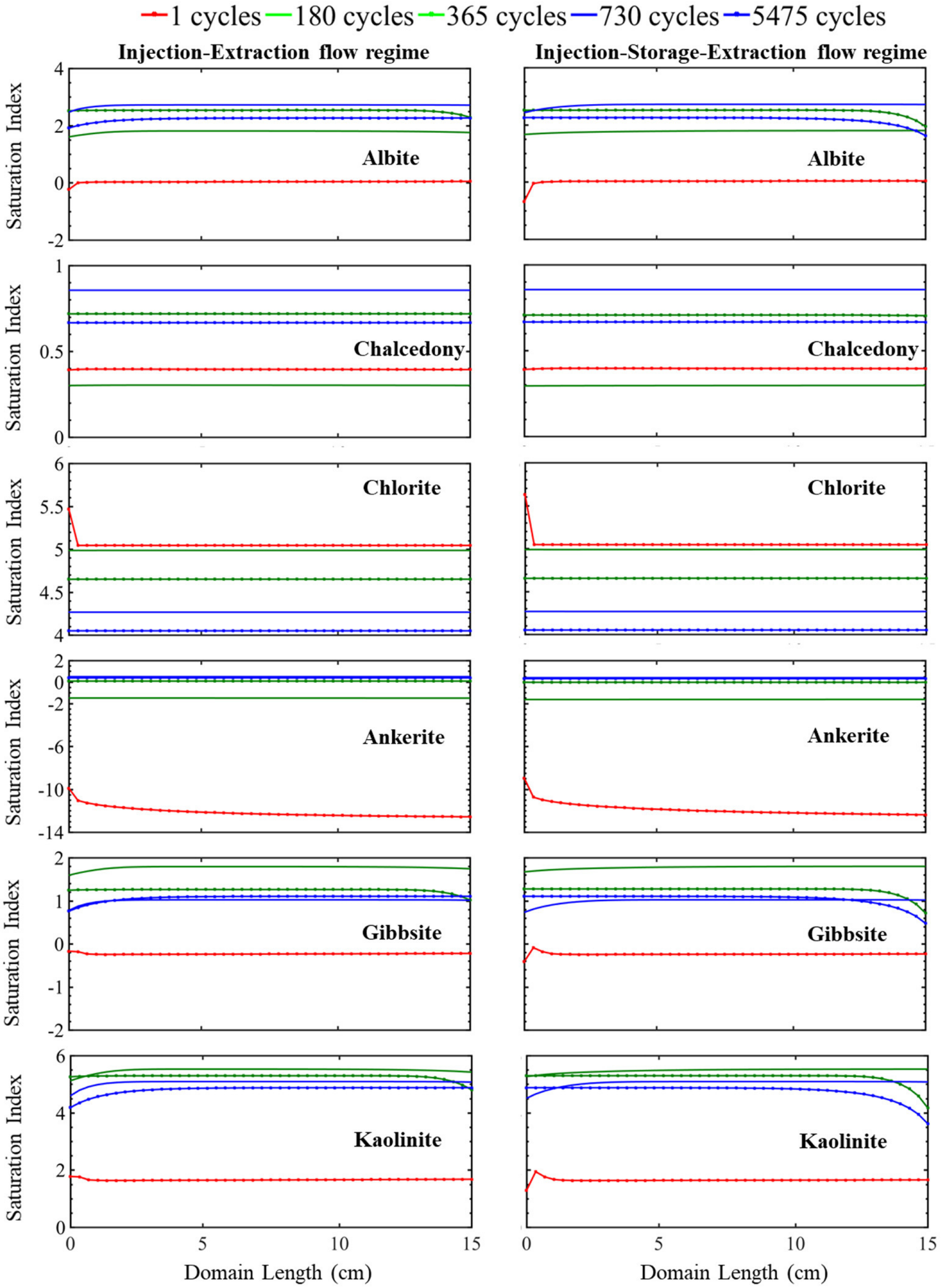


FIGURE 3 | The simulated evolution of mineral saturation index across the simulation domain at certain times over 15 years for the injection-extraction flow regimes (left) and injection–storage–extraction flow regimes (right). A positive saturation index indicates precipitation potential and negative precipitation index indicates dissolution potential. The dotted red line is 1 day, green 0.5 years, dotted green 1 year, blue 2 years, and dotted blue 15 years.

schedule. Released ions from dissolving minerals create conditions favoring precipitation of similar secondary mineral phases (**Figure 3**). After 1.5 years, quartz, which was initially stable within the domain, starts to precipitate. K-feldspar remains stable in both systems throughout the simulation timeline.

Porosity

The simulated evolution of porosity for the continuous and periodic schedules is shown in **Figure 4**. As evident in this figure, energy storage results in a non-uniform increase in porosity throughout the simulation domain. The porosity evolution depends on the rate and extent of dissolving minerals.

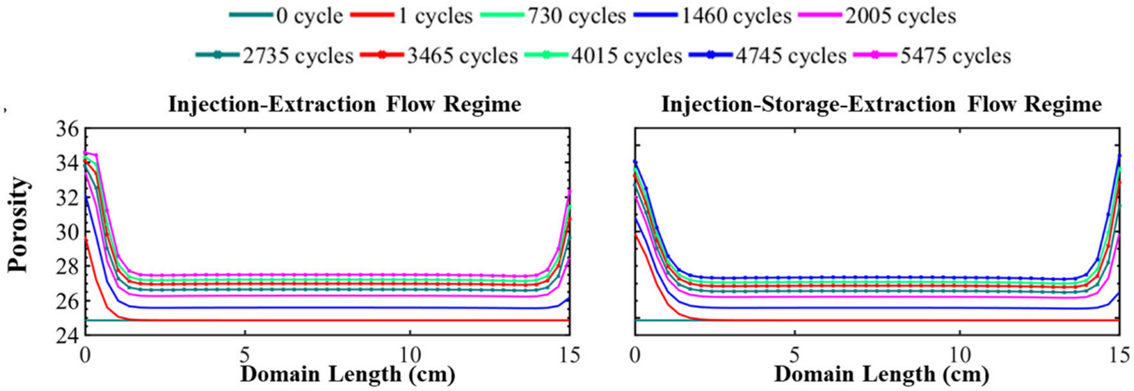


FIGURE 4 | The simulated evolution of porosity across the simulation domain over 15 years for the injection–extraction flow regime (left) and injection–storage–extraction flow regime (right). 0 years (dark green) is the initial condition and 15 years (dotted magenta) is the final condition. Red is 1 day, light green 2 years, blue 4 years, magenta 5.5 years, dotted dark green 7.5 years, dotted red 9.5 years, dotted light green 11 years, dotted blue 13 years, and dotted magenta 15 years.

Here, calcite, siderite, smectite, and muscovite dissolve. While muscovite completely dissolves, its small volume fraction (0.23%) does not largely contribute to porosity evolution. Instead, porosity is predominantly controlled by calcite, smectite, and quartz. Smectite dissolves relatively uniformly throughout the simulation domain, increasing porosity relatively uniformly by 3.54%. Conversely, quartz precipitates uniformly across the domain and reduces the porosity by 4.58%. It should be noted, however, that the extent of precipitation may be overestimated with the TST approach that does not account for nucleation and growth (Bourg et al., 2015). Calcite dissolves non-uniformly throughout the domain, increasing porosity by 5.55–7.24% and the evolution of porosity reflects the variation in calcite volume fraction over time. From the simulation results, it can be seen that elevated ion concentrations in the recycling brine create conditions favorable for precipitation of other secondary mineral phases. However, the impact of secondary minerals on porosity evolution is insignificant where the largest volume fraction of a secondary mineral is three order of magnitude lower than the primary minerals in the system.

The main differences in the porosity evolution between the two operational regimes occurs near the boundaries of the simulation. After the first injection of acidified brine, the difference in porosity between the systems is 1.33 percentage points with a porosity of 29.86% in the periodic system and 31.19% in the continuous system near the injection well. Brine saturation, the duration of flow, and flow rate all impact the rate of mineral dissolution and precipitation. The intense mineral dissolution in the first cycle, which occurs to greater extremes in the continuous schedule, results in elevated saturation conditions in the brine that control further reaction in the system. Thus, the subsequent cycling of the brine results in slightly more dissolution of minerals in the periodic schedule in comparison to the continuous schedule. As a result, the initial 1.33 percentage point difference in porosity reduces over time. After 15 years of operation, the porosity difference near the injection well is only 0.38 percent point difference. At the end of the simulation, the difference in porosity furthest from the injection well is 1.71 percentage point where the final porosity is 33.08% in the continuous schedule and 34.79% in the periodic schedule. This difference is because there is less dissolution during extraction for the continuous schedule due to the effect of prolonged periods of reaction during injection which further saturate the brine as compared to the periodic schedule. In addition, the lack of transport during the storage period builds up high concentrations of species from dissolving minerals while species involved in mineral precipitation reactions are depleted, reducing the driving forces for dissolution and precipitation following the storage period (Fig S4). Similar porosity evolution occurs in the middle locations in both systems where there is a moderate increase in porosity throughout the middle of the domain (**Figure 4**) in both operational schedules even after 15 years of operation.

Most of the porosity change occurs within the first year of system operation, regardless of operational schedule. Overall, the system attained 86 and 91% of its final porosity within the first year. This can serve as a guideline to know that using CO_2 as a cushion gas will require adequate monitoring during the early stages of the project, but extensive continued modification of the formation is less likely following this period.

Other System Considerations

Variations in Flow Rate

The previous discussion is based on the comparison of model systems with equal pore volumes of acidified brine flowing through the domain during each cycle. This results in an increased injection rate for the system with the storage period. However, the periodic schedule system may instead operate with the same injection rate as in the continuous schedule as a result of the energy production rate or attempts to control the bottom hole pressure condition. Hence, this section compares the mineral evolution in the reservoir for the periodic schedule designed for equal pore volume of flow (discussed above) with a new low injection rate condition designed for equal flow rates as compared to the continuous schedule. Since we have identified higher reactivity during the early stages of cycling the working gas in this system, this investigation was carried out considering a 4 month study period to understand the impact of flowrate on the mineral evolution in the system. The relative volume fractions and relative porosity plot in **Supplementary Figure 5** are generated from the ratio of the volume fractions and porosity, respectively, during the high flow rate condition to the low flowrate condition.

Comparing the evolution of mineral volume fractions for the two different flow rates, we see that smectite and other less reactive minerals like quartz and K-feldspar are insensitive to changes in flowrate. The evolution of the relative volume fraction of muscovite shows that more muscovite dissolves at the beginning of the simulation in the location closest to the injection well under the high flow rate condition. As a result, less muscovite dissolves further away from the injection well given the higher species saturation of the brine which limits further muscovite dissolution. Overall, however, the difference in the amount of muscovite dissolution is small (3.5%) during this period. The dissolution of highly reactive minerals like siderite and calcite increases with increasing flowrate. This increase is very significant for calcite at the location closest to the injection well. The evolution of the relative volume fraction of calcite reinforces the idea that low flowrates are required to sustain the reactivity throughout the domain as less calcite dissolution occurs at the location furthest from the injection well in the high flow rate conditions. Siderite dissolution, however, is slightly higher throughout the domain in the high flow rate conditions. Overall, the impact of flow rate on porosity is limited. Higher flowrates favor increased porosity near the injection well resulting in a 2% increase in porosity in comparison to the lower flow rate conditions. Conversely, lower flow rates favor porosity increase in locations furthest from the injection well. There is limited difference in the porosity evolution for the two flowrates at the central location of the domain. While some differences in the extent of reaction occur with different flow rates, the overall change in porosity in the system is similar and limited, mainly increasing near the injection well.

Potential Permeability Evolution

This study has shown that cycling of acidified brine presents the potential for porosity change in the formation especially near the injection well. Furthermore, this porosity evolution has been shown to be somewhat sensitive to the variation in flowrate. Here, the resulting change in permeability along the domain is estimated using the Kozeny-Carmen porosity-permeability relationship (Carman, 1997) as shown below.

$$K = \frac{\varnothing^3}{c(1-\varnothing)^2 S^2} \tag{3}$$

Where K is permeability, $\varnothing$ is porosity, c is Kozeny constant, and S is specific surface area. The Kozeny-Carmen equation parameters was obtained from a study on the pore network and permeability evolution of the Paluxy formation has the value of c as 2.8 and S as $0.083\mu m^{-1}$ (Bensinger and Beckingham, 2020). Hence, the initial porosity of 24.83% was used to calculate an initial permeability of 1,414.3mD. These values were then used to develop the plot of the percentage change in porosity and percentage change in permeability along the domain after the 15 years study period in **Supplementary Figure 6**. The values of the percentage change in porosity varies from 9.79 to 28.62% and permeability varies from 31.13 to 71.40% for the periodic flow regime. The range for the percentage change in porosity for the continuous schedule is 10.21–28.63% and permeability is 32.94–72.65%. However, the change in permeability depends on the spatial distribution of mineral reactions within the pore network where pore network model simulations of reactions in the Paluxy formation in Bensinger and Beckingham (2020) found increases in porosity to 34.8% can result in permeability of ∼ 1100–7000mD. As already mentioned, the higher values of percentage change in porosity and permeability occur at the extremes of the domain. The total impacted domain length is 2.64 cm in a 15 cm domain length of porous media where there is more than an 11% change in porosity under the two flow regime. The changes in porosity and permeability presented here could be higher in a more reactive formation matrix and lower with a less reactive mineral assemblage.

CONCLUSION

Utilizing CO_2 as a cushion gas compressed energy storage in porous formations introduces new uncertainties in terms of the reactivity between the injected gas and surrounding matrix. Earlier reactive transport simulations have suggested that the extent of reactions between injected CO_2, brine, and formation minerals will be limited in compressed energy storage system in comparison to geologic CO_2 sequestration (Iloejesi and Beckingham, 2021). Here, variations in operational regime on geochemical reactions at the plume boundary have been explored.

Flow in energy storage systems with and without storage periods is distinctly different. Without a storage period, there is a constant cyclic flow away from and toward the injection well as excess energy is injected and extracted, coupled with production and demands. With a storage period, there is the addition of a stagnant or slow flow period when energy is neither injected (stored) or extracted (produced). The evolution of CO_2-bine–mineral interactions behaves differently during this storage period. Reactive transport simulations here considering geochemical reactions for an injection–storage–extraction operational schedule and injection–extraction operational schedule reveal some differences in the evolution of mineral volume fractions, predominantly for carbonate minerals. The carbonate dissolution is greater in the periodic operational schedule than with the continuous operational schedule further away from the injection well. This is because of slower dissolution rates during the storage period makes the brine to attain lower degree of saturation with respect to dissolving mineral phases. This facilitate increased dissolution near the near the source of injection during recycling of the brine in the injection–storage–extraction system in comparison to the injection–extraction system. During recycling, the source of the brine is located downstream "ghost cell" which is the furthest grid cell from the injection well. However, the overall difference in simulated porosity between the two systems is only within the range of 0.38–1.71%. Hence the operational schedule is not anticipated to largely impact the geochemical evolution of formation at the cushion gas-brine boundary. It should be noted, however, that reactivity may additionally be influenced by advective driven mixing of the cushion gas with the brine. This may result in additional variations between the operational regimes where increased mixing will occur in the continuous operational schedule.

The majority of geochemical reactions, and porosity change, at the cushion gas-brine boundary occur within the first year of operation for systems with and without storage. As such, adequate monitoring of the system at early stages can be helpful to avoid unexpected operational hazards. In both systems, porosity evolves non-uniformly with the largest increases near the injection well. Increases in porosity may reduce operational efficiencies by promoting migration of the injected gas away from the injection well if permeability additionally increases. This would require injection of additional cushion gas to maintain operational pressures. If CO_2 is used as the cushion gas, however, this may have the additional benefit of storing larger quantities of CO_2. Away from the injection well, larger increases in porosity are expected at the edge of the impacted domain, dictated by the edge of the pressure plume. Between these boundaries, the porosity changes are anticipated to be much less significant. As a result, concerns about the injectivity and working gas production rate during the operation of the storage plant could be manageable. The 15 years of operation simulated here suggest that the overall extent and impact of geochemical reactions in compressed energy storage systems utilizing CO_2 as a working gas is limited, regardless of operational schedule and associated induced flow conditions. It should be noted, however, that in a more reactive mineral assemblage, like basalt formations (Kanakiya et al., 2017), a higher degree of reactivity may be observed and the conclusions here may not apply.

DATA AVAILABILITY STATEMENT

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

CODE AVAILABILITY

The simulation code for the study is available upon request.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design, formal analysis, and investigation. The data collection and preliminary analysis and first draft of the manuscript was done by CI and all authors commented on previous versions of the manuscript. The funding for the project, supervision, reviewing and editing and resources for the projects was done by LB. All authors read and approved the final manuscript.

FUNDING

This work was supported by Auburn University by the Presidential Awards for Interdisciplinary Research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frwa.2021.689404/full#supplementary-material

Supplementary Material | The plot of volume fraction of each mineral on its axes over two cycles and 4 months. The plot of the saturation index of potential secondary minerals. The plot of each species evolution on its axes over two cycles and 4 months. Porosity plot at three mineral grid locations over two cycles.

REFERENCES

Aghahosseini, A., and Breyer, C. (2018). Assessment of geological resource potential for compressed air energy storage in global electricity supply. *Energy Conv. Manag.*169, 161–173. doi: 10.1016/j.enconman.2018.05.058

Alami, A. H., Hawili, A. A., Hassan, R., Al-Hemyari, M., and Aokal, K. (2019). Experimental study of carbon dioxide as working fluid in a closed-loop compressed gas energy storage system. *Renew. Energy* 134, 603–611. doi: 10.1016/j.renene.2018.11.046

Alkattan, M., Oelkers, E. H., Dandurand, J. L., and Schott, J. (1998). An experimental study of calcite and limestone dissolution rates as a function of pH from −1 to 3 and temperature from 25 to 80°C. *Chem. Geol.* 151, 199–214. doi: 10.1016/S0009-2541(98)00080-1

Allen, K. (1985). CAES: the underground portion. *IEEE Trans. Power Appar. Syst.* 4, 809–812. doi: 10.1109/TPAS.1985.319078

Allen, R. D. (1981). *Basis for compressed air energy storage (CAES) field test at Pittsfield, Illinois*. No. PNL-SA-9447; CONF-811066-6. Pacific Northwest Lab., Richland, WA (USA).

Allen, R. D., Doherty, T. J., Erikson, R. L., and Wiles, L. E. (1983). *Factors affecting storage of compressed air in porous-rock reservoirs*. No. PNL-4707. Pacific Northwest Lab., Richland, WA (USA). doi: 10.2172/6270908

Álvarez-Herránz, A., Balsalobre, D., Cantos, J. M., and Shahbaz, M. (2017). Energy innovations-GHG emissions nexus: fresh empirical evidence from OECD countries. *Energy Policy* 101, 90–100. doi: 10.1016/j.enpol.2016.11.030

Amram, K., and Ganor, J. (2005). The combined effect of pH and temperature on smectite dissolution rate under acidic conditions. *Geochim. Cosmochim. Acta* 69, 2535–2546. doi: 10.1016/j.gca.2004.10.001

Apergis, N., and Payne, J. (2010). Renewable energy consumption and economic growth: evidence from a panel of OECD countries. *Energy policy* 38, 656–660. doi: 10.1016/j.enpol.2009.09.002

Aslani, A., Naaranoja, M., and Zakeri, B. (2012). The prime criteria for private sector participation in renewable energy investment in the Middle East (case study: Iran). *Renew. Sustain. Energy Rev.* 16, 1977–1987. doi: 10.1016/j.rser.2011.12.015

Bachu, S. (2000). Sequestration of CO_2 in geological media: criteria and approach for site selection in response to climate change. *Energy Convers. Manag.* 41, 953–970. doi: 10.1016/S0196-8904(99)00149-1

Bauer, S., Pfeiffer, T., Boockmeyer, A., Dahmke, A., and Beyer, C. (2015). Quantifying induced effects of subsurface renewable energy storage. *Energy Proc.* 76, 633–641. doi: 10.1016/j.egypro.2015.07.885

Beckingham, L. E., Steefel, C. I., Swift, A. M., Voltolini, M., Yang, L., Anovitz, L. M., et al. (2017). Evaluation of accessible mineral surface areas for improved prediction of mineral reaction rates in porous media. *Geochim. Cosmochim. Acta* 205, 31–49. doi: 10.1016/j.gca.2017.02.006

Beckingham, L. E., and Winningham, L. (2019). Critical knowledge gaps for understanding water–rock–working phase interactions for compressed energy storage in porous formations. *ACS Sustain. Chem. Eng.* 8, 2–11. doi: 10.1021/acssuschemeng.9b05388

Bensinger, J., and Beckingham, L. E. (2020). CO_2 storage in the Paluxy formation at the Kemper County CO_2 storage complex: pore network properties and simulated reactive permeability evolution. *Int. J. Greenhouse Gas Control* 93:102887. doi: 10.1016/j.ijggc.2019.102887

Bevan, J., and Savage, D. (1989). The effect of organic acids on the dissolution of K-feldspar under conditions relevant to burial diagenesis. *Mineral. Mag.* 53, 415–425. doi: 10.1180/minmag.1989.053.372.02

Bourg, I. C., Beckingham, L. E., and DePaolo, D. J. (2015). The nanoscale basis of CO_2 trapping for geologic storage. *Environ. Sci. Technol.* 49, 10265–10284. doi: 10.1021/acs.est.5b03003

Brady, P. V., and Walther, J. V. (1990). Kinetics of quartz dissolution at low temperatures. *Chem. Geol.* 82, 253–264. doi: 10.1016/0009-2541(90)90084-K

Carden, P. O., and Paterson, L. (1979). Physical, chemical and energy aspects of underground hydrogen storage. *Int. J. Hydrogen Energy* 4, 559–569. doi: 10.1016/0360-3199(79)90083-1

Carman, P. C. (1997). Fluid flow through granular beds. *Chem. Eng. Res. Des.* 75, S32–S48. doi: 10.1016/S0263-8762(97)80003-2

Chandler, J. (2009). Trendy solutions: Why do states adopt sustainable energy portfolio standards? *Energy Policy* 37, 3274–3281. doi: 10.1016/j.enpol.2009.04.032

Chen, W. M., Kim, H., and Yamaguchi, H. (2014). Renewable energy in eastern Asia: renewable energy policy review and comparative SWOT analysis for promoting renewable energy in Japan, South Korea, and Taiwan. *Energy Policy* 74, 319–329. doi: 10.1016/j.enpol.2014.08.019

Chien, T., and Hu, J. L. (2008). Renewable energy: an efficient mechanism to improve GDP. *Energy Policy* 36, 3045–3052. doi: 10.1016/j.enpol.2008.04.012

Cui, G., Wang, Y., Rui, Z., Chen, B., Ren, S., and Zhang, L. (2018). Assessing the combined influence of fluid–rock interactions on reservoir properties and injectivity during CO_2 storage in saline aquifers. *Energy* 155, 281–296. doi: 10.1016/j.energy.2018.05.024

Dávila, G., Cama, J., Gali, S., Luquot, L., and Soler, J. M. (2016). Efficiency of magnesium hydroxide as engineering seal in the geological sequestration of CO_2. *Int. J. Greenhouse Gas Control* 48, 171–185. doi: 10.1016/j.ijggc.2016.01.031

DePaolo, D. J., and Cole, D. R. (2013). Geochemistry of geologic carbon sequestration: an overview. *Rev. Mineral. Geochem.* 77, 1–14. doi: 10.2138/rmg.2013.77.1

Duan, Z., Sun, R., Zhu, C., and Chou, I. M. (2006). An improved model for the calculation of CO_2 solubility in aqueous solutions containing Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, and SO_4^{2-}. *Mar. Chem.* 98, 131–139. doi: 10.1016/j.marchem.2005.09.001

Dunn, B., Kamath, H., and Tarascon, J. M. (2011). Electrical energy storage for the grid: a battery of choices. *Science* 334, 928–935. doi: 10.1126/science.1212741

Eckroad, S., and Gyuk, I. (2003). *EPRI-DOE Handbook of Energy Storage for Transmission and Distribution applications.* Washington, DC: Electric Power Research Institute, Inc, 3–35.

Fischer, S., Liebscher, A., Wandrey, M., and C. O.2S. I. N. K., Group. (2010). CO_2-brine–rock interaction—First results of long-term exposure experiments at in situ P–T conditions of the Ketzin CO_2 reservoir. *Geochemistry* 70, 155–164. doi: 10.1016/j.chemer.2010.06.001

Fleming, M. R., Adams, B. M., Randolph, J. B., Ogland-Hand, J. D., Kuehn, T. H., Buscheck, T. A., et al. (2018). "High efficiency and large-scale subsurface energy storage with CO_2," in *43rd Workshop on geothermal reservoir engineering, Stanford, CA*.

Golubev, S. V., Bénézeth, P., Schott, J., Dandurand, J. L., and Castillo, A. (2009). Siderite dissolution kinetics in acidic aqueous solutions from 25 to 100°C and 0 to 50 atm pCO_2. *Chem. Geol.* 265, 13–19. doi: 10.1016/j.chemgeo.2008.12.031

Gunter, W. D., Perkins, E. H., and McCann, T. J. (1993). Aquifer disposal of CO_2-rich gases: reaction design for added capacity. *Energy Convers. Manag.* 34, 941–948. doi: 10.1016/0196-8904(93)90040-H

Haar, N., and Theyel, G. (2006). US electric utilities and renewable energy: drivers for adoption. *Int. J. Green Energy* 3, 271–281. doi: 10.1080/01971520600704043

Helgeson, H. C., and Kirkham, D. H. (1974). Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures; II, Debye-Huckel parameters for activity coefficients and relative partial molal properties. *Am. J. Sci.* 274, 1199–1261. doi: 10.2475/ajs.274.10.1199

Huang, M. Y., Alavalapati, J. R., Carter, D. R., and Langholtz, M. H. (2007). Is the choice of renewable portfolio standards random?. *Energy Policy* 35, 5571–5575. doi: 10.1016/j.enpol.2007.06.010

Huq, F., Haderlein, S. B., Cirpka, O. A., Nowak, M., Blum, P., and Grathwohl, P. (2015). Flow-through experiments on water–rock interactions in a sandstone caused by CO_2 injection at pressures and temperatures mimicking reservoir conditions. *Appl. Geochem.* 58, 136–146. doi: 10.1016/j.apgeochem.2015.04.006

IEA (2019). SDG7: Data and Projections. *SDG7: Data and Projections*. Available online at: https://www.iea.org/data-and-statistics/?country=WORLD&fuel=Renewables%20and%20waste&indicator=SDG72 (accessed August 15, 2020)

Iloejesi, C. O., and Beckingham, L. E. (2021). Assessment of geochemical limitations to utilizing CO_2 as a cushion gas in compressed energy storage systems. *Environ. Eng. Sci.* 38, 115–126. doi: 10.1130/abs/2021SE-362383

Jenner, S., Chan, G., Frankenberger, R., and Gabel, M. (2012). What drives states to support renewable energy? *Energy J.* 33, 1–12. doi: 10.5547/01956574.33.2.1

Kanakiya, S., Adam, L., Esteban, L., Rowe, M. C., and Shane, P. (2017). Dissolution and secondary mineral precipitation in basalts due to reactions with carbonic acid. *J. Geophys. Res. Solid Earth* 122, 4312–4327. doi: 10.1002/2017JB014019

Kharaka, Y. K., and Cole, D. R. (2011). Geochemistry of geologic sequestration of carbon dioxide. *Front. Geochem. Contrib. Geochem. Study Earth* 2011, 133–174. doi: 10.1002/9781444329957.ch8

Kim, J., Choi, J., and Park, K. (2015). Comparison of nitrogen and carbon dioxide as cushion gas for underground gas storage reservoir. *Geosyst. Eng.* 18, 163–167. doi: 10.1080/12269328.2015.1031916

Knauss, K. G., and Wolery, T. J. (1988). The dissolution kinetics of quartz as a function of pH and time at 70 C. *Geochim. Cosmochim. Acta* 52, 43–53. doi: 10.1016/0016-7037(88)90055-5

Lin, B., and Zhu, J. (2019). The role of renewable energy technological innovation on climate change: empirical evidence from China. *Sci. Total Environ.* 659, 1505–1512. doi: 10.1016/j.scitotenv.2018.12.449

Lund, P. D. (2007). Effectiveness of policy measures in transforming the energy system. *Energy Policy* 35, 627–639. doi: 10.1016/j.enpol.2006.01.008

Lyon, T. P., and Yin, H. (2010). Why do states adopt renewable portfolio standards?: An empirical investigation. *Energy J.* 31, 133–157. doi: 10.5547/ISSN0195-6574-EJ-Vol31-No3-7

Ma, J., Li, Q., Kempka, T., and Khn, M. (2019). Hydromechanical response and impact of gas mixing behavior in subsurface CH_4 storage with CO_2-based cushion gas. *Energy Fuels* 33, 6527–6541. doi: 10.1021/acs.energyfuels.9b00518

Mai, T., Cole, W., Lantz, E., Marcy, C., and Sigrin, B. (2016). *Impacts of federal tax credit extensions on renewable deployment and power sector emissions*. No. NREL/TP-6A20-65571. National Renewable Energy Lab. (NREL), Golden, CO (United States). doi: 10.2172/1239642

Marques, A. C., Fuinhas, J. A., and Manso, J. P. (2011). A quantile approach to identify factors promoting renewable energy in European countries. *Environ. Resour. Econ.* 49, 351–366. doi: 10.1007/s10640-010-9436-8

Mensah, C. N., Long, X., Boamah, K. B., Bediako, I. A., Dauda, L., and Salman, M. (2018). The effect of innovation on CO 2 emissions of OCED countries from 1990 to 2014. *Environ. Sci. Pollut. Res.* 25, 29678–29698. doi: 10.1007/s11356-018-2968-0

Nathenson, M., and Guffanti, M. (1988). Geothermal gradients in the conterminous United States. *J. Geophys. Res. Solid Earth* 93, 6437–6450. doi: 10.1029/JB093iB06p06437

Oelkers, E. H., Schott, J., Gauthier, J. M., and Herrero-Roncal, T. (2008). An experimental study of the dissolution mechanism and rates of muscovite. *Geochim. Cosmochim. Acta* 72, 4948–4961. doi: 10.1016/j.gca.2008.01.040

Oldenburg, C. M. (2003). Carbon dioxide as cushion gas for natural gas storage. *Energy Fuels* 17, 240–246. doi: 10.1021/ef020162b

Oldenburg, C. M., and Pan, L. (2013). Porous media compressed-air energy storage (PM-CAES): theory and simulation of the coupled wellbore–reservoir system. *Oldenburg C. M. Transp. Porous Media* 97, 201–221. doi: 10.1007/s11242-012-0118-6

Pashin, J. C., Achang, M., Chandra, A., Folaranmi, A., Martin, S., Meng, J., et al. (2018). "The Paluxy Formation in the East-Central Gulf of Mexico Basin: geology of an Ultra-Giant Anthropogenic CO_2 Sink," in *AAPG ACE* 2018.

Petrusak, R., Cyphers, S., Bumgardner, S., Hills, D., Pashin, J., and Esposito, R. (2010). "Saline reservoir storage in an active oil field: extracting maximum value from existing data for initial site characterization; Southeast Regional Carbon Sequestration Partnership (SECARB) phase III," in *SPE international conference on CO_2 capture, storage, and utilization*. Society of Petroleum Engineers. doi: 10.2118/139700-MS

Pfeiffer, W. T., and Bauer, S. (2015). Subsurface porous media hydrogen storage–scenario development and simulation. *Energy Proc.* 76, 565–572. doi: 10.1016/j.egypro.2015.07.872

Pfeiffer, W. T., Beyer, C., and Bauer, S. (2017). Hydrogen storage in a heterogeneous sandstone formation: dimensioning and induced hydraulic effects. *Petrol. Geosci.* 23, 315–326. doi: 10.1144/petgeo2016-050

Qin, F., and Beckingham, L. E. (2019). Impact of image resolution on quantification of mineral abundances and accessible surface areas. *Chem. Geol.* 523, 31–41. doi: 10.1016/j.chemgeo.2019.06.004

Qin, F., and Beckingham, L. E. (2021). The impact of mineral reactive surface area variation on simulated mineral reactions and reaction rates. *Appl. Geochem.* 124, 104852. doi: 10.1016/j.apgeochem.2020.104852

Reysa, G. (2005). Ground Temperatures as a Function of Location, Season, and Depth. Build It Solar.[Online]. Available online at: https://www.builditsolar.com/Projects/Cooling/EarthTemperatures.htm (accessed May 01, 2019).

Robinson, H., and Davis, T. (2012). "Seismic reservoir characterization of distributary channel sandstones in the Lower Cretaceous Paluxy reservoir, Delhi Field, Louisiana," in *SEG Technical Program Expanded Abstracts 2012*, 1–6. Society of Exploration Geophysicists. doi: 10.1190/segam2012-0164.1

Rul Internal Revenue Service (2009). Credit for Carbon Dioxide Sequestration Under Section 45Q. Available online at: https://www.irs.gov/irb/2009-44_IRB#NOT-2009-83 (accessed September 10, 2020)

Sadorsky, P. (2009a). Renewable energy consumption and income in emerging economies. *Energy Policy* 37, 4021–4028. doi: 10.1016/j.enpol.2009.05.003

Sadorsky, P. (2009b). Renewable energy consumption, CO_2 emissions and oil prices in the G7 countries. *Energy Econ.* 31, 456–462. doi: 10.1016/j.eneco.2008.12.010

Steefel, C. I., Appelo, C. A., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., et al. (2015). Reactive transport codes for subsurface environmental simulation. *Computat Geosci.* 19, 445–478. doi: 10.1007/s10596-014-9443-x

Steefel, C. I., and Molins, S. (2009). CrunchFlow: Software for modeling multicomponent reactive flow and transport. User's manual. Berkeley: Lawrence Berkeley National Laboratory.

Succar, S., and Williams, R. H. (2008). Compressed air energy storage: theory, resources, and applications for wind power. *Princeton Environmental Institute Report* 8:81.

van der Linden, S. (2006). Bulk energy storage potential in the USA, current developments and future prospects. *Energy* 31, 3446–3457. doi: 10.1016/j.energy.2006.03.016

Walters, A. B. (1976). "Technical and environmental aspects of underground hydrogen storage," in *1st World Hydrogen Energy Conference*, vol. 2, pp. 2B_65–2B_79.

Wang, B., Wang, Q., Wei, Y. M., and Li, Z. P. (2018). Role of renewable energy in China's energy security and climate change mitigation: An index decomposition analysis. *Renew. Sustain. Energy Rev.* 90, 187–194. doi: 10.1016/j.rser.2018.03.012

Xu, T., Apps, J. A., Pruess, K., and Yamamoto, H. (2007). Numerical modeling of injection and mineral trapping of CO_2 with H_2S and SO_2 in a sandstone formation. *Chem. Geol.* 242, 319–346. doi: 10.1016/j.chemgeo.2007.03.022

Xu, T., E., Sonnenthal, N., Spycher, and Zheng, L. (2017). TOUGHREACT V3. 32 Reference Manual: a Parallel Simulation Program for Non-isothermal Multiphase Geochemical Reactive Transport. *Lawrence Berkeley National Laboratory, Report LBNL-Draft, Berkeley, Calif* ,

Zhang, L., Soong, Y., Dilmore, R., and Lopano, C. (2015). Numerical simulation of porosity and permeability evolution of Mount Simon sandstone under geological carbon sequestration conditions. *Chem. Geol.* 403, 1–12. doi: 10.1016/j.chemgeo.2015.03.014

Zhang, S., and DePaolo, D. J. (2017). Rates of CO_2 mineralization in geological carbon storage. *Accounts Chem. Res.* 50, 2075–2084. doi: 10.1021/acs.accounts.7b00334

Zhang, Z., and Huisingh, D. (2017). Carbon dioxide storage schemes: technology, assessment and deployment. *J. Cleaner Prod.* 142, 1055–1064. doi: 10.1016/j.jclepro.2016.06.199

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.

frontiers in Water

REVIEW
published: 29 September 2021
doi: 10.3389/frwa.2021.696630

Drivers for Intermittent Water Supply in India: Critical Review and Perspectives

*Anujkumar Ghorpade, Abhishek Kumar Sinha and Pradip P. Kalbar**

Center for Urban Science and Engineering, Indian Institute of Technology Bombay, Mumbai, India

Intermittent Water Supply (IWS) is prevalent in most developing countries. Specifically, in India, IWS is existent throughout the country. Many studies focus on documenting the effects of IWS, and rarely the drivers of the IWS regime are studied. In this study, a systematic literature review was conducted on IWS studies around the globe. The various causes for IWS were documented. Then, by studying India's typical water supply system (WSS) configuration, the vicious cycle of IWS in India is discussed. Further, the drivers of IWS were identified and elaborated with the causing mechanisms. This knowledge will help devise strategies and solutions for improving the IWS in India and other developing countries with similar socio-economic conditions.

Keywords: continuous water supply, India, intermittent water supply, vicious cycle, water supply system design, water supply system operation

Edited by:
Auroop Ratan Ganguly,
Northeastern University, United States

Reviewed by:
Katarzyna Pietrucha-Urbanik,
Rzeszów University of Technology, Poland
Ramakrishna Tipireddy,
Pacific Northwest National Laboratory (DOE), United States

****Correspondence:***
Pradip P. Kalbar
kalbar@iitb.ac.in

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *17 April 2021*
Accepted: *31 August 2021*
Published: *29 September 2021*

Citation:
Ghorpade A, Sinha AK and Kalbar PP (2021) Drivers for Intermittent Water Supply in India: Critical Review and Perspectives. Front. Water 3:696630. doi: 10.3389/frwa.2021.696630

INTRODUCTION

Piped water supply is the safest way to provide potable water to consumers in adequate quantity and recommended quality. The planning, design, construction, operation, and maintenance of a Water Supply System (WSS) requires comprehensive efforts from multiple stakeholders. Due to reliable design and operation in developed countries as per the standards, Continuous Water Supply (CWS) prevails. However, the situation is diametrically opposite in developing countries. Although designs of WSSs are as per stipulated standards and manuals, the operation of the WSS is not as expected in design (Abu-Madi and Trifunovic, 2013). Especially in the South Asian countries, 100% Intermittent Water Supply (IWS) exists (Charalambous and Laspidou, 2017). Typically, in India, the design is based on the CWS guidelines, and the system operates as the IWS. A piped water supply distributing water to consumers for less than 24 h a day is termed as IWS (Mokssit et al., 2018).

In India, all the WSSs are operated as an IWS, with supply duration varying from 1 to 6 h/day (Ahluwalia et al., 2011). The IWS results in inequitable supply, increased Non-Revenue Water (NRW), and deteriorated water quality. Most of the WSS in India fail to comply with the service level benchmark (**Table 1**). The standard is based on Central Public Health and Environmental Engineering Organization (CPHEEO) norms, an advisory board to the Ministry of Urban Development, Government of India.

Many past studies (Klingel, 2012; Galaitsi et al., 2016; Kumpel and Nelson, 2016; Mokssit et al., 2018; Simukonda et al., 2018) have focused on the causes and the consequences of IWS worldwide. **Table 2** illustrates the past review studies on IWS, which depicts the studies' objective and significant findings. Additionally, few studies (Andey and Kelkar, 2009; Kumpel and Nelson, 2013) reported the comparison of CWS and IWS with the study area as India.

TABLE 1 | Service level benchmarks for water supply in India (MoUD, 2012).

Sr. No.	Indicator	Unit	Benchmark	Average
1	Coverage of connections	%	100	50.2
2	Per Capita Supply	lpcd	135	69.2
3	Metering of connections	%	100	13.3
4	Non-Revenue water	%	20	32.9
5	Continuity of supply	Hours	24	3.1
6	Quality and Treatment	%	100	81.7

Another set of studies (Elala et al., 2011; Mellor et al., 2016) reported the IWSs impact on water quality in India. Apart from reporting the causes and consequences of the IWS, studies reported the methodology for the direct transition of IWS to the CWS (Jayaramu et al., 2015; Hastak et al., 2017; Burt et al., 2018). However, these studies were based on the pilot scale level, and achieving CWS throughout the WSS is a distant dream in developing countries (Mokssit et al., 2018; Kalbar and Gokhale, 2019).

Limited literature is available on the improvement of existing IWS. To solve the IWS issues, understanding the root cause of the intermittency is of prime importance. The causes and consequences of IWS enlisted in the past literature are primarily generic. A detailed understanding of the interaction between various drivers responsible for IWS is necessary to propose interventions for improving existing IWS. To the best of our knowledge, a comprehensive description of various drivers responsible for IWS in India is missing in the literature. The present study captures the different design, operation, economic, and institutional arrangement issues of WSS responsible for IWS in India. Based on the field condition and operational hydraulics, various drivers of IWS are discussed. It is hoped this study will bring more clarity toward understanding the drivers responsible for IWS in India.

The paper includes five sections. After this introduction, the second section is about the review methodology adopted in the study. The third section is focused on the analysis of various studies on IWS. The drivers for IWS in India are discussed in section four. Finally, the conclusion of the study is reported in section Vicious cycle of WSS failure in India.

REVIEW METHODOLOGY

In this study, only peer-reviewed journal articles were selected as the novelty of work is given priority in the peer-reviewed journal articles (Owens et al., 2020). The journal articles were retrieved from the Scopus database. The search string used was "Intermittent Water Supply," and the search was conducted for "Article title, abstract, and keywords." A total of 1080 articles were reported from 1873 to 2020, and all the articles were screened in detail. The first screening level was based on reading the title, abstract, keywords, and excluding the duplication, which resulted in 201 publications. A detailed reading of the title, abstract, introduction, and conclusion was done in the second level of screening, and 91 articles were selected for the detailed review.

The articles were classified based on three criteria: design and analysis, operation, and Economic aspects (revenue). The main aim of these studies was used as criteria while classifying the articles. It was observed that studies reported before 2002 (and in the years 2004 and 2005) were not entirely focusing on IWS, hence excluded from the review process. As shown in **Figure 1**, a substantial amount of work is done on the IWS operation aspect. As expected, a significant number of studies have considered WSS in developing countries, especially India. Additionally, based on knowledge gained through numerous field visits to WSSs across India, the existing layout, design and operational practices are discussed, followed by the drivers for IWS in India.

STUDIES ON IWS

The past studies (91 screened articles) on IWS are discussed based on three criteria Design and Analysis, Operation, and Revenue. The bibliographic information of the articles included in the survey is provided in the Supplementary Information (**Supplementary Table SI-1**). The category wise summary of each study is compiled in **Supplementary Tables SI-2.1–2.3** of the Supplementary Information.

Design and Analysis of IWS

Design and analysis are an essential part of WSS irrespective of the supply regime. Out of 91 articles screened, only three studies focused on a new set of design guidelines for IWS (**Table 3**). However, the guidelines are not available in the public domain. From the analysis point of view, most of the studies (almost 50%) focused on simulating the process of partial pipe filling in IWS. Although Pressure Driven Analysis (PDA) was used to simulate the pipeline's partial flow condition, Demand Driven Analysis (DDA) is a standard practice to simulate the WDN with IWS. Simulating partial flow conditions is one of the steps for identifying the problematic areas in the network. Once the problems are identified, the next step is to improve the existing network.

The studies (**Table 3**) on rehabilitation of existing networks with IWS were based on the theoretical model. Validation of the proposed theoretical models using various WDN is missing. Also, the proposed models failed to capture the actual ground reality of the IWS. Instead of studies focusing on simulating the actual ground reality of IWS, studies on direct conversion of IWS to CWS were encountered. While modeling IWS, emphasis should be given to analyzing household storage's impact on the overall network performance. Overall, a holistic design and analysis approach covering all the elements of the actual operation

TABLE 2 | Review studies on IWS.

Author	Objective	Findings
Klingel (2012)	To enlist the technical causes and impacts of IWS	• Deficiency in the system concept, system knowledge, system planning, and infrastructure management are the leading causes of IWS. • Infrastructure deterioration, loss of water, inequitable supply, and reduced water quality are some of the impacts of IWS. • Need for a detailed study on each aspect of IWS.
Galaitsi et al. (2016)	To study the complex interrelation between the causes and impacts of IWS	• Lack of consistency in the definition of IWS. The authors proposed three types of intermittency, predictable irregular, and unreliable intermittency. • Predicting the type of intermittency is essential for appropriate interventions.
Kumpel and Nelson (2016)	To examine the IWS status in the world, nature of IWS, and its effect on water quality.	• Intrusions, backflow, biofilms, loose deposits, and microbial growth are the main reasons for the degradation of water quality in IWS. • Improvement in water quality monitoring in IWS is required. • Need for descriptive research to understand the IWS practices and methods responsible for water quality deterioration.
Mokssit et al. (2018)	To propose a methodology for assessing service quality of WSS with IWS regime	• The interrelation between causes and problems of IWS gives rise to a vicious cycle. • Based on availability, quantity, quality, accessibility, and affordability of water, the service quality of WSS with IWS regime can be determined.
Simukonda et al. (2018)	To review key water supply intermittency casual factors, problems, and options for improvement	• IWS tree comprising the root cause and problems due to IWS. • Appropriate design and operation interventions can improve IWS, instead of impractical direct conversion of IWS to CWS.
Bautista-de los Santos et al. (2019)	To review the impact of IWS on the drinking water microbiome	• A combination of lab studies and field studies is necessary to understand IWS impact on water quality.
Al-Washali et al. (2020)	To review different assessment methods of water loss component for IWS	• The accuracy and uncertainty of methods is directly proportional to each other for calculating the water loss.

of WSS in developing countries is required to minimize the system's failure.

Operation of IWS

The operation of a WSS based on IWS results in various problems like increased Non-Revenue Water (NRW), inequitable supply, deterioration of water quality, etc. In relation to this, a significant amount of studies (**Table 4**) are reported to either quantify the problems or ways to improve the efficiency of IWS. The studies conducted regarding the operational aspect of the IWS focused on enlisting impacts of IWS, reducing the leakages, reasons for deterioration of water quality, improving the supply equity, and converting to the CWS. Among the impacts of IWS, leakage is the most widely studied area. The different components vulnerable to the leakage in the WSS and various leakage quantification techniques suitable for IWS are reported in the literature (**Supplementary Table SI-2.2**). However, the proposed methodologies are labor-intensive, costly, and tested on a pilot scale, resulting in replicability issues. Hence, the exact quantification of water loss in IWS remains a tedious task.

From the water quality aspect in IWS, most of the studies reported the household storage tank is the main component responsible for the deterioration of water quality. Limited studies (**Supplementary Table SI-2.2**) highlighted the pathway for improving the water quality in IWS.

Besides the leakage and water quality, inequitable water supply is a significant drawback in IWS. However, studies on replicable methodology to achieve equity in IWS are scarce (SI-2.2). The limited literature (**Table 4**) on methods to achieve equitable water supply in the IWS suggests a lack of willingness to improve the IWS's performance.

The focus is on direct conversion of IWS to CWS (Klingel and Nestmann, 2014; Jayaramu et al., 2015; Hastak et al., 2017). However, the studies focus on a city or town pilot area for conversion to CWS. The scaling up of direct conversion to the CWS throughout the city is difficult without completely understanding the ground reality and gradually improving the existing system (Kalbar and Gokhale, 2019). Moreover, the lack of consistency in defining benchmarking parameters and reliability in the data collection framework makes it challenging to utilize the benchmarking data to directly implement the interventions (Rawas et al., 2020).

Revenue in IWS

Like all other infrastructure projects, revenue generation in WSS is the driving force for successful operation of a scheme. However, the failure of WSS with IWS leads to a lack of consumer satisfaction, which minimizes willingness to pay and ultimately reduces revenue generation.

Due to the reduction in efficiency of WSS with IWS, the consumers water demand is not fulfilled. Hence, consumers

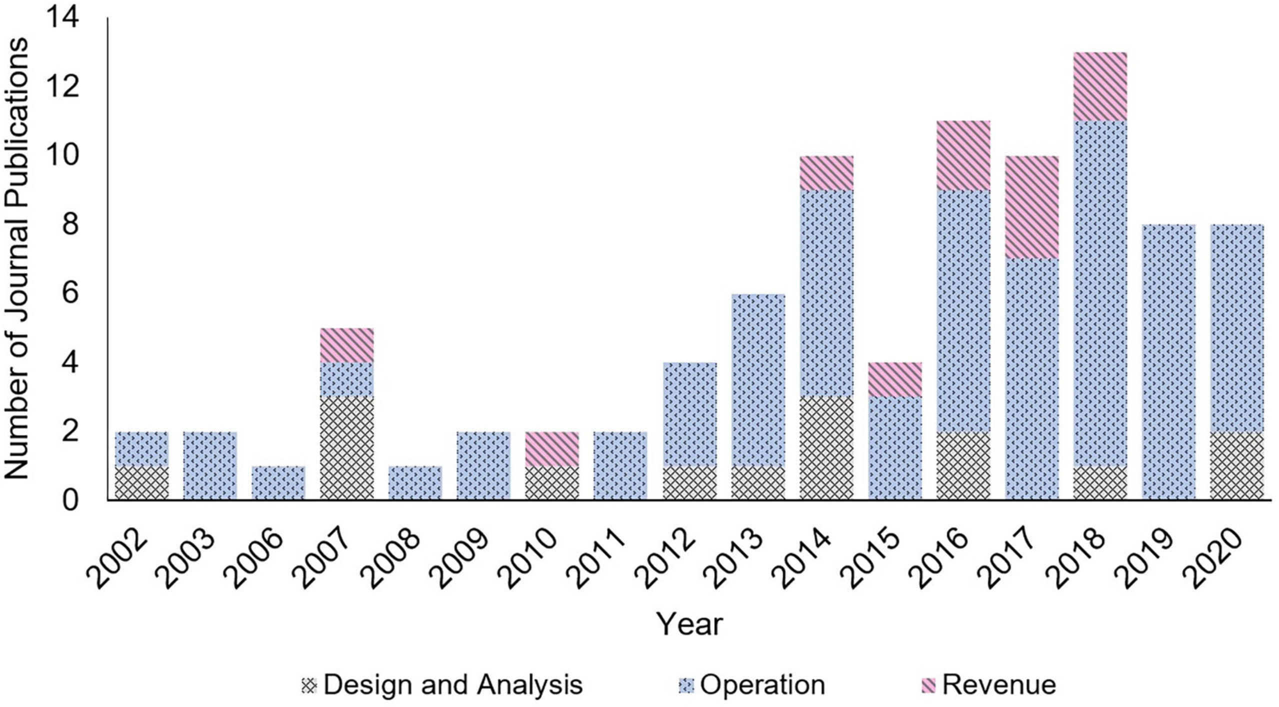


FIGURE 1 | Screened articles on Intermittent Water Supply from 2002 to 2020.

spend on household storage, pumps, underground tanks to store water in IWS. The additional expenditure on the solution mentioned above is termed as coping cost to consumers (McIntosh, 2003). In relation to coping costs, most of the studies (**Supplementary Table SI-2.3**) classified the various expenditure consumer have to incur on various components to satisfy the demand in IWS. Moreover, factors responsible for the willingness to pay from the consumer end were discussed in the literature. From service providers point of view, exact quantification of water consumption is vital to generate appropriate water bills. In IWS, accurate determination of water consumption is difficult to achieve (Jayaramu et al., 2016; Kumpel et al., 2017). A methodology to accurately quantify the household's actual consumption with IWS was proposed by Guragai et al. (2018) using data loggers in Kathmandu, Nepal. However, the scaling up of the proposed methodology is difficult, as a considerable investment is required to install and maintain the data loggers at a larger scale.

IWS IN INDIA

This section describes the configuration, design norms, and the actual operation of WSSs in India. The actual ground reality will help understand prevailing IWS both in urban and rural settings in India.

Configuration of WSSs in India

A typical layout of WSS with a surface water source in India is illustrated in **Figure 2A**. Raw water from the source is transported to the Water Treatment Plant (WTP) through gravity or using pumps based on the terrain. The pipe carrying the water from the source to WTP is typically called the Rising(pumped)/Gravity main. The treated water from WTP is then stored in a Master Balancing Reservoir (MBR). From the MBR, water is transported to different Elevated Storage Reservoir (ESR) through a pipe network known as a transmission network. Ideally, MBR is situated at a high altitude such that water is distributed to all the ESRs under gravity. Preferably, the ESR is supposed to be at the center of the serving area. Water from ESR to the consumer is delivered through a network of pipelines termed distribution network. ESR is provided with a separate inlet and outlet pipe. Different valves are provided at appropriate locations to maintain the efficiency of the system. The configuration is such that water is transported in a single direction from source to consumer.

The difference between the WSS in developed countries and India was studied based on technical norms for water supply laid by the America Water Works Association (AWWA). As per AWWA manual M31 (AWWA American Water Works Association, 2008), the most efficient WSS configuration is one with a floating Service Reservoir (SR) situated at the end of the consumer center (**Figure 2B**). In other words, the consumer center should be between the pump station and SR. Water from the pumping station is pumped to both the consumer center and the SR. Unlike the ESRs in India, SR has only a single pipe to fill and withdraw water. At the time of emergency like fire demand, water is provided to the consumer center from both pump station and SR. Whereas, at the time of the off-peak period,

TABLE 3 | Studies contributing to design and analysis aspect of IWS.

References	Location	Issue of IWS addressed
Vairavamoorthy and Elango (2002)	N/A	Design
Vairavamoorthy et al. (2007a,b)	India	Design
Fontanazza et al. (2007), De Marchis et al. (2010), Campisano et al. (2018)	Italy	Analysis—Partial pipe filling process
Nyende-Byakika et al. (2012)	N/A	Analysis—Partial pipe filling process
Nyende-Byakika et al. (2013)	Uganda	Analysis—Partial pipe filling process
Lieb et al. (2016)	N/A	Analysis—Partial pipe filling process
Mohan and Abhijith (2020)	India	Analysis—Partial pipe filling process
Manohar and Mohan Kumar (2014), Mohapatra et al. (2014), Neelakantan et al. (2014)	India	Analysis- Rehabilitation of existing WSS with IWS and conversion to CWS
Haddad et al. (2016)	Iran	Analysis- Rehabilitation of existing WSS with IWS and conversion to CWS
El Achi and Rouse (2020)	Jordan	Analysis- Rehabilitation of existing WSS with IWS and conversion to CWS

TABLE 4 | Studies contributing to operation aspect of IWS.

References	Location	Issue of IWS addressed
Andey and Kelkar (2007, 2009), Kumar et al. (2018)	India	Impacts on WSS and consumers
Christodoulou and Agathokleous (2012), Agathokleous and Christodoulou (2016), Agathokleous et al. (2017)	Cyprus	Impacts on WSS and consumers
Abu-Madi and Trifunovic (2013)	Jordan	Impacts on WSS and consumers
Fontanazza et al. (2013)	Italy	Impacts on WSS and consumers
Fan et al. (2014)	China	Impacts on WSS and consumers
Al-Ghamdi and Gutub (2002), Al-Ghamdi (2011), Haider et al. (2019)	Saudi Arabia	Quantification of leakages
Criminisi et al. (2009), De Marchis et al. (2013)	Italy	Quantification of leakages
Tamari and Ploquet (2012)	Mexico	Quantification of leakages
Zyoud et al. (2016), Zyoud and Fuchs-Hanusch (2019)	Palestine	Quantification of leakages
Mastaller and Klingel (2017), Taylor et al. (2019)	India	Quantification of leakages
Al-Washali et al. (2020), Aboelnga et al. (2018)	Jordan	Quantification of leakages
Al-Washali et al. (2019)	Yemen	Quantification of leakages
Coelho et al. (2003)	Jordan, Lebanon, Palestine, United Kingdom, and Portugal	Water quality
Tokajian and Hashwa (2003), Ayoub and Malaeb (2006)	Lebanon	Water quality
Elala et al. (2011), Kumpel and Nelson (2013), Kumpel and Nelson (2014), Ercumen et al. (2015), Khadse et al. (2016), Taylor et al. (2018)	India	Water quality
Haddad et al. (2014)	NA	Water quality
Hernandez-Lopez et al. (2016)	N/A	Water quality
Bivins et al. (2017)	N/A	Water quality
Erickson et al. (2017)	Panama	Water quality
Alazzeh et al. (2019)	Palestine	Water quality
Rubino et al. (2019)	Mexico	Water quality
Li et al. (2020)	China	Water quality
Sakomoto et al. (2020)	Uganda	Water quality
Rosenberg et al. (2008)	Jordan	Inequitable water supply
Ameyaw et al. (2013)	N/A	Inequitable water supply
Freni et al. (2014)	Italy	Inequitable water supply
Gottipati and Nanduri (2014)	N/A	Inequitable water supply
Solgi et al. (2015)	Iran	Inequitable water supply
Ilaya-Ayza et al. (2017)	Bolivia	Inequitable water supply
Strijdom et al. (2017)	South Africa	Inequitable water supply

(Continued)

water is pumped in SR and stored. Thus, the SR is floating on the network, with the flexibility to store and deliver water based on the consumer center's demand. As shown in **Figure 2B**, the WSS configuration is one of the primary reasons for its high reliability in developed countries. In addition to the layout, the design norms are an essential factor responsible for WSS's performance. The subsequent section will highlight the design norms of WSS in India and developed countries.

WSS Design Norms in India

In India, the WSSs are designed as per the CPHEEO norms. The guidelines include the per capita demand, minimum residual pressure, minimum pipe diameter, the design period of various components, peak factor, and demand pattern. **Table 5** enlists the different norms as per CPHEEO and AWWA Manual M31. Even though both the standards are based on CWS, there is a gap between standards laid by two agencies. Apart from the minimum residual criteria and per capita norm, the significant difference is in fire flow demand. In the USA, the pipe diameter is designed based on the maximum day demand and fire flow demand (Walski, 2014). The pipe diameter should satisfy the combined demand with the minimum residual pressure shown in **Table 5**. Whereas in India, additional fire demand is considered for calculating the storage capacity of ESR. Fire demand is not included in the design of the distribution network. Moreover, Smith and Liu (2020) enlisted the minimum pressure standards around the world. The dissimilarity in the design

TABLE 4 | Continued

References	Location	Issue of IWS addressed
Klingel and Nestmann (2014)	Algeria	Conversion to CWS
Jayaramu et al. (2015), Hastak et al. (2017)	India	Conversion to CWS
Ilaya-Ayza et al. (2016, 2018)	Bolivia	Conversion to CWS
David et al. (2020)	Mexico	Conversion to CWS
Cronk and Bartram (2018)	Honduras, Nicaragua, and Panama	Database management
Kaminsky and Kumpel (2018)	N/A	Database management
Rawas et al. (2020)	Peru	Database management

norms between developed and developing nations makes the WSS in the developed countries less vulnerable to failure than the developing countries.

Apart from the design parameters mentioned in **Table 5**, the transmission network between MBR and ESR is designed for a 1.2 to 1.5 peak factor in India. On the distribution side, the consumer demand pattern is considered a diurnal pattern, peaking in the morning and evening. Population forecast is done for the design period of 30 years (CPHEEO, 1999). The various methods mentioned in the CPHEEO manual are used to arrive at the forecasted population, and per capita demand (**Table 5**) is multiplied to meet the future demands. MBR and ESR's capacity is finalized based on the mass curve method, based on the supply and demand pattern throughout the day. Furthermore, Demand Driven Approach (DDA) is used to design transmission and distribution networks.

The WSS design norms in India cover all the fundamental aspects. A Detailed Project Report (DPR) is formed based on the design norms and submitted to concerned authorities for design approval. The WSS construction is carried out based on the DPR, followed by its commissioning and operation. The following section will highlight the actual operation of WSS in India.

Actual Operation of WSS in India

The operation of WSS in India is intermittent, both in transmission and distribution networks. The operation is different from the design guidelines used. In the following subsection, the operation of the transmission and distribution network is discussed separately.

Operation of the Transmission System

The transmission network is designed based on CWS, assuming that all the downstream ESRs will be filled simultaneously. Conversely, the ESRs are filled in a staggered manner with the help of control valves (ON/OFF valves). A peak factor between 1.2 and 1.5 is used to design the transmission network CPHEEO, 1999. However, the transmission network's staggering operation increases the peak factor, deterioration of the pipeline, increasing the water loss.

The design life of a transmission network is 30 years CPHEEO, 1999. Accordingly, the demand is calculated, and pipe diameter is determined and commissioned at the site. The buffer capacity available at the initial phase of the design period is utilized to expand the network unscientifically. Such a scenario results in over withdrawal in the transmission network to fulfill the increased demand. The uncontrolled withdrawal causes partially filling of the ESRs especially, at the tail end ESRs.

Furthermore, to counter the ground undulations, Break Pressure Tanks (BPT) are used at a higher elevation to reduce the energy consumption by converting the pumped flow into gravity flow. However, due to the inappropriate location of BPT, the actual required discharge through the pipe is not achieved, aggravating the issue of partial filling of ESRs. Finally, illegal tapings on the main transmission line are commonly found in India. The unauthorized connections reduce the system pressure, which already fails to deliver the desired flow.

Operation of the Distribution System

In India, a large-scale centralized approach is used to design WSS (**Figure 2A**). Large ESRs with a capacity between 5–10 ML and 10–25 ML (Kalbar and Gokhale, 2019) are designed and constructed to cater to fluctuations between demand and supply. As the operation is based on IWS, the large storage tanks develop a perception of the operators that an infinite volume of water is available, which results in the expansion of the distribution network beyond the hydraulic capacity (influence) of the tank. Furthermore, due to land constraints in urban areas, ESRs are generally located far away from the serving area, resulting in an inefficient distribution network. The consumers at the end of network are deprived of water, and the consumers near the tank receive ample water.

The general practice for the operation of WDN is to divide the service area with the help of control valves. The supply time is bifurcated with the use of control valves. The division is based on operator experience, and the aim is to supply water to consumers irrespective of the residual pressure, resulting in inequitable distribution. Sometimes, the WDN from different storage tanks is interconnected, and it becomes difficult to demarcate the exact location of the pipelines. In other words, the concept of isolated District Metered Area (DMA) is failed at the ground level.

Moreover, as discussed in the previous section, the distribution network is designed based on the peak factor as per the diurnal demand pattern. As water is supplied to a particular area with the help of a valve for a limited time, the peak factor exerted is higher than the one considered in the design. For example, if the water is supplied to an area for 2 h a day, the actual peak factor exerted in the respective pipe network will be 12. The high peak factor results in the deterioration of pipes and increased NRW.

VICIOUS CYCLE OF WSS FAILURE IN INDIA

In the previous section, the actual design and operation of WSS in India were discussed. Further, it is necessary to

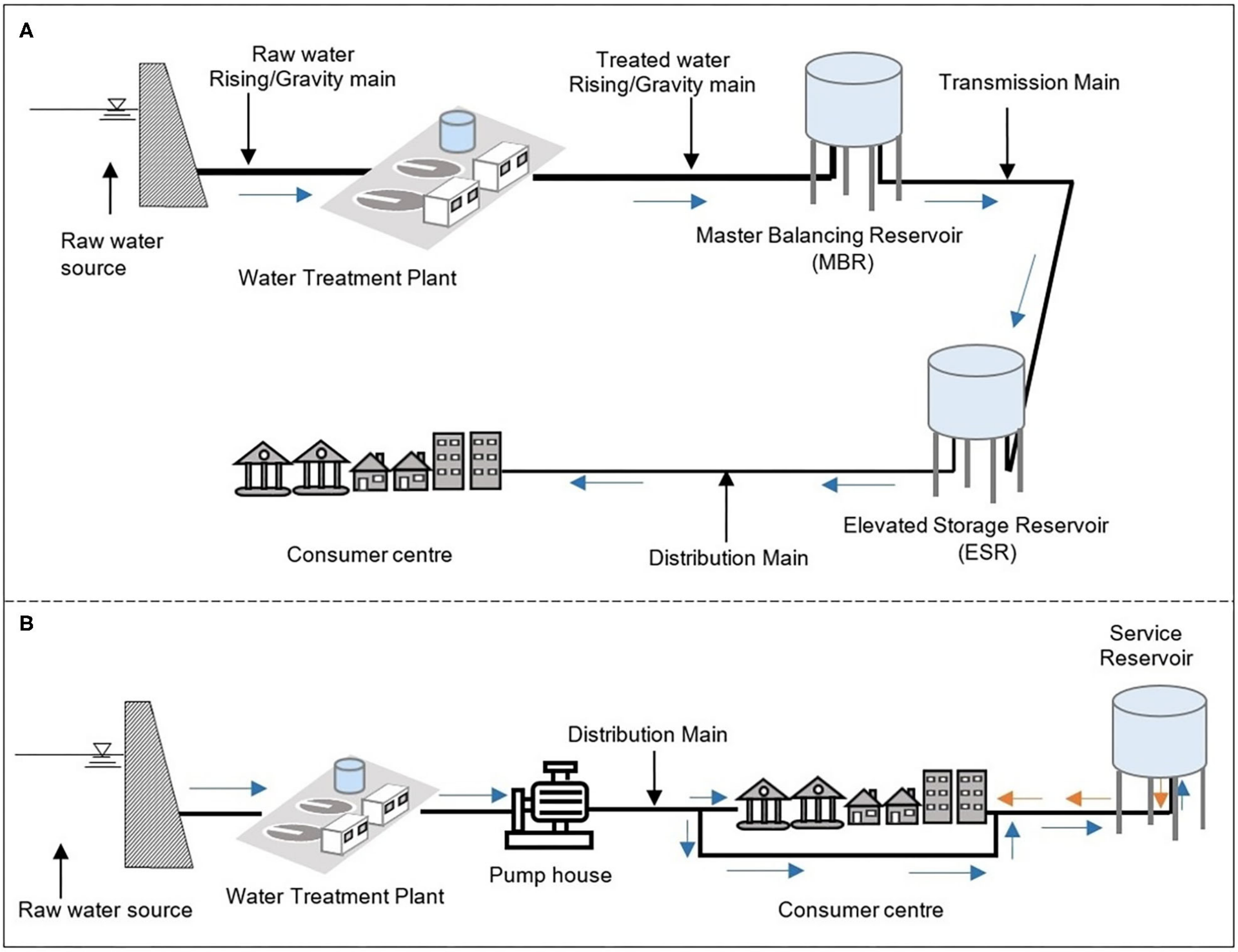


FIGURE 2 | (A) Typical layout of WSS with a surface water source in India, **(B)** Recommended WSS layout by AWWA manual M31.

understand the relationship between the various aspects of WSS in India causing the existence of current IWS. As shown in **Figure 3**, a vicious cycle is formed between the operation and maintenance, infrastructure condition, service delivery, consumer satisfaction, and revenue aspect of a WSS. Lack of proper operation and maintenance deteriorates the infrastructure condition, which reduces the service quality in terms of quantity and quality. The decreased service quality negatively impacts consumer satisfaction; reduced consumer satisfaction results drop in revenue generation. In turn, the depleted revenues affect the operation and maintenance of the WSS, which again initiates the vicious cycle of failure.

Various factors are responsible for the formation of the vicious cycle (**Figure 3**). The operation and maintenance of a WSS is dependent on the aspects like asset management, DMA, network expansion, and so on. Failure of one of the factors results in the WSS's degraded operation and forms the vicious cycle. Similarly, the infrastructure condition is governed by the planning, design, and analysis practices along with the operation and maintenance of the scheme. The infrastructure condition and reliability decide the service quality of WSS and ultimately dictates consumer satisfaction.

TABLE 5 | Design norms for WSS in India and the USA.

Sr.no.	Parameter	CPHEEO manual	AWWA manual
1.	Minimum residual pressure (m)	7	14
2.	Maximum residual pressure (m)	17	69
3.	Minimum diameter (mm)	150/100	150/100
4.	Peak factor	3	2.5
5.	Minimum fire flow (lps)	Neglected	15
6.	Per capita demand (lpcd)	135,70,40	560

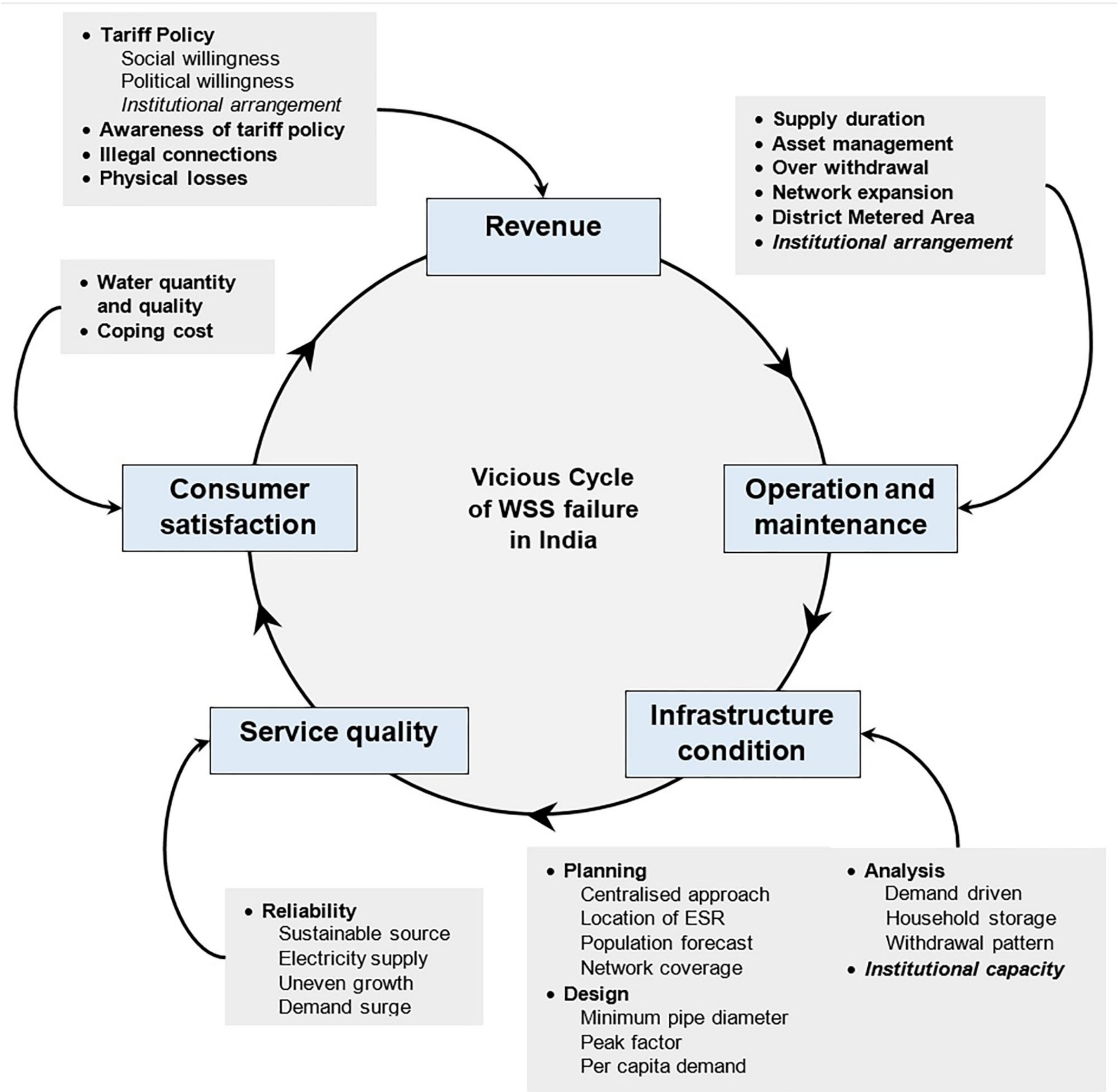


FIGURE 3 | Vicious cycle of WSS failure in India.

Additionally, the coping costs govern consumer satisfaction. The revenue generation aspect relies on the tariff policy, which is influenced by social and political willingness. Furthermore, illegal connections, physical losses, and consumer satisfaction impact the revenue generation of a WSS. Finally, the shared point of all the factors is the institutional capacity of the water supply sector in India. The institutional capacity affects significant aspects like planning, design, analysis, operation, and tariff policy of a WSS. The vicious cycle (**Figure 3**) can be converted to a virtuous cycle if all the factors are maintained as per standards. However, understanding the different drivers of the formation of a vicious cycle is vital. The following section will discuss the drivers responsible for the failure of WSS in India.

DRIVERS FOR VICIOUS CYCLE OF WSS FAILURE IN INDIA

As discussed in section IWS in India, there is a gap between WSS's design and actual operation in India. The design is based on CWS, whereas the operation on IWS. The difference in the design and operation results in the failure of WSS in India. Section Vicious cycle of WSS failure in India illustrated the vicious cycle of WSS failure in India. In this section, the drivers responsible for forming a vicious cycle are discussed based on the field conditions and operational hydraulics.

Design and Analysis

In India, a highly centralized approach is adopted for planning a WSS (section Configuration of WSSs in India). The concept of the ideal location of ESR (center of the serving area) fails at the ground level due to land constraint, resulting in hydraulic failure of the WDN. Additionally, population forecast by standard methods fails to capture the uncertainties, leads to the faulty demand forecast. The design norms of WSS in India are discussed in section WSS design norms in India. The gap between the design and actual operation of WSS makes the design standard unrealistic. For example, there might be a difference between the required diameter at the tail end of the network based on intermittent operation and the minimum pipe diameter criteria as per the standards (**Table 5**).

Furthermore, as discussed in sections WSS design norms in India and Actual operation of WSS in India, WSS's design is based on the peak factor as per CWS. In contrast, the actual peak exerted on the network is higher due to the intermittent operation. Moreover, the per capita demand norm does not account for the changed lifestyle, especially in rural areas. The penetration of new technologies and instruments in rural areas has changed the lifestyle, increasing the per capita water demand.

The next step after the design of WSS is the analysis of the designed network. DDA is predominantly used to simulate the network. In DDA, demand at nodes is assumed to be completely satisfied irrespective of the node's residual pressure (Berardia et al., 2014). However, due to IWS in India, wide variation in the residual pressure is observed, and accordingly, the disparity in the nodal outflow is recorded. To simulate the variation in nodal flow regarding residual pressure, Pressure Driven Analysis (PDA) is preferred and used in recent times (Sayyed et al., 2015; Mahmoud et al., 2017). Even though PDA gives satisfactory results in terms of variation in nodal outflow with respect to residual pressure, it is not easy to estimate the actual withdrawal from the node. The service connection pipe in the Indian WSS is a crucial component to calculate the actual withdrawal from the system. Also, household storage is a significant component of IWS. The analysis methods adopted for WSSs in India ignore the impact of household storage and the service connection pipe. The exclusion of these two vital components fails to capture the ground reality. Hence, it is not easy to simulate the actual scenario of WSSs in India.

The factors mentioned above related to planning, design, and analysis directly impact the infrastructure condition (**Figure 3**). Additionally, the operation and maintenance affect the infrastructure condition; the subsequent section will discuss the drivers for unsuccessful operation and maintenance of WSS in India.

Operation and Maintenance

In India, the intermittent operation of WSS has become standard practice. As discussed in section Operation of the distribution system, the concept of DMA is failed at the actual operation of a scheme. The maintenance activity of the pipelines becomes challenging with the lack of DMA in the network (Diao et al., 2013). In addition to the maintenance activities, conducting a water audit to quantify the actual water loss becomes tedious. As discussed in section Operation of the distribution system, the transmission network's partial flow condition results in incomplete filling of the ESRs. In India, the ESR is bypassed at many places, and the transmission network is directly connected to the distribution network.

Additionally, the refilling of pipelines requires a large volume of water in every cycle of supply time. Water loss in the network is intensified due to the filling and refilling of the pipelines. The Hydraulic Grade Line (HGL) at the storage tanks is one of the governing factors for calculating the residual pressure at nodes. At the design phase, the starting HGL at the storage tank is assumed to be above the Low Supply Level (LSL). In the actual operation, after sometime, the outlet of the storage tank is partially filled due to uncontrolled withdrawal from the tanks. Hence, the HGL sets below the LSL of the tank. The difference between the actual and design HGL value negatively impacts the performance of the distribution network.

For effective operation and maintenance of infrastructure, the fundamental prerequisite is maintaining the infrastructure database. The date of the laying of the pipeline is of vital importance in WSS. In urban areas of India, most of the functional pipelines are very old. In a deteriorated condition, the lack of a proper database regarding the pipe age and rehabilitation activity exacerbates the deterioration of the network. The application of Geographical Information System (GIS) in the database management of WSS is yet to be realized by most of the authorities. The operation practice like short duration supply, lack of DMA, unplanned expansion, uncontrolled withdrawal, and lack of asset management deteriorate WSS's infrastructure condition in India.

Consumer Satisfaction

Consumer satisfaction depends upon the service quality of the WSS. The service quality is decided based on the infrastructure condition and reliability of the WSS. As discussed in the previous sections, the shortcomings in the planning, design, analysis, and operation of WSS in India decline the infrastructure condition. Even after overcoming the issues mentioned above, the improvement in service quality is not assured. The reliability of the WSS is another aspect that governs the service quality. Interrupted electricity supply, uneven population growth, seasonal variation, drought conditions, and unexpected surge in water demand (e.g., large social gatherings) hampers the reliability of the WSS.

In India's rural areas, pumps used in lifting the water from the source often fail within the design period due to interrupted electricity supply. The pump failure disrupts the operation of the scheme, aggravating existing problems of operation and maintenance. Additionally, haphazard growth, especially in urban areas, increases the load on essential utilities like water supply. However, the WSS is not designed to cater to such an increase in demand. Lack of sustainable raw water source to fulfill the increased demand is another vital issue related to WSS's reliability.

As a result of IWS, consumers have to invest in household storage tanks to fulfill the water demand until the next supply cycle. Consumers adopt an underground storage tank, overhead storage tank, large vessels, etc., as per the financial capabilities (Cook et al., 2016). In addition to the storage facility, consumers spend on the pumps to fetch water from the system due to inadequate residual pressures. In India, to counter the health issues caused by IWS, it has become common to install a household water filtration unit (Ercumen et al., 2015). Consumers with unfulfilled water demand like rural areas or slum populations in the urban area adopt alternate sources like groundwater, tanker feed water, bottled water, community water filters and fulfill their water demand. Different alternatives adopted by consumers to accomplish the water demand are known as coping strategies. The investment made by the consumers on different coping strategies is termed coping cost. In addition to the payment of water bills, consumers spend a substantial amount of their earnings on different coping strategies (Yepes et al., 2001; Cook et al., 2016; Gurung et al., 2017). Apart from the coping costs, consumers have to modify the daily schedule as per the supply hours. In most households, one of the members has to stay at home to store water in the respective supply hours, reducing the job prospect for one household member. Hence, to overcome the ill effects of IWS, consumers spend extra time and money.

The quantification of service quality can be done by monitoring the water quantity and quality received by consumers. In addition to it, the quantification of coping cost will lead to determining the service quality, in turn, consumer satisfaction. Consumer satisfaction is directly proportional to the willingness to pay toward the water bill. Further, studies (Burt and Ray, 2014) have shown that willingness to pay is more to reliable water supply than the continuous water supply.

Revenue

Revenue generation plays a vital role in the success of any infrastructure project. Water is considered the basic need for the survival of human beings. Hence, people's general perspective is that the water tariff should be as minimal as possible, or water should be provided free of cost (Aggarwal et al., 2013). However, as a significant investment is required to design, construct, operate, and maintain a WSS, revenue generation is of prime importance for the service providers. In India, the coping costs and the working hours spent on satisfying the water demand dictate the willingness to pay for water bills in a household. As discussed in section Consumer satisfaction, consumer satisfaction is not achieved in most of the WSS in India, resulting in reduced willingness to pay.

From the tariff on water supply point of view, the tariff structure varies among different cities. The tariff policy is finalized by the state government, making it a politically driven decision (Aggarwal et al., 2013). In India, the average tariff rate is below the production cost (McKenzie and Ray, 2009). For example, the Municipal Council of Greater Mumbai (MCGM) spends Rs. 20 (0.27 USD[1]) per thousand liters for treatment of water, whereas only Rs. 5 (0.06 USD) per thousand liters is charged to the consumers (Singh, 2021).

Furthermore, some service providers opt for flat rate tariffs others follow telescopic rates for tariff collection. A flat rate policy is adopted for unmetered connection, block tariff for metered connection. The block's price and bifurcation vary among the cities; McKenzie and Ray (2009) have enlisted the various tariff structure in major Indian cities. In a flat rate policy, the tariff is based on the size of the service connection. In contrast, the block tariff requires household water meters. However, mechanical water meters in IWS fail to quantify the exact water consumption, and after a while, the meters become defective (Walter et al., 2018). Also, the Electromagnetic and Ultrasonic water meters require high operation and maintenance costs (Research Design Standards Organization, 2015; Li and Chong, 2019). Additionally, the insufficient coverage, unequal distribution, low pressure in the WDN escalate the illegal connections, theft of water from the system. Such malpractices increase the burden of revenue loss on the utilities and the lack of exact quantification of water consumption.

Finally, automation is considered the most relevant solution for increasing the efficiency of WSSs. Different automation techniques like Programmable Logic Control (PLC), Supervisory Control and Data Acquisition (SCADA), valve actuators, level sensors, etc., are adopted at a different scale in India. However, the operation and maintenance of such high-end solutions are costly. As discussed above, the willingness to pay is directly proportional to consumer satisfaction. Hence, generating financial returns of the considerable investment in automation becomes an uphill task.

Institutional capacity

In India, as per the 73rd Amendment to the constitution supplying clean drinking water is the state's responsibility (Ministry of Jal Shakti Government of India, 2019). WSS's institutional structure in India is fragmented; different states have a varied institutional framework (Ahmed and Araral, 2019). Most of the states have the Urban Local Body (ULB) as the responsible authority for managing WSS. In contrast, others have a centralized board to manage the water supply (Government of India, 2009). An excellent institutional system has well-trained staff, a financially sustainable tariff structure, an improved revenue collection system, accountability, transparency, and public participation (McIntosh, 2003; Bakker et al., 2008). A weak institutional framework leads to the failure of WSS (Government of India, 2009). In most ULBs or water supply boards, technical and managerial skills are scarce (Klingel, 2012; Mathur, 2017). Instead of suitable capacity building at the institute level, the planning, design, and construction of WSS are outsourced to private vendors. Only the operation and maintenance of the scheme are done by ULBs or water supply boards.

---

[1] Exchange rate of currency 1 INR=0.013 USD 16th March 2021.

The central government of India allocates funds to states for the construction of WSS. At the state level, money from various national level schemes is used to increase the number of tanks and pipelines; there is no cognizance of WSS expenditure. Separate funds are also allocated for capacity building, used for various training programs for the engineer. However, the actual issues and appropriate solutions are rarely discussed in the training programs.

Furthermore, the dynamics between two different institutions also impact the operation of WSS. In Maharashtra, India, the institutional arrangement for water supply is complex (SI-2.4). Maharashtra Jeevan Pradhikaran (MJP), a parastatal body, is responsible for planning, designing, and constructing WSS. Post construction, MJP handover the scheme to respective local bodies. ULBs are accountable for the scheme's operation and maintenance in urban areas; Zilha Parishad and Gram Panchayat in the rural area are responsible for the scheme's operation and maintenance. The relation between different institutions defines the fate of the WSS. Moreover, the dynamics of a civil servant and the engineers at various institutes also impact the water distribution (Bawa, 2011).

In addition to it, the water tariff structure fixed by the authority is unrealistic, institutions fail to recover the operating cost of the scheme (Golin et al., 2015). The application of GIS for maintaining the revenue collection database is yet to be incorporated by all the institutions. As discussed in the earlier section, lack of proper database management escalates the problem of cost recovery. Furthermore, the actual issues and appropriate solutions are rarely discussed in the training programs used for capacity building. In the wake of such a scenario, the consultant's unrealistic design gets approved, and the scheme enters the vicious cycle of failure (**Figure 3**).

To summarize, a robust institutional capacity will improve all aspects of WSS, including planning, design, analysis, operation, and tariff policy. The recently launched Jal Jeevan Mission (JJM) by the Government of India has focused on strengthening the institutional capacity with a four tier institutional setup at the National, State, District, and Village level. Also, strong database management and community involvement are made mandatory for states (Ministry of Jal Shakti Government of India, 2019). However, the provision for appointment of the Project Management Unit (PMU) defies the purpose. In PMU, the institutions are free to appoint consultants for the design and operation of WSS. As the focus will be on outsourcing, vendors will provide high-end solutions which are generally suitable for developed nations. The issue of the lack of appropriate interventions to improve WSS in India will remain unresolved.

LEARNINGS AND WAY FORWARD

In the present study, a detailed analysis of drivers for IWS in India is presented. It is clear that the drivers for IWS are not independent and reinforce each other. One of the crucial aspects of IWS is how to operate the WSS considering the field conditions. As water is supplied intermittently, household storage is a common phenomenon in IWS. Additionally, over withdrawal from the network is also predominant in IWS. However, such scenarios are hardly considered during the design and planning phase, resulting in the WSS's failure. Bhave and Gupta (2000) proposed hydraulic devices such as orifice plate to restrict the over withdrawal from the WSS.

The ultimate goal of a WSS is to provide water to consumers in adequate quantity and prescribed quality in CWS mode. However, with the present scenario of operating the WSS, directly converting the IWS to CWS will result in more problems, as such approach will not tackle the drivers to IWS. This will not effectively use the investment made for achieving CWS as the WSS will again move back to IWS with such an approach of direct transition. There should be an intermediate step of improving the performance of existing IWS such that the vicious cycle of failure (**Figure 3**) is broken. To improve the existing IWS, the application of appropriate design and operation interventions is required. As discussed in the present study, limited studies proposed the alternative design and operation methodology for IWS. Kalbar and Gokhale (2019) described few design and operation interventions, viz. shaft, manifold, and multi-outlet tanks, for improving the efficiency of existing IWS. Furthermore, Ghorpade et al. (2021) reported multi-outlet tanks' superiority over conventional single outlet tanks in terms of hydraulic and operational aspects.

Once the performance of IWS is improved and consumer satisfaction is achieved, a gradual transition to CWS is possible. Such a transition will result in sustainable CWS. To summarize, the approach of directly converting IWS to CWS will lead to reinforcing the vicious cycle of failure. To overcome such situations, an intermediate step is necessary to bridge the gap between IWS to CWS.

CONCLUSIONS

IWS is the most common operational regime of WSSs in India. In this study, the drivers for the IWS in India were identified based on studies on field conditions and operational hydraulics. Initially, an SLR was conducted to understand the various aspects of IWS and the variety of work done to improve IWS. This was followed by a detailed study of WSS in India, including the layout, planning, design, analysis, and actual operation. It was found that the WSSs of India are trapped in the vicious cycle, wherein failure of one factor initiates the decline in the performance of another factor, and ultimately the system fails to deliver a satisfactory water supply to the consumers. Application of solutions for improving the performance of WSSs in India without understanding the vicious cycle and drivers of IWS will lead to the continuation of the vicious cycle of failure. Also, direct conversion of WSS operating in IWS mode to CWS is not recommended as it will not address the drivers for IWS and WSS will again convert back to IWS. The present study will help understand the interrelation between various drivers responsible for the vicious cycle of IWS. Once a thorough understanding of the various drivers is developed in the practitioners, they can develop strategies to overcome the vicious cycle of IWS. The success of national level missions like JJM (Rural and Urban) and any future programmes depend upon overcoming the vicious cycle of IWS.

AUTHOR CONTRIBUTIONS

AG with significant contributions from PK designed the research idea and the basis of the content for the manuscript. AG wrote the first draft of the manuscript. AS provided inputs about the comparison of WSS between developed and developing nations. PK, AG, and AS revised and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The funding received under the sponsored research project from the Department of Science and Technology (DST), India—Water Technology Initiative (WTI) −2017 DST/TM/WTI/2K17/39.

ACKNOWLEDGMENTS

The first author acknowledges the scholarship received under the Teaching Assistant Category at Indian Institute of Technology Bombay by the Ministry of Human Resource and Development (MHRD), Government of India. Technical discussions with Mr. Pradeep Gokhale, Retired Superintendent Engineer, Maharashtra Jeevan Pradhikaran, Maharashtra, India, helped to bring clarity to this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frwa.2021.696630/full#supplementary-material

REFERENCES

Aboelnga, H., Saidan, M., Al-Weshah, R., Sturm, M., Ribbe, L., and Frechen, F. B. (2018). Component analysis for optimal leakage management in Madaba, Jordan. *J. Water Suppl. Res. Technol*. 67, 384–396. doi: 10.2166/aqua.2018.180

Abu-Madi, M., and Trifunovic, N. (2013). Impacts of supply duration on the design and performance of intermittent water distribution systems in the West Bank. *Water Int.* 38, 263–282. doi: 10.1080/02508060.2013.794404

Agathokleous, A., Christodoulou, C., and Christodoulou, S. E. (2017). Influence of intermittent water supply operations on the vulnerability of water distribution networks. *J. Hydroinform.* 19, 838–852. doi: 10.2166/hydro.2017.133

Agathokleous, A., and Christodoulou, S. (2016). Vulnerability of urban water distribution networks under intermittent water supply operations. *Water Resourc. Manage.* 30, 4731–4750. doi: 10.1007/s11269-016-1450-3

Aggarwal, V., Maurya, N., and Jain, G. (2013). Pricing urban water supply. *Environ. Urbaniz. Asia* 4, 221–241. doi: 10.1177/0975425313477768

Ahluwalia, I. J., Munjee, N., Mor, N., Vijayanunni, M., Mankad, S., Lall, R., and Sankaran, H. (2011). *Report on Indian urban Infrastructure and Services*. Ministry of Urban Development, New Delhi, India.

Ahmed, M., and Araral, E. (2019). Water Governance in India: evidence on Water Law, policy, and administration from eight Indian states. *Water* 11:2071. doi: 10.3390/w11102071

Alazzeh, S., Galaitsi, S. E., Bishara, A., Al-Azraq, N., and Durant, J. L. (2019). Impacts of intermittent water supply on water quality in two Palestinian refugee camps. *Water* 11:670. doi: 10.3390/w11040670

Al-Ghamdi, A. S. (2011). Leakage–pressure relationship and leakage detection in intermittent water distribution systems. *J. Water Supply Res. Technol.* 60, 178–183. doi: 10.2166/aqua.2011.003

Al-Ghamdi, A. S., and Gutub, S. A. (2002). Estimation of leakage in the water distribution network of the Holy City of Makkah. *J. Water Supply Res. Technol.* 51, 343–349. doi: 10.2166/aqua.2002.0031

Al-Washali, T., Sharma, S., Al-Nozaily, F., Haidera, M., and Kennedy, M. (2019). Monitoring nonrevenue water performance in intermittent supply. *Water.* 11:1220. doi: 10.3390/w11061220

Al-Washali, T., Sharma, S., Lupoja, R., Fadhl, A. N., Haidera, M., and Kennedy, M. (2020). Assessment of water losses in distribution networks: Methods, applications, uncertainties, and implications in intermittent supply. Resources, *Conserv. Recycl.* 152:104515. doi: 10.1016/j.resconrec.2019.104515

Ameyaw, E. E., Memon, F. A., and Bicik, J. (2013). Improving equity in intermittent water supply systems. *J. Water Supply Res. Technol.* 62, 552–562. doi: 10.2166/aqua.2013.065

Andey, S. P., and Kelkar, P. S. (2007). Performance of water distribution systems during intermittent versus continuous water supply. *J. Am. Water Works Assoc.* 99, 99–106. doi: 10.1002/j.1551-8833.2007.tb08011.x

Andey, S. P., and Kelkar, P. S. (2009). Influence of intermittent and continuous modes of water supply on domestic water consumption. *Water Resourc. Manage.* 23, 2555–2566. doi: 10.1007/s,11269-008-9396-8

AWWA American Water Works Association (2008). *Distribution System Requirements for Fire Protection, Vol. 31*. Denver, CO: American Water Works Association.

Ayoub, G. M., and Malaeb, L. (2006). Impact of intermittent water supply on water quality in Lebanon. *Int. J. Environ. Pollut.*, 26(4), 379–397. doi: 10.1504/IJEP.2006.009328

Bakker, K., Kooy, M., Shofiani, N. E., and Martijn, E. J. (2008). Governance failure: rethinking the institutional dimensions of urban water supply to poor households. *World Develop.* 36, 1891–1915. doi: 10.1016/j.worlddev.2007.09.015

Bautista-de los Santos, Q. M., Chavarria, K. A., and Nelson, K. L. (2019). Understanding the impacts of intermittent supply on the drinking water microbiome. *Curr. Opin. Biotechnol.* 57, 167–174. doi: 10.1016/j.copbio.2019.04.003

Bawa, Z. (2011). Where is the state? how is the state? Accessing water and the state in Mumbai and Johannesburg. *J. Asian Afr. Stud.* 46, 491–503. doi: 10.1177/0021909611403707

Berardia, L., Laucellia, D., and Giustolisia, O. (2014). Accounting for local water storages in assessing WDN supply capacity. *Proc. Engi.* 70, 142–151. doi: 10.1016/j.proeng.2014.02.017

Bhave, P. R., and Gupta, R. (2000). Design, performance and operation of regional rural water supply systems. *J. Indian Water Works Assoc.* 32, 273–278.

Bivins, A. W., Sumner, T., Kumpel, E., Howard, G., Cumming, O., Ross, I., and Brown, J. (2017). Estimating infection risks and the global burden of diarrheal disease attributable to intermittent water supply using QMRA. *Environ. Sci. Technol.* 51, 7542–7551. doi: 10.1021/acs.est.7b01014

Burt, Z., Ercümen, A., Billava, N., and Ray, I. (2018). From intermittent to continuous service: Costs, benefits, equity and sustainability of water system reforms in Hubli-Dharwad, India. *World Develop.* 109, 121–133. doi: 10.1016/j.worlddev.2018.04.011

Burt, Z., and Ray, I. (2014). Storage and non-payment: persistent informalities within the formal water supply of Hubli-Dharwad, India. *Water Alternat.* 7:183.

Campisano, A., Gullotta, A., and Modica, C. (2018). Using EPA-SWMM to simulate intermittent water distribution systems. *Urban Water J.* 15, 925–933. doi: 10.1080/1573062X.2019.1597379

Charalambous, B., and Laspidou, C. (2017). *Dealing with the Complex Interrelation of Intermittent Supply and Water Losses*. London: IWA Publishing.

Christodoulou, S., and Agathokleous, A. (2012). A study on the effects of intermittent water supply on the vulnerability of urban water distribution networks. *Water Sci. Technol. Water Supply* 12, 523–530. doi: 10.2166/ws.2012.025

Coelho, S. T., James, S., Sunna, N., Abu Jaish, A., and Chatila, J. (2003). Controlling water quality in intermittent supply systems. *Water Sci. Technol. Water Supply* 3, 119–125. doi: 10.2166/ws.2003.0094

Cook, J., Kimuyu, P., and Whittington, D. (2016). The costs of coping with poor water supply in rural Kenya. *Water Resourc. Res.* 52, 841–859. doi: 10.1002/2015WR017468

CPHEEO (1999). *Manual on Water supply and Treatment. 3rd Edn*. New Delhi: Ministry of Urban Development, Government of India.

Criminisi, A., Fontanazza, C. M., Freni, G., and Loggia, G. L. (2009). Evaluation of the apparent losses caused by water meter under-registration in intermittent water supply. *Water Sci. Technol.* 60, 2373–2382. doi: 10.2166/wst.2009.423

Cronk, R., and Bartram, J. (2018). Identifying opportunities to improve piped water continuity and water system monitoring in Honduras, Nicaragua, and Panama: evidence from Bayesian networks and regression analysis. *J. Clean. Product.* 196, 1–10. doi: 10.1016/j.jclepro.2018.06.017

David, C., García, M., and Navarro Gómez, C. J. (2020). Analysis of hybrid demand pattern on a water distribution network with transition from intermittent to continuous water supply in Riberas de Sacramento, Chihuahua. *Water Pract. Technol.* 15, 213–224. doi: 10.2166/wpt.2020.013

De Marchis, M., Fontanazza, C. M., Freni, G., La Loggia, G., Napoli, E., and Notaro, V. (2010). A model of the filling process of an intermittent distribution network. *Urban Water J.* 7, 321–333. doi: 10.1080/1573062X.2010.519776

De Marchis, M., Fontanazza, C. M., Freni, G., La Loggia, G., Notaro, V., and Puleo, V. (2013). A mathematical model to evaluate apparent losses due to meter under-registration in intermittent water distribution networks. *Water Sci. Technol. Water Supply* 13, 914–923. doi: 10.2166/ws.2013.076

Diao, K., Zhou, Y., and Rauch, W. (2013). Automated creation of district metered area boundaries in water distribution systems. *J. Water Resourc. Plann. Manage.* 139, 184–190. doi: 10.1061/(ASCE)WR.1943-5452.0000247

El Achi, N., and Rouse, M. J. (2020). A hybrid hydraulic model for gradual transition from intermittent to continuous water supply in Amman, Jordan: a theoretical study. *Water Supply* 20, 118–129. doi: 10.2166/ws.2019.142

Elala, D., Labhasetwar, P., and Tyrrel, S. F. (2011). Deterioration in water quality from supply chain to household and appropriate storage in the context of intermittent water supplies. *Water Sci. Technol. Water Supply* 11, 400–408. doi: 10.2166/ws.2011.064

Ercumen, A., Arnold, B. F., Kumpel, E., Burt, Z., Ray, I., Nelson, K., and Colford Jr, J. M. (2015). Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: a matched cohort study in urban India. *PLoS Med* 12:*e*1001892. doi: 10.1371/journal.pmed.1001892

Erickson, J. J., Smith, C. D., Goodridge, A., and Nelson, K. L. (2017). Water quality effects of intermittent water supply in Arraiján, Panama. *Water Res.* 114, 338–350. doi: 10.1016/j.watres.2017.02.009

Fan, L., Liu, G., Wang, F., Ritsema, C. J., and Geissen, V. (2014). Domestic water consumption under intermittent and continuous modes of water supply. *Water Resourc. Manage.* 28, 853–865. doi: 10.1007/s11269-014-0520-7

Fontanazza, C. M., Freni, G., and La Loggia, G. (2007). Analysis of intermittent supply systems in water scarcity conditions and evaluation of the resource distribution equity indices. *WIT Trans. Ecol. Environ.* 103:70591. doi: 10.2495/WRM070591

Fontanazza, C. M., Freni, G., Loggia, G. L., Notaro, V., and Puleo, V. (2013). Evaluation of the water scarcity energy cost for users. *Energies*, 6, 220–234. doi: 10.3390/en6010220

Freni, G., De Marchis, M., and Napoli, E. (2014). Implementation of pressure reduction valves in a dynamic water distribution numerical model to control the inequality in water supply. *J. Hydroinform.* 16, 207–217. doi: 10.2166/hydro.2013.032

Galaitsi, S. E., Russell, R., Bishara, A., Durant, J. L., Bogle, J., and Huber-Lee, A. (2016). Intermittent domestic water supply: a critical review and analysis of causal-consequential pathways. *Water* 8:274. doi: 10.3390/w8070274

Ghorpade, A., Sinha, A. K., and Kalbar, P. (2021). Multi-outlet storage tanks to improve water distribution networks in India. *Urban Water J.* 19, 1-9. doi: 10.1080/1573062X.2021.1914117

Golin, C., Cox, M., Brown, M., and Thomas, V. (2015). The water efficiency gap. *Sustain. Water Resourc. Manage.* 1, 315–324. doi: 10.1007/s40899-015-0025-4

Gottipati, P. V., and Nanduri, U. V. (2014). Equity in water supply in intermittent water distribution networks. *Water Environ. J.* 28, 509–515. doi: 10.1111/wej.12065

Government of India (2009). *Guidance Notes for Continuous Water Supply (24-7 Supply). A Guide to Project Preparation, Implementation and Appraisal*. New Delhi: Ministry of Urban Development.

Guragai, B., Hashimoto, T., Oguma, K., and Takizawa, S. (2018). Data logger-based measurement of household water consumption and micro-component analysis of an intermittent water supply system. *J. Clean. Product.* 197, 1159–1168. doi: 10.1016/j.jclepro.2018.06.198

Gurung, Y., Zhao, J., Kumar, K. C., B., Wu, X., Suwal, B., and Whittington, D. (2017). The costs of delay in infrastructure investments: a comparison of 2001 and 2014 household water supply coping costs in the Kathmandu Valley, Nepal. *Water Resourc. Res.* 53, 7078–7102. doi: 10.1002/2016WR019529

Haddad, M., Mcneil, L., and Omar, N. (2014). Model for predicting disinfection by-product (DBP) formation and occurrence in intermittent water supply systems: Palestine as a case study. *Arab. J. Sci. Eng.* 39, 5883–5893. doi: 10.1007/s13369-014-1200-x

Haddad, O., Hoseini-Ghafari, S., Solgi, M., and Loáiciga, H. A. (2016). Intermittent urban water supply with protection of consumers' welfare. *J. Pipeline Syst. Eng. Pract.* 7:04016002. doi: 10.1061/(ASCE)PS.1949-1204.0000231

Haider, H., Al-Salamah, I. S., Ghazaw, Y. M., Abdel-Maguid, R. H., Shafiquzzaman, M., and Ghumman, A. R. (2019). Framework to establish economic level of leakage for intermittent water supplies in arid environments. *J. Water Resourc. Plann. Manage.* 145:05018018. doi: 10.1061/(ASCE)WR.1943-5452.0001027

Hastak, S., Labhasetwar, P., Kundley, P., and Gupta, R. (2017). Changing from intermittent to continuous water supply and its influence on service level benchmarks: a case study in the demonstration zone of Nagpur, India. *Urban Water J.* 14, 768–772. doi: 10.1080/1573062X.2016.1240808

Hernandez-Lopez, R. D., Tzatchkov, V. G., Martin-Dominguez, A., and Alcocer-Yamanaka, V. H. (2016). Study of hydraulics and mixing in roof tanks used in intermittent water supply. *J. Water, Sanit. Hygiene Develop.* 6, 547–554. doi: 10.2166/washdev.2016.147

Ilaya-Ayza, A. E., Campbell, E., Pérez-García, R., and Izquierdo, J. (2016). Network capacity assessment and increase in systems with intermittent water supply. *Water* 8:126. doi: 10.3390/w8040126

Ilaya-Ayza, A. E., Martins, C., Campbell, E., and Izquierdo, J. (2017). Implementation of DMAs in intermittent water supply networks based on equity criteria. *Water* 9:851. doi: 10.3390/w9110851

Ilaya-Ayza, A. E., Martins, C., Campbell, E., and Izquierdo, J. (2018). Gradual transition from intermittent to continuous water supply based on multi-criteria optimization for network sector selection. *J. Comput. Appl. Mathemat.* 330, 1016–1029. doi: 10.1016/j.cam.2017.04.025

Jayaramu, K. P., Burt, Z., and Manoj Kumar, B. (2015). A study of the consumption pattern in a continuous water service demonstration zone and bulk water demand forecasting for Hubli-Dharwad, India. *J. Water Sanit. Hygiene Develop.* 5, 201–212. doi: 10.2166/washdev.2015.096

Jayaramu, K. P., Manoj Kumar, B., and Prasanna Rashmi, K. K. (2016). Improving cost recovery in urban water supply service: an experience from India. *Water Policy* 18, 685–707. doi: 10.2166/wp.2015.134

Kalbar, P., and Gokhale, P. (2019). Decentralized infrastructure approach for successful water supply systems in India: use of multi-outlet tanks, shafts and manifolds. *J. Water Supply Res. Technol. Aqua* 68, 295–301. doi: 10.2166/aqua.2019.158

Kaminsky, J., and Kumpel, E. (2018). Dry pipes: Associations between utility performance and intermittent piped water supply in low and middle income countries. *Water* 10:1032. doi: 10.3390/w10081032

Khadse, G. K., Patni, P. M., Talkhande, A. V., and Labhasetwar, P. K. (2016). Surveillance of the chemical and microbial quality of drinking water for safe water supply in an urban area. *J. Water Supply Res. Technol.* 65, 220–233. doi: 10.2166/aqua.2016.062

Klingel, P. (2012). Technical causes and impacts of intermittent water distribution. *Water Sci. Technol. Water Supply* 12, 504–512. doi: 10.2166/ws.2012.023

Klingel, P., and Nestmann, F. (2014). From intermittent to continuous water distribution: a proposed conceptual approach and a case study of Béni Abbès (Algeria). *Urban Water J.* 11, 240–251. doi: 10.1080/1573062X.2013.765493

Kumar, T., Post, A. E., and Ray, I. (2018). Flows, leaks and blockages in informational interventions: a field experimental study of Bangalore's water sector. *World Develop.* 106, 149–160. doi: 10.1016/j.worlddev.2018.01.022

Kumpel, E., and Nelson, K. L. (2013). Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Res.* 47, 5176–5188. doi: 10.1016/j.watres.2013.05.058

Kumpel, E., and Nelson, K. L. (2014). Mechanisms affecting water quality in an intermittent piped water supply. *Environ. Sci. Technol.* 48, 2766–2775. doi: 10.1021/es405054u

Kumpel, E., and Nelson, K. L. (2016). Intermittent water supply: prevalence, practice, and microbial water quality. *Environ. Sci. Technol.* 50, 542–553. doi: 10.1021/acs.est.5b03973

Kumpel, E., Woelfle-Erskine, C., Ray, I., and Nelson, K. L. (2017). Measuring household consumption and waste in unmetered, intermittent piped water systems. *Water Resourc. Res.* 53, 302–315. doi: 10.1002/2016WR019702

Li, H., Cohen, A., Li, Z., Lv, S., He, Z., Wang, L., and Zhang, X. (2020). Intermittent water supply management, household adaptation, and drinking water quality: a comparative study in Two Chinese Provinces. *Water* 12:1361. doi: 10.3390/w12051361

Li, X. J., and Chong, P. H. J. (2019). Design and implementation of a self-powered smart water meter. *Sensors* 19:4177. doi: 10.3390/s19194177

Lieb, A. M., Rycroft, C. H., and Wilkening, J. (2016). Optimizing intermittent water supply in urban pipe distribution networks. *SIAM J. Appl. Math.* 76, 1492–1514. doi: 10.1137/15M1038979

Mahmoud, H. A., Savi,ć, D., and Kapelan, Z. (2017). New pressure-driven approach for modeling water distribution networks. *J. Water Resourc. Plann. Manage.* 143:04017031. doi: 10.1061/(ASCE)WR.1943-5452.0000781

Manohar, U., and Mohan Kumar, M. S. (2014). Modeling equitable distribution of water: dynamic inversion-based controller approach. *J. Water Resourc. Plann. Manage.* 140, 607–619. doi: 10.1061/(ASCE)WR.1943-5452.0000368

Mastaller, M., and Klingel, P. (2017). Adapting the IWA water balance to intermittent water supply and flat-rate tariffs without customer metering. *J. Water Sanitat. Hyg. Dev*. 7, 396–406. doi: 10.2166/washdev.2017.116

Mathur, S. (2017). Public-private partnership for municipal water supply in developing countries: Lessons from Karnataka, India, urban water supply improvement project. *Cities* 68, 56–62. doi: 10.1016/j.cities.2017.05.007

McIntosh, A. (2003). *Asian Water Supplies: Reaching the Urban Poor*. Mandaluyong: Asian Development Bank.

McKenzie, D., and Ray, I. (2009). Urban water supply in India: status, reform options and possible lessons. *Water Policy* 11, 442–460. doi: 10.2166/wp.2009.056

Mellor, J., Kumpel, E., Ercumen, A., and Zimmerman, J. (2016). Systems approach to climate, water, and diarrhea in Hubli-Dharwad, India. *Environ. Sci. Technol.* 50, 13042–13051. doi: 10.1021/acs.est.6b02092

Ministry of Jal Shakti and Government of India, New Delhi (2019). "*Operational guidelines for the implementation of Jal Jeevan Mission* (Har Ghar Jal)."

Mohan, S., and Abhijith, G. R. (2020). Hydraulic analysis of intermittent water-distribution networks considering partial-flow regimes. *J. Water Resourc. Plann. Manage.* 146:04020071. doi: 10.1061/(ASCE)WR.1943-5452.0001246

Mohapatra, S., Sargaonkar, A., and Labhasetwar, P. K. (2014). Distribution network assessment using EPANET for intermittent and continuous water supply. *Water Resourc. Manage.* 28, 3745–3759. doi: 10.1007/s11269-014-0707-y

Mokssit, A., De Gouvello, B., Chazerain, A., Figuères, F., and Tassin, B. (2018). Building a methodology for assessing service quality under intermittent domestic water supply. *Water* 10:1164. doi: 10.3390/w10091164

MoUD (2012). *Service levels in Urban Water and Sanitation Sector. 1st edn*. New Delhi: Ministry of Urban Development, Government of India.

Neelakantan, T. R., Rammurthy, D., Smith, S. T., and Suribabu, C. R. (2014). Expansion and upgradation of intermittent water supply system. *Asian J. Appl. Sci.* 7, 470–485. doi: 10.3923/ajaps.2014.470.485

Nyende-Byakika, S., Ndambuki, J. M., and Ngirane-Katashaya, G. (2013). Modelling of pressurised water supply networks that may exhibit transient low pressure–open channel flow conditions. *Water Pract. Technol.* 8, 503–514. doi: 10.2166/wpt.2013.054

Nyende-Byakika, S., Ngirane-Katashaya, G., and Ndambuki, J. M. (2012). Modeling flow regime transition in intermittent water supply networks using the interface tracking method. *Int. J. Physic. Sci.* 7, 327–337. doi: 10.5897/IJPS11.873

Owens, C. E., Angles, M. L., Cox, P. T., Byleveld, P. M., Osborne, N. J., and Rahman, M. B. (2020). Implementation of quantitative microbial risk assessment (QMRA) for public drinking water supplies: Systematic review. *Water Res.* 174:115614. doi: 10.1016/j.watres.2020.115614

Rawas, F., Bain, R., and Kumpel, E. (2020). Comparing utility-reported hours of piped water supply to households' experiences. npj Clean *Water* 3, 1–9. doi: 10.1038/s41545-020-0053-y

Research Design and Standards Organization (2015). *Ministry of Railways, Government of India*. Guidelines on water meters.

Rosenberg, D. E., Talozi, S., and Lund, J. R. (2008). Intermittent water supplies: challenges and opportunities for residential water users in Jordan. *Water Int.* 33, 488–504. doi: 10.1080/02508060802474574

Rubino, F., Corona, Y., Jiménez Pérez, J. G., and Smith, C. (2019). Bacterial contamination of drinking water in Guadalajara, Mexico. *Int. J. Environ. Res. Public Health* 16:67. doi: 10.3390/ijerph16010067

Sakomoto, T., Lutaaya, M., and Abraham, E. (2020). Managing water quality in intermittent supply systems: the case of Mukono Town, Uganda. *Water* 12:806. doi: 10.3390/w12030806

Sayyed, M. A. H. A., Gupta, R., and Tanyimboh, T. T. (2015). Noniterative application of EPANET for pressure dependent modelling of water distribution systems. *Water Resourc. Manage.* 29, 3227–3242. doi: 10.1007/s11269-015-0992-0

Simukonda, K., Farmani, R., and Butler, D. (2018). Intermittent water supply systems: causal factors, problems and solution options. *Urban Water J.* 15, 488–500. doi: 10.1080/1573062X.2018.1483522

Singh, L. (2021). *BMC Needs Sustainable Revenue Sources like Independent Fee Revision Authority as Long Term Plan*. New Delhi: The Indian Express. Available online at: https://indianexpress.com/article/cities/mumbai/bmc-needs-sustainable-revenue-sources-like-independent-fee-revision-authority-as-long-term-plan-7228292/lite/

Smith, K., and Liu, S. (2020). Methodology for evaluating city-level energy footprint for water distribution systems. *J. Clean. Product.* 2020:125463. doi: 10.1016/j.jclepro.2020.125463

Solgi, M., Bozorg Haddad, O., Seifollahi-Aghmiuni, S., and Loáiciga, H. A. (2015). Intermittent operation of water distribution networks considering equanimity and justice principles. *J. Pipeline Syst. Eng. Pract.* 6:04015004. doi: 10.1061/(ASCE)PS.1949-1204.0000198

Strijdom, L., Speight, V., and Jacobs, H. E. (2017). An assessment of sub-standard water pressure in South African potable distribution systems. *J. Water Sanit. Hygiene Develop.* 7, 557–567. doi: 10.2166/washdev.2017.227

Tamari, S., and Ploquet, J. (2012). Determination of leakage inside buildings with a roof tank. *Urban Water J.* 9, 287–303. doi: 10.1080/1573062X.2012.660959

Taylor, D. D., Slocum, A. H., and Whittle, A. J. (2018). Analytical scaling relations to evaluate leakage and intrusion in intermittent water supply systems. *PloS ONE* 13:e0196887. doi: 10.1371/journal.pone.0196887

Taylor, D. D., Slocum, A. H., and Whittle, A. J. (2019). Demand satisfaction as a framework for understanding intermittent water supply systems. *Water Resour*. Res. 55, 5217–5237. doi: 10.1029/2018WR024124

Tokajian, S., and Hashwa, F. (2003). Water quality problems associated with intermittent water supply. *Water Sci. Technol.* 47, 229–234. doi: 10.2166/wst.2003.0200

Vairavamoorthy, K., and Elango, K. (2002). Guidelines for the design and control of intermittent water distribution systems. *Waterlines* 21, 19–21. doi: 10.3362/0262-8104.2002.041

Vairavamoorthy, K., Gorantiwar, S. D., and Mohan, S. (2007a). Intermittent water supply under water scarcity situations. *Water Int.* 32, 121–132. doi: 10.1080/02508060708691969

Vairavamoorthy, K., Yan, J., Galgale, H. M., and Gorantiwar, S. D. (2007b). IRA-WDS: A GIS-based risk analysis tool for water distribution systems. *Environ. Model. Softw*. 22, 951–965. doi: 10.1016/j.envsoft.2006.05.027

Walski, T. (2014). "How does water distribution design really work?," in *World Environmental and Water Resources Congress 2014,* 375–382.

Walter, D., Mastaller, M., and Klingel, P. (2018). Accuracy of single-jet and multi-jet water meters under the influence of the filling process in intermittently operated pipe networks. *Water Sci. Technol. Water Supply* 18, 679–687. doi: 10.2166/ws.2017.149

Yepes, G., Ringskog, K., and Sarkar, S. (2001). The high costs of intermittent water service. *J. Indian Water Works Assoc.* 33, 99–115.

Zyoud, S. H., and Fuchs-Hanusch, D. (2019). Comparison of several decision-making techniques: a case of water losses management in developing countries. *Int. J. Inform. Technol. Decis. Mak*. 18, 1551–1578. doi: 10.1142/S0219622019500275

Zyoud, S. H., Kaufmann, L. G., Shaheen, H., Samhan, S., and Fuchs-Hanusch, D. (2016). A framework for water loss management in developing countries under fuzzy environment: Integration of Fuzzy AHP with Fuzzy TOPSIS. *Expert Syst. Appl.* 61, 86–105. doi: 10.1016/j.eswa.2016.05.016

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.

frontiers in Water

BRIEF RESEARCH REPORT
published: 13 December 2021
doi: 10.3389/frwa.2021.734361

Enabling Automatic Detection of Anomalies in Wastewater: A Highly Simplified Approach to Defining "Normal" in Complex Chemical Mixtures

Alfred P. Navato[1]* *and Amy V. Mueller*[1,2]

[1] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA, United States, [2] Department of Marine and Environmental Sciences, Northeastern University, Boston, MA, United States

OPEN ACCESS

Edited by:
Abbas Roozbahani,
University of Tehran, Iran

Reviewed by:
Sumit Purohit,
Pacific Northwest National Laboratory (DOE), United States
Ananda Tiwari,
University of Helsinki, Finland

***Correspondence:**
Alfred P. Navato
navato.a@northeastern.edu

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *01 July 2021*
Accepted: *12 November 2021*
Published: *13 December 2021*

Citation:
Navato AP and Mueller AV (2021) Enabling Automatic Detection of Anomalies in Wastewater: A Highly Simplified Approach to Defining "Normal" in Complex Chemical Mixtures. Front. Water 3:734361. doi: 10.3389/frwa.2021.734361

Wastewater treatment demands management of influent conditions to stabilize biological processes. Generally wastewater collection systems lack advance warning of approaching water parcels with anomalous characteristics, which could then be diverted for testing or pre-treatment. A major challenge in achieving this goal is identifying anomalies against the complex chemical background of wastewaters. This work evaluates unsupervised clustering methods to characterize "normal" wastewater characteristics, using $>$17 months of 10-min resolution absorbance spectrometry data collected at an operating wastewater treatment facility. Comparison of results using K-means, GMM, Hierarchical, and DBSCAN clustering shows minimal intra-cluster variability achieved using K-means. The four K-means clusters include three representing 99% of samples, with the remaining cluster ($<$0.3% of samples) representing atypical measurements, demonstrating utility in identifying both underlying modalities of wastewater characteristics and outliers. K-means clustering provides a better separation than grouping based on factors such as month, precipitation, or flow (with 25% overlap at 1-σ level, compared to 93, 93, and 83%, respectively) and enables identification of patterns that are not visible in factor-driven grouping, e.g., shows that summer and November months have a characteristic type of behavior. When evaluated with respect to wastewater influent changes occurring during the SARS-CoV-2 pandemic, the K-means approach shows a distinct change in strength of diurnal patterns when compared to non-pandemic periods during the same season. This method may therefore be useful both as a tool for fast anomaly detection in wastewaters, contributing to improved infrastructure resilience, as well for providing overall analysis of temporal patterns in wastewater characteristics.

Keywords: wastewater, anomaly detection, machine learning, K-means clustering, Gaussian Mixture Model, hierarchical clustering, DBSCAN, SARS-CoV-2

1. INTRODUCTION

Municipal wastewater treatment and water resource recovery facilities leverage biological processes to remove carbon and nutrients, a critical aspect for meeting discharge permit limits that protect downstream ecosystems. While this is more cost effective and produces less hazardous waste then chemical treatment alone, it requires maintenance of the health of the microbial community to ensure proper results. Because re-establishing the process microbiome after a collapse is slow (up to ~weeks), there is interest in monitoring wastewater influent for anomalies that may indicate the presence of a toxic chemical or even a drastic change in influent character, to which the plant operator can respond, for example, by diverting flows to a holding basin while detailed chemical tests are conducted and/or adjustments (such as blending) are done before sending the wastewater to the treatment train. Yet "wastewater" is in reality a complex mixture of hundreds to thousands of different molecules, the concentrations of which can fluctuate based on community behaviors, industrial plant operations, precipitation (which can lead to dilution of wastewaters through stormwater infiltration), etc. (Thomas et al., 1999; Baurès et al., 2007; Tsoumanis et al., 2010; Schilperoort et al., 2012; Loos et al., 2013). Therefore, detecting an ability to properly detect anomalies is predicated on first being able to properly define the "normal" background against which to compare.

Initial approaches have worked toward this goal by detecting sudden changes in sensor measurements, with some robustness offered by using multi-wavelength spectrometer data [e.g., differential spectrum detection (Langergraber et al., 2003, 2006; Daniel et al., 2008), change detection using Autoregressive Integrated Moving Average (ARIMA) modeling (Daniel et al., 2008)]. However, sudden changes can occur due to non-problematic occurrences (e.g., operating hours of industrial or large office facilities). Therefore, a step toward defining "normal" has been to develop deterministic forecasting models based on external factors known to affect the chemical composition of wastewater, e.g.,time of day (Thomas, 2017), land use type (Lourenço et al., 2006; Baurès et al., 2007; Tsoumanis et al., 2010), and precipitation (Vaillant et al., 1999). Major limitations of such models are the complexity (i.e., relationship between factors and wastewater—for example, modeled behavior of citizens and businesses on a typical Monday must be differentiated from a holiday Monday, with numerous exceptional cases to be managed) and the need for data sources to ingest to drive the model.

Advances in pattern detection and recognition have resulted in new techniques for direct anomaly detection being applied to wastewater environments, e.g., anomaly detection through use of a deep autoencoder model on sewer flow data (Russo et al., 2020) and fault detection through application of a deep neural network to data from 12 chemical and operational sensors at an operating wastewater treatment plant (Mamandipoor et al., 2020). Related attempts on small datasets (months or less) have started to approach characterizing the complex background, e.g., clustering on 24-h of wastewater spectral data (Chow et al., 2018) and most recently, forecasting using various statistical and machine learning techniques on a few months of treatment train sensor data (flow, temperature, pH, and NO_3^-/NH_4^+) rather than external factors (as described above), with a neural network outperforming the other techniques in quantitatively predicting future sensor readings (Cicceri et al., 2021).

These recent approaches suggest promise in moving toward defining "normal" as a way to understand the natural variability in wastewater—thereby suggesting ways to assess anomalies by difference, which may provide more robustness in the long run given the challenges of tuning data-trained algorithms to detect rare occurrences. However, it is not yet clear how well existing approaches will generalize, particularly given the strong seasonal effects driving wastewater characteristics and short time frames of datasets used in most studies. Therefore, this paper explores the use of unsupervised clustering methods to define "normal" in a chemically complex system, taking wastewater as the test case given the strong motivation for anomaly detection in these collection and treatment systems, and leveraging a long-term dataset to assess the consistency (or variability) of wastewater over these operationally-relevant time scales. The results are compared to more typical factor-based approaches, such as those described above, for context and then evaluated under the very different operational conditions experienced during 2020 during the SARS-CoV-2 pandemic. The manuscript closes by providing an assessment of how this methodology may contribute to the overarching goal of providing resiliency to wastewater treatment facilities through a reliable method for characterizing conditions in wastewaters.

2. METHODS

While a variety of sensing mechanisms can be used to assess water quality, absorbance spectrometry was selected for this study due to the wide range of chemicals that can be detected, the relevance of detectable analytes to wastewater, e.g., carbon oxygen demand (Gruber, 2006), nitrate and nitrite (Tsoumanis et al., 2010), and total suspended solids (Lepot et al., 2016), and the results of previous studies showing variability in wastewater character to be detectable using this method (Rieger et al., 2006; Lourenço et al., 2012; Carreres-Prieto et al., 2020). The approach described here has three key aspects: (1) multi-year field data collection, (2) use of typical and novel approaches to identifying underlying characteristics of the wastewater mixture, (3) evaluation of these approaches under a unique case study scenario, i.e., reconfiguration of societal behavior during the 2020 SARS-CoV-2 pandemic.

2.1. Data Collection, Pre-processing, and Normalization

An s::can spectro::lyser using a 5 mm path length was deployed at the Upper Blackstone Clean Water facility (UBCW) in the influent to the secondary treatment train (before the biological process but after removal of large solids), recording absorption at 208 wavelengths (220–737.5 nm, in steps of 2.5 nm) at 10-min intervals. UBCW has a capacity of 31 million gallons per day (MGD) and serves approximately 250,000 customers in the Worcester, MA area. Re-zeroing of the device (to Millipore Ultrapure water) and data download were performed approximately every 10 days. The dataset used for this work covers the period from January 7, 2020 to June 7, 2021 (total of 70,060 measurements, i.e., the complete dataset is a matrix of size 208 by 70,060).

Spectral data were pre-processed to compensate for lens fouling (i.e., buildup of material on the pathlength surfaces between cleanings) and to mitigate zeroing effects. Lens fouling is assumed to build up linearly; the offset between average signals within the 24 h after one cleaning and the 24 h immediately preceding the subsequent cleaning is linearly interpolated across the sampling period and subtracted from the sample data. As even small air bubbles can affect the zeroing of the device, essentially introducing an offset to subsequently measured spectra, spectral data are typically shifted by identifying one wavelength that has minimal and/or constant response to enable intercomparison between sampling periods. For this work the absorbance at the longest wavelength (least response in wastewater) was set to zero for all measurements. Finally, because the spectral shape, rather than intensity, was hypothesized to be important in identifying underlying "normal" modes, each spectrum was scaled such that the maximum absorption measurement was set to 1 after the adjustments made for lens fouling and zeroing described above [a normalization step previously shown to reduce impact of flowrate/dilution (Vaillant et al., 2002)]. The post-processed matrix (corrected for fouling and normalized using offset/scaling) of size 208 by 70,060 constituted the input to all algorithms described below.

To provide a comparison to existing methods leveraging known drivers of wastewater variability, auxiliary data were recorded for the following: facility inflow (provided by UBCW), precipitation [retrieved from the National Ocean Atmospheric Administration's Integrated Surface Data database (Smith et al., 2011), station ID 725100-94746-2021 at Worcester Regional Airport], and date/time (hour of day) of measurement.

2.2. Baseline Analysis

To provide a baseline for comparison, the spectral data were grouped based on known drivers (listed above), with the average and standard deviation of each group calculated. Binning within groups is as follows: date (binned by month), hour of day (3-h bins), flow (five bins each containing 20% of data, sorted from lowest to highest flow), and precipitation (dry weather, < 0.75 in/h, and ≥ 0.75 in/h).

2.3. Data-Driven Group Identification

Many unsupervised clustering algorithms exist, each relying on different assumptions to represent the underlying structure of unlabeled data. Therefore, a comparative analysis of clustering models was performed, with the dual goals of producing clusters that were both distinct (low overlap between clusters) and tight (low within-cluster variability). Within-cluster variability was characterized by the sum of the squared Euclidean distances between each data point and the cluster centroid (Equation 1).

$$W = \sum_{k=1}^{K} \sum\nolimits_{x_i \in C_k} \parallel x_i - \overline{\mathrm{x}}_k \parallel^2 \tag{1}$$

where (x_i,...,x_N) is the set of observations, K is the number of clusters, $\overline{x}_k$ is the centroid of cluster C_k (coordinate-wise average of data points in C_k), and W is the sum of the within-cluster variabilities. Normalized x_i were used to promote clustering based on spectral shape rather than absolute magnitude.

Four methods were selected, each summarized briefly here with citations to the methodologies provided: K-means (Likas et al., 2003), Gaussian Mixture Model (GMM) (Reynolds, 2009), Hierarchical (Johnson, 1967), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering (Khan et al., 2014). K-means clustering groups data based on their proximity to centroids that are iteratively adjusted, with the K parameter identifying the number of centroids (i.e., number of clusters). GMM is similar to K-means but initializes, and then iteratively fits, K Gaussian distributions to the data and uses likelihood values to assign observations to clusters. Hierarchical clustering uses a similarity measurement (for this study, Euclidean distance) to group data into a linkage structure with a cutoff parameter determining the amount of similarity required to cluster measurements. DBSCAN clusters are identified based on the density of observations (the proximity of measurements to each other within the vector space), with the ϵ parameter establishing the maximum distance between observations allowed for clustering. The initial centroids for K-means and GMM were selected using the K++ algorithm. All clustering models were developed in Matlab R2021a (Mathworks, Natick, MA) using the Statistics and Machine Learning Toolbox.

Selecting the appropriate number of clusters was accomplished using the "Elbow method" (Yuan and Yang, 2019) with the number of clusters initially set to 1 and incremented until the point of diminishing returns (using W as the performance metric) was observed. The resulting trained classifier at the elbow point was used in this study for comparison with performance of the other classifiers, and relative performance of different classifiers was evaluated using W as the metric.

2.4. Case Study: 2020 SARS-CoV-2 Pandemic

As a case study, the selected best clustering algorithm was used to classify all collected data and specifically to evaluate the extent to which changes in patterns could be observed between days with no pandemic restrictions and days when pandemic restrictions were in place. In 2020, the SARS-CoV-2 pandemic caused major disruptions to societal behavior as local and federal governments imposed restrictions on business operations and movement of individuals. In Massachusetts the most severe restrictions were in place from March 17 to July 6, 2020. To minimize seasonal effects, this period was compared to March 17–June 7, 2021 during which time the restrictions were significantly loosened (approaching, though not yet at, "business as usual").

3. RESULTS

3.1. Wastewater Baselining

The mean and standard deviation ($\pm 1\sigma$) of the spectra grouped by known factors and their normalized transformations are shown in **Figure 1**. Typical relationships are visible in the data, e.g., higher flows tend to be more dilute (result in lower absorbance) than lower flows, and a similar effect is seen during storm events with intensity greater than 0.75 in/h. However, the clusters are neither tight nor distinct, with considerable overlap of the groups for all four factors considered. This level of residual uncertainty is typical and leads to similar uncertainty in associated forecasts which therein presents challenges for anomaly detection.

3.2. Unsupervised Clustering

The four utilized algorithms (K-means, GMM, Hierarchical, DBSCAN) identified different numbers of clusters (4, 4, 4+outliers, 2+outliers) and resulted in different W metrics (9.0×10^3, 1.1×10^4, 9.3×10^3, 2.2×10^4). For Hierarchical clustering, for computational tractability the number of clusters was iteratively increased by 1,000 to identify the elbow point at 6,000 measurements, however it should be noted that the four largest clusters comprise 89% of all measurements and most of the remaining 5,996 clusters consist of a single data point. Therefore, only the top four Hierarchical clusters were used in comparative analyses, with the remaining clusters treated as outliers. The K-means clustering trained classifier was selected as the optimal clustering method for this study, however results are reported both for this method and comparing generated clusters between methods.

Figures 2A,B show the K-means generated clusters, plotted, respectively, as normalized and unnormalized spectra. Two key observations follow. First, clusters 1–3 show similar characteristics, while cluster 4 has a very distinct shape, significantly greater standard deviation, and a minimal number of classified measurements. This implies this cluster represents atypical measurements, e.g., potentially requiring follow-up investigation by an operator to assess possible explanations (more discussion below). Second, when comparing the normalized K-means clusters (**Figure 2B**) to baseline groupings (**Figures 1E–H**), it is clear that the K-means cluster are more distinct: K-means clusters have only 25% overlap (at the 1-σ level), compared to 93, 99, 83, and 93% overlap for month-based, hour-based, flow-based, and precipitation-based groupings (number reported for normalized data). That is, for a new measurement, one may achieve a much higher level of confidence in assessing whether it is "normal" by comparing to these three clusters rather than requiring a deterministic model integrating effects of these multiple other factors (further implications discussed below).

To determine how "meaningful" the K-means identified clusters are, the similarity (or lack thereof) between clusters identified by different algorithms were compared. This was quantified by calculating the proportions of the K-means clusters within the clusters of each other technique (**Figure 3**, showing only top 4 clusters for Hierarchical which represent 94% of all measurements). K-means and GMM (similar techniques) result in very similarly defined clusters. Hierarchical clusters show substantial alignment with K-means clusters 2 and 3, with some movement of measurements to cluster 1. DBSCAN clustering appears to have been driven primarily by the atypical K-means cluster 4 data and are unlikely to be operationalizeable. In general, these results show consistency in the clustering where results are interpretable.

3.3. Clustering in Context of Known Factors and Case Study

To provide insight with respect to standard practice (**Figure 1**), K-means clusters are presented as a function of known factors in **Figures 4A–D**. **Figure 4A** shows that variability happens even at the scale of hours, with occurrence of the clusters roughly similar regardless of time of day. **Figure 4D** shows that the strong seasonal driver of wastewater is also visible in cluster proportional occurrence, although data from additional years would help clarify what fraction of this might be due to the exceptional conditions of 2020-21. These distinct seasonal patterns could help further refine anomaly detection, e.g., by weighting relative likelihoods of different clusters based on time of year.

Figures 4E,F show, respectively, data separated by time where there were no restrictions and under pandemic restrictions. While these effects are also certainly visible as a function of flow or other measures, visualizing the wastewater characteristics in the framework of the (small number of) K-means clusters provides a very simple lens through which to visualize the drastic differences between the 2 years.

4. DISCUSSION

The presented analysis can provide insight on several dimensions: on clustering techniques in comparison to current approaches, between different clustering approaches, and on use of K-means specifically for studying wastewater. On the first point, a major benefit of the clustering approach is decoupling the need to forecast multiple determining factors (e.g., precipitation, flow rates as driven by societal behavior) from an ability to set expectations for the "normal" compositional characteristics of wastewater inflow. This suggests a fast and simple method for development of an "early warning" system for anomalous inflows, i.e., simply by determining the likelihood that a given measurement falls within one of the 3 identified clusters. Operators then have an ability to set thresholds based on the risk tolerance of their facility, providing a balance between number of alerts (suggestions for further investigation of wastewater composition) and likelihood of detection of a potentially dangerous inflow condition.

Clustering of spectra can also be used to observe trends or patterns that may be difficult to visualize using other methods. For instance, some very strong patterns are observed in the monthly data (**Figure 4D**), with summer months consisting almost exclusively of K-means cluster 1 and November almost exclusively K-means cluster 3. This is in contrast with the "raw"

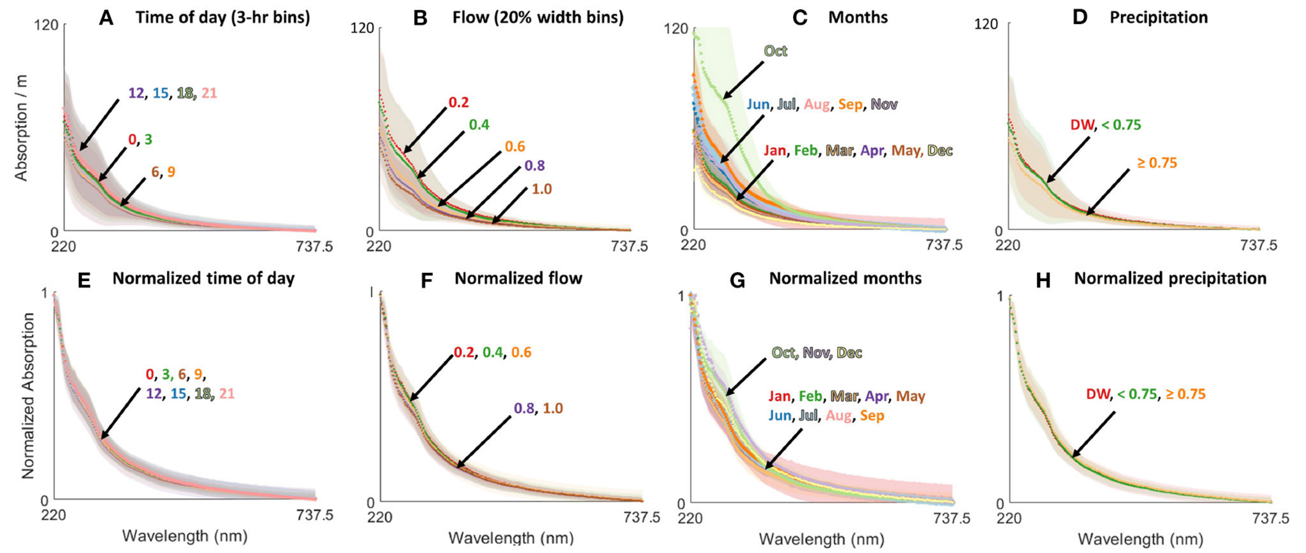


FIGURE 1 | Average wastewater spectra categorized by the following factors; **(A)** time of day (3-h bins), **(B)** flow (20% width bins), **(C)** month, and **(D)** precipitation (DW is dry weather, numbers are in/h). Normalized spectra categorized by the following factors; **(E)** time of day, **(F)** flow, **(G)** month, and **(H)** precipitation with binning as for **(A–D)**.

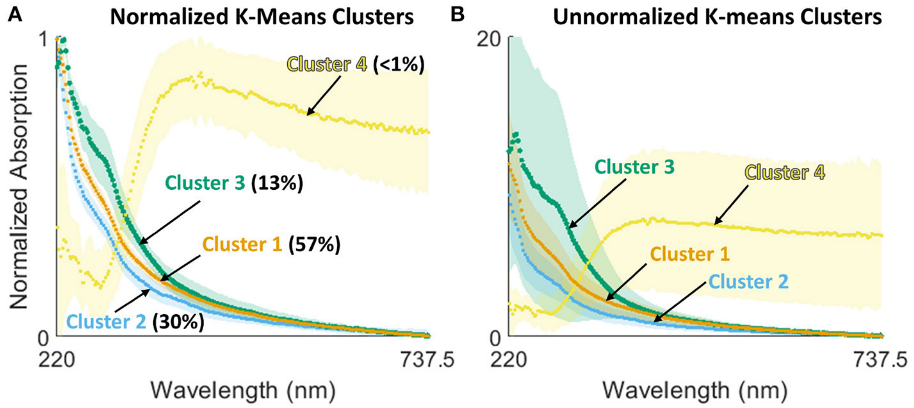


FIGURE 2 | K-means cluster spectra mean and standard deviation as **(A)** normalized measurements (labeled with percentage of measurements within that cluster) and **(B)** unnormalized measurements.

month-based analysis (**Figure 1C**) where only October is clearly differentiable from the rest of the data. Further while it could be speculated that the patterns relating to summer months tie to precipitation patterns, the distribution of clusters as a function of flow and of precipitation both show only weak relationships. This suggests that the seasonal patterns are more complex than dilution alone. For instance, there may be different influences of business operational seasons, temperatures of water within wastewater collection systems, etc. that collectively drive a changing wastewater character. The patterns revealed in comparing spring seasons in 2020 (pandemic restrictions) and 2021 (relatively normal operations) also highlight a new way to visualize the complex impacts of the changes in societal behavior on wastewater composition. A major change during spring 2020 was the closing of a wide range of businesses and movement for individuals to work from home (a major effect for the urban population of Worcester, MA); this reflects in a lessening of the strength in diurnal patterns and a redistribution of different cluster frequencies despite similar inflow magnitudes.

The classification algorithm comparison (**Figure 3**) demonstrates that, while the underlying assumptions of each approach do affect the fine details of the developed clusters, all approaches achieve some level of agreement even in this unsupervised context. This provides some level of comfort that the groupings are "real," i.e., that despite the fact that wastewater characteristics lie across a continuum, there are modalities which are statistically more common. The within-cluster variability also provides an ability to assess likelihood of a new sample

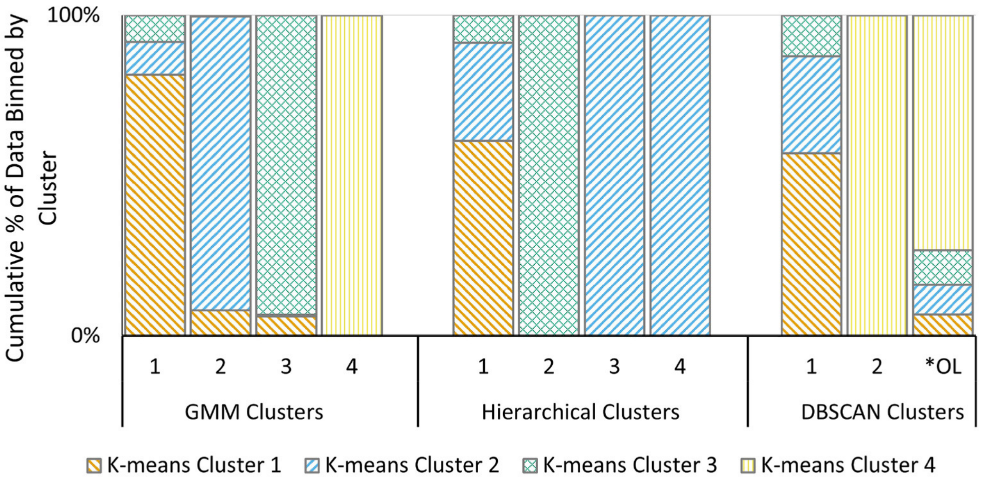


FIGURE 3 | Proportion of K-means clusters assigned to clusters generated by alternative techniques: GMM clustering, Hierarchical clustering, and DBSCAN clustering. *OL, Outliers.

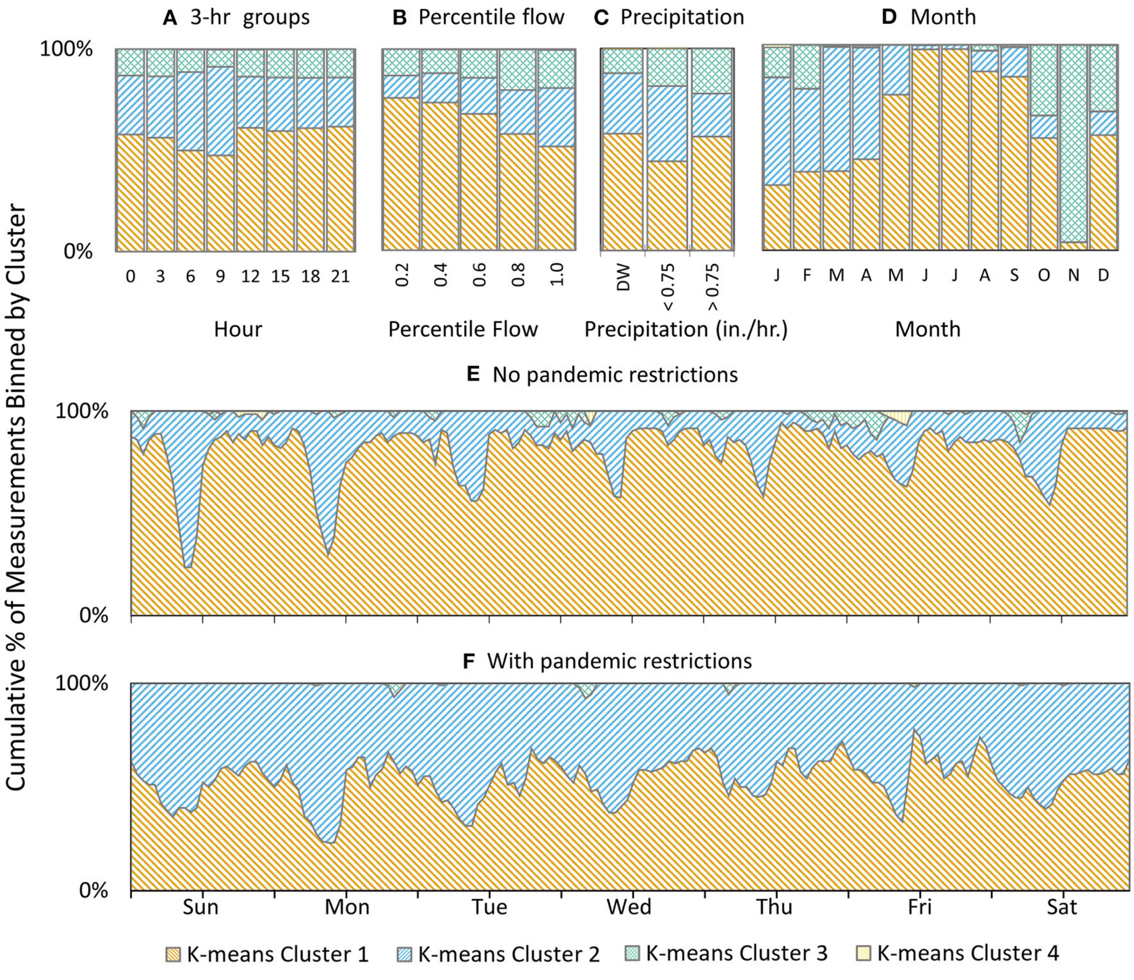


FIGURE 4 | Distribution of K-means clusters categorized by known factors, **(A)** time of day (3-h bins), **(B)** flow (20% width bins), **(C)** precipitation, **(D)** months, **(E)** no pandemic restrictions (March 17–June 7, 2021), and **(F)** with pandemic restrictions (March 17–July 6 2020.

matching each cluster, and therefore inversely, the likelihood of not belonging to any cluster, and use of these statistics from multiple dissimilar clustering approaches could therefore provide additional information to support anomaly detection. Finally, the results show that the distance-based approaches are strongly affected by outliers, which may prevent detection of nuanced patterns in the remaining data. For instance, in Hierarchical clustering 89% of observations are classified into the first cluster and in DBSCAN the largest cluster contains 99% of observations. Calibrating such approaches for wastewater applications may therefore require highly cleaned datasets or further research to evaluate applicability for anomaly detection.

Finally, K-means clustering specifically demonstrated utility in both clustering and anomaly detection in this application, identifying three relatively distinct modalities of wastewater (clusters 1–3) and relative outliers (cluster 4) (**Figure 2**). While the exact driver of the atypical points in cluster 4 has not been identified, the spectral shape does not resemble wastewater, and in fact the shape is more similar to readings taken when the sensor pathlength was not completely submerged in water – i.e., visual inspection confirms that these have been appropriately flagged as different from the three wastewater clusters. The K-means results are therefore appropriate to underpin a real-time alert system to assess the wastewater chemical characteristics and/or to verify sensor operations—a critically needed capability for managing wastewater treatment facilities.

In conclusion, this study developed a new approach to understanding modalities of "normal" in complex chemical mixtures by leveraging unsupervised clustering algorithms to identify underlying patterns, specifically building toward an online real-time approach for flagging chemical anomalies in wastewater systems. K-means clustering successfully identified relatively distinct and tight clusters in addition to a small fraction of atypical samples from approximately 1.5 years of data collected from an operating wastewater treatment facility. Visualizing the frequency of these clusters as a function of time and other variables, including societal behavior change driven by the SARS-CoV-2 pandemic, demonstrates a relatively simple approach to identifying patterns and changes in those patterns.

Moving forward, integration of data from other types of sensors could further refine the cluster definitions, and particularly could suggest cost-beneficial tradeoffs such as reducing the number of wavelengths that must be monitored to achieve classification in favor of less expensive hardware. Ultimately by achieving cost tradeoffs, it may be possible to distribute this type of sensor throughout the wastewater collection networks, therein providing operators with both longer lead times on anomaly warnings and more granular understanding of chemical changes throughout the system, ultimately providing more resiliency in operations of these critical infrastructure systems.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

AN and AM conceptualized the research. AN performed data acquisition, data analysis, model development, and prepared the manuscript with contribution from all authors. AM supervised the project and reviewed and revised the manuscript. Both authors have read and approved the content of the manuscript.

FUNDING

This work was supported by Northeastern University Faculty startup funds.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Upper Blackstone Clean Water facility and staff, especially Drs. Edris Taher and Ken Pousland. We would also like to thank Prof. Aron Stubbins at Northeastern University for use of the s::can instrumentation and Kevin Ryan for assistance in instrumentation operations.

REFERENCES

Baurès, E., Hélias, E., Junqua, G., and Thomas, O. (2007). Fast characterization of non domestic load in urban wastewater networks by UV spectrophotometry. *J. Environ. Monit.* 9, 959. doi: 10.1039/b704061j

Carreres-Prieto, D., García, J. T., Cerdán-Cartagena, F., and Suardiaz-Muro, J. (2020). Wastewater quality estimation through spectrophotometry-based statistical models. *Sensors* 20:5631. doi: 10.3390/s20195631

Chow, C. W., Liu, J., Li, J., Swain, N., Reid, K., and Saint, C. P. (2018). Development of smart data analytics tools to support wastewater treatment plant operation. *Chemom. Intell. Lab. Syst*. 177, 140–150. doi: 10.1016/j.chemolab.2018.03.006

Cicceri, G., Maisano, R., Morey, N., and Distefano, S. (2021). A machine learning approach for anomaly detection in environmental iot-driven wastewater purification systems. *Int. J. Environ. Ecol. Eng*. 15, 123–130.

Daniel, O., Denieul, M.-P., and Lemoine, C. (2008). "Novel parametric and non-parametric approach for the online and real-time evaluation of the variability of an effluent (EVE)," in *18th European Symposium on Computer Aided Process Engineering* (Lyon), 1–7.

Gruber, G. (2006). Practical aspects, experiences and strategies by using UV/VIS sensors for long-term sewer monitoring. *Water Pract. Technol*. 1, 21–26. doi: 10.2166/wpt.2006020

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika* 32, 241–254. doi: 10.1007/BF02289588

Khan, K., Rehman, S. U., Aziz, K., Fong, S., Sarasvady, S., and Vishwa, A. (2014). "DBSCAN: past, present and future," in *5th International Conference on the Applications of Digital Information and Web Technologies, ICADIWT 2014* (Bangalore: IEEE), 232–238.

Langergraber, G., Broeke, J. V. D., Lettl, W., and Weingartner, A. (2006). Real-time detection of possible harmful events using UV/vis spectrometry. *Spectroscopy Eur*. 18, 4–7.

Langergraber, G., Fleischmann, N., Hofstaedter, F., Weingartner, A., and Lettle, W. (2003). Detection of (unusual) changes in wastewater composition using UV/VIS spectroscopy. *Wastewater* (IWA Publishing), 1–8. Available online at: https://www.iwapublishing.com/books/iwa-publishing-digital-reference-library

Lepot, M., Torres, A., Hofer, T., Caradot, N., Gruber, G., Aubin, J. B., et al. (2016). Calibration of UV/Vis spectrophotometers: a review and comparison of different methods to estimate TSS and total and dissolved COD concentrations in sewers, WWTPs and rivers. *Water Res*. 101, 519–534. doi: 10.1016/j.watres.2016.05.070

Likas, A., Vlassis, N., and Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern Recognit*. 36, 451–461. doi: 10.1016/S0031-3203(02)00060-2

Loos, R., Carvalho, R., António, D. C., Comero, S., Locoro, G., Tavazzi, S., et al. (2013). EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents. *Water Res*. 47, 6475–6487. doi: 10.1016/j.watres.2013.08.024

Lourenço, N. D., Chaves, C. L., Novais, J. M., Menezes, J. C., Pinheiro, H. M., and Diniz, D. (2006). UV spectra analysis for water quality monitoring in a fuel park wastewater treatment plant. *Chemosphere* 65, 786–791. doi: 10.1016/j.chemosphere.2006.03.041

Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., and Pinheiro, H. M. (2012). Bioreactor monitoring with spectroscopy and chemometrics: a review. *Anal. Bioanal. Chem*. 404, 1211–1237. doi: 10.1007/s00216-012-6073-9

Mamandipoor, B., Majd, M., Sheikhalishahi, S., Modena, C., and Osmani, V. (2020). Monitoring and detecting faults in wastewater treatment plants using deep learning. *Environ. Monit. Assess*. 192:148. doi: 10.1007/s10661-020-8064-1

Reynolds, D. A. (2009). *Gaussian Mixture Models*. Technical report, Encyclopedia of Biometrics.

Rieger, L., Langergraber, G., and Siegrist, H. (2006). Uncertainties of spectral *in situ* measurements in wastewater using different calibration approaches. *Water Sci. Technol*. 53, 187–197. doi: 10.2166/wst.2006.421

Russo, S., Disch, A., Blumensaat, F., and Villez, K. (2020). "Anomaly detection using deep autoencoders for *in-situ* wastewater systems monitoring data," in *10th IWA Symposium on Systems Analysis and Integrated Assessment* (Copenhagen).

Schilperoort, R. P., Dirksen, J., Langeveld, J. G., and Clemens, F. H. (2012). Assessing characteristic time and space scales of in-sewer processes by analysis of one year of continuous in-sewer monitoring data. *Water Sci. Technol*. 66, 1614–1620. doi: 10.2166/wst.2012.115

Smith, A., Lott, N., and Vose, R. (2011). The integrated surface database: recent developments and partnerships. *Bull. Am. Meteorol. Soc*. 92, 704–708. doi: 10.1175/2011BAMS3015.1

Thomas, M.-F. (2017). "Chapter 10: industrial wastewater," in *UV-Visible Spectrophotometry of Water and Wastewater* (Amsterdam), 317–348.

Thomas, O., El Khorassani, H., Touraud, E., and Bitar, H. (1999). TOC versus UV spectrophotometry for wastewater quality monitoring. *Talanta* 50, 743–749. doi: 10.1016/S0039-9140(99)00202-7

Tsoumanis, C. M., Giokas, D. L., and Vlessidis, A. G. (2010). Monitoring and classification of wastewater quality using supervised pattern recognition techniques and deterministic resolution of molecular absorption spectra based on multiwavelength UV spectra deconvolution. *Talanta* 82, 575–581. doi: 10.1016/j.talanta.2010.05.009

Vaillant, S., Pouet, M., and Thomas, O. (2002). Basic handling of UV spectra for urban water quality monitoring. *Urban Water* 4, 273–281. doi: 10.1016/S1462-0758(02)00019-5

Vaillant, S., Pouet, M. F., and Thomas, O. (1999). Methodology for the characterization of heterogeneous fractions in wastewater. *Talanta* 50, 729–736. doi: 10.1016/S0039-9140(99)00200-3

Yuan, C., and Yang, H. (2019). Research on k-value selection method of k-means clustering algorithm. *Journal* 2, 226–235. doi: 10.3390/j2020016

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.

ORIGINAL RESEARCH
published: 16 December 2021
doi: 10.3389/frwa.2021.717852

Monitoring Lake Levels From Space: Preliminary Analysis With SWOT

Akhilesh S. Nair[1], Nitish Kumar[1], J. Indu[1,2]* and B. Vivek[3]

[1] Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India, [2] Interdisciplinary Programme in Climate Studies, Indian Institute of Technology Bombay, Mumbai, India, [3] Land and Water Management Research Group, Centre for Water Resources Development and Management (CWRDM), Kozhikode, India

Lakes are an essential component of biogeochemical processes, and variations in lake level are regarded as indicators of climate change. For more than a decade, satellite altimetry has successfully monitored variation in water levels over inland seas, lakes, rivers, and wetlands. Through altimetry, the surface water levels are measured at varying temporal scales depending on the orbit cycle of the satellite. The futuristic mission of Surface Water and Ocean Topography (SWOT) scheduled to be launched in year 2022 shall offer the spatial coverage and resolution suitable for water level estimation and volume calculation in small water bodies like lakes worldwide. With a radar interferometer in Ka-band, SWOT proposes to provide two-dimensional maps of water heights 21 days repeat orbit configuration. Cycle average SWOT datasets for land will be developed with higher temporal resolution, with temporal resolution varying geographically. This work assesses the potential of SWOT for monitoring water volumes over a case study lake by analyzing SWOT like synthetic data produced using the SWOT simulator developed by the Centre National d'Etudes Spatiales (CNES). With SWOT relying on a novel technology, the initial 90 days of this mission after launch shall focus on an extensive calibration and validation. Firsthand results of SWOT-simulated water levels and volumes are presented over a case study region in the tropical band, namely, Pookode Lake, in the ecologically fragile district of Wayanad, Kerala, India. It is the second-largest freshwater lake in Kerala that is being affected by anthropogenic activities, causing huge depletion in lake water storage in the last four decades. Our analysis indicated that the lake region is subjected to a rise in temperature of 0.018°C per year. We further assess the potential of remote sensing and SWOT data to monitor water storage of Pookode Lake, which is undergoing a rapid change. Results show that the proxy water surface elevations have immense potential in scientific studies pertaining to lake monitoring across the world. Overall, the study shows the potential of SWOT for monitoring the variability of water levels and volumes in this region.

Keywords: SWOT, interferometry, climate change, lakes, water levels

Edited by:
Auroop Ratan Ganguly,
Northeastern University, United States

Reviewed by:
Sumit Purohit,
Pacific Northwest National Laboratory (DOE), United States
Ramakrishna Tipireddy,
Pacific Northwest National Laboratory (DOE), United States

****Correspondence:***
J. Indu
indusj@civil.iitb.ac.in

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *31 May 2021*
Accepted: *01 November 2021*
Published: *16 December 2021*

Citation:
Nair AS, Kumar N, Indu J and Vivek B (2021) Monitoring Lake Levels From Space: Preliminary Analysis With SWOT. Front. Water 3:717852. doi: 10.3389/frwa.2021.717852

INTRODUCTION

Surface water bodies are important sources of water for civilization, and they have a direct impact on ecosystems (Dudgeon et al., 2006; Palmer et al., 2015). Changes in terrestrial water storage have a substantial impact on the hydrological cycle. Among these, lake system embodies a complex interaction between atmosphere, surface, and underground water, which responds strongly to climatic conditions. Changes in stored volume of surface water have geodynamical implications for Earth's rotation and have been shown to be proxy indicators of local and regional climate (Raymond et al., 2013; Seekell et al., 2014). Lake monitoring is one of the main objectives of the World Meteorological Organization (WMO) and the Global Climate Observing System (GCOS) because lakes are primary indicators of climate change (Adrian et al., 2009; Schindler, 2009). Climate change-related changes in precipitation and air temperature have an immediate impact on lake water storage (Wit and Stankiewicz, 2006). This makes monitoring temporal shifts in lake water volumes critical (Rahman and Di, 2017; Bonsal et al., 2019). While hydrometric stations provide the most reliable information on water levels, there exists a dearth of dense gauging stations for global lake monitoring (Pavelsky et al., 2014). This is particularly a concern for developing countries like India, where the primary focus is on monitoring of major rivers and large reservoirs using hydrometric stations (Gleason and Hamdan, 2017). For more than one decade, satellite altimetry missions (Topex-Poseidon, ERS-1, ERS-2, Jason-1, Envisat) have successfully monitored water levels from space (da Silva et al., 2014) using a plethora of sensors operating in visible region (Jiang et al., 2014; Nair and Indu, 2020), microwave region like Synthetic Aperture Radar (SAR) (Zeng et al., 2017; Nair and Indu, 2020), and a combination of both sensors (Crétaux and Birkett, 2006; da Silva et al., 2014; Bioresita et al., 2019; Grippa et al., 2019).

Existing studies by da Silva et al. (2014) demonstrated the ability of Envisat altimeter data to assess water storage in Amazon basins; in another study by Duan and Bastiaanssen (2013), variation in lake volume was evaluated using various altimetry databases and laser altimeters for different lakes in the United States, Ethiopia, and the Netherlands. Similarly, Crétaux et al. (2016) used altimetry measurements to observe the effect of climate change on variation in lake water storage over the Tibetan plateau. Despite the high potential of altimetry measurements, a main drawback is the narrow swath of altimeters leading to low spatial sampling and data acquisition problems for small and medium lakes (Grippa et al., 2019). This shortcoming can be overcome by using Interferometric SAR (InSAR), which uses SAR data to determine relative water level changes in lakes. Siles et al. (2020) used this method to measure water level changes in lakes and wetlands across Canada. Despite these measures, maintaining a continuous lake monitoring inventory is daunting due to the difficulty in obtaining optical imagery, radar, and altimetry measurements for a particular lake body within a specific temporal window.

The future Surface Water and Ocean Topography (SWOT) mission, which is set to launch in 2022, will be the first of its kind to have a bistatic SAR interferometric wide swath altimeter. SWOT mission is a collaborative effort between NASA and the Centre National d'Études Spatiales (CNES), along with support from the Canadian Space Agency (CSA) and the United Kingdom Space Agency (Biancamaria et al., 2016). The SWOT satellite will have an onboard Ka-band SAR, which has a great potential for detecting water surface elevation and slope for rivers and lakes (Biancamaria et al., 2016). It will benefit from both radar altimeters for water level detection and SAR imagery for high spatial resolution. The SWOT will capture water surfaces of rivers (width more than 100 m) and lakes (surface area more than 0.0625 m^2), with a revisit period of 21 days, globally between 78°S and 78°N (Pavelsky et al., 2014).

Previous studies on SWOT characteristics have concentrated on river hydrology, such as the ability to estimate streamflow (Frasson et al., 2017), streamflow assimilation (Oubanas et al., 2018), and bathymetry (Yoon et al., 2012). Very few studies have examined the contribution of SWOT in lake hydrology: Grippa et al. (2019) demonstrated SWOT capabilities to monitor lakes in Sahel in Africa, and Bergeron et al. (2020) indicated SWOT capabilities for measuring lakes in Canada. Many of these studies are mostly focused on the United States and South America. Temporally, SWOT mission shall sample each location at least twice every month with the exact timing depending mostly on the latitude. With SWOT relying on a novel technology, this mission shall require an in-depth calibration and validation (CalVal). In the year 2022, SWOT mission is known to fly a specific orbit dedicated to the CalVal activities, which is expected to last for 90 days upon launch. CalVal orbit is designed in such a manner that the satellite will follow the same track every day. Keeping these factors in mind, our research looks at Pookode Lake in Kerala's Wayanad area to evaluate the efficacy of synthetic SWOT data. This is due to the fact that SWOT will capture Pookode Lake in CalVal orbit, giving daily data for the lake during the CalVal phase.

With the high potential in SWOT for hydrological studies of reservoirs and rivers, it is crucial to check the efficiency of the SWOT dataset before its launch. Previous studies have tried to simulate SWOT-like data using different ways such as by corrupting errors into merged observations from satellite altimetry, satellite images, and gauges (Lee et al., 2010); a different approach is to use simulators to generate data using orbit pass plan and expected errors (Rodriguez and Moller, 2004; Solander et al., 2016). SWOT hydrology simulator is available from both the collaborators of the mission, i.e., from Jet Propulsion Laboratory (JPL) and CNES.

In the present study, we utilize the CNES Large-Scale SWOT Hydrology Simulator to simulate data over Pookode Lake. Furthermore, the difference in water storage is first computed using interferometric SAR for Pookode Lake. In the second step, synthetic SWOT data are generated for this lake and evaluated for its ability to compute lake volume. This is the first research to demonstrate the capabilities of remote sensing to detect lake volume fluctuations and the potential of future SWOT mission for lake hydrology in India.

STUDY REGION

This research work has been carried out on the second largest lake in Kerala, a state in southern India (**Figure 1**). Pookode Lake is a fresh water body encompassing an area of 0.085 sq. km. It is situated at an altitude of 770 m above mean sea level, and the main source of water is precipitation from the southwest monsoon (during June, July, August, and September) and the northeast monsoon (during October, November, and December).

The lake region receives a mean annual rainfall of 4,000 mm (Sandeep et al., 2012) and is subjected to a maximum temperature of 35°C in summer to 7°C in winter. Along with the major source of water from precipitation, the lake is also fed by two streams of the Kabini river. Because of its primary source of precipitation, it is critical to track fluctuations in lake volume, as Kerala has experienced two extreme floods in a row in 2018 and 2019.

SWOT HYDROLOGY SIMULATOR

The aim of the SWOT satellite is to provide hydrologists with unprecedented observations of terrestrial water bodies in order to track changes in water volume. These data are crucial for determining surface water availability and preparing for major water-related disasters like floods and droughts. Readers are recommended to the SWOT mission page for further information on the SWOT satellite's goal pertaining to hydrology (https://swot.jpl.nasa.gov/science/hydrology/). Except for a few places in Greenland's interiors and portions of the Sahara Desert, SWOT will catch inland water bodies worldwide. Users will have access to three types of SWOT datasets: pixel cloud, raster, and vector. The pixel cloud will contain geolocated heights, backscatter, geophysical fields, and flags over tile of 64×64 km^2. Raster products will provide water surface height, area, water fraction, backscatter, and geophysical information at 100 and 250-m resolution. The vector shapefiles will include information such as water surface elevation, slope, width, and calculated discharge. The further section provides details on how proxy SWOT point cloud data is generated for Pookode Lake.

In this study, CNES SWOT large-scale hydrology simulator (CNES, 2020) is used to generate synthetic SWOT data. The Water Surface Elevation (WSE) is corrupted with SWOT measurement errors and spatial sampling based on SWOT orbital configuration to produce synthetic SWOT results. White noise, dark water effect, satellite positioning error, geolocation error, and errors attributable to atmospheric attenuation are all factors that are taken into account by the SWOT simulator. The measurement noise of the SWOT interferometric phase difference is a prime example of random errors. The system SNR (Signal-to-Noise Ratio), the duration of the interferometric baseline, and the processing algorithm all play a role in the random error contribution. The dark water effect uncertainty refers to the inaccuracy that occurs when a water pixel is incorrectly categorized as land because of dark radar returns. Satellite positing errors are caused by a lack of knowledge about the spacecraft's roll angle, changes in the baseline due to thermal contraction or expansion, system timing mistakes, and phase errors produced by the antennas or electronics. Geolocation errors occur in satellite images when pixels are incorrectly georeferenced. Layover effects incurred to near-nadir instruments are ignored in SWOT simulator used for present study. However, due to the strong contrast between water and land observed in Ka-band, the SWOT mission is anticipated to have a low contribution of error from layover (Elmer et al., 2020). Despite its limitations in representing the complete SWOT error budget, the large-scale SWOT simulator is a reliable platform for generating synthetic SWOT data for this study. To generate synthetic SWOT observations, the SWOT simulator uses a lake water extent shapefile with true WSE (details provided in Extraction of Water Level Using Interferometric SAR section) as one of the attributes (CNES, 2020).

To generate SWOT data for time corresponding to before (2015) and after (2020) flood events in Kerala, lake water extent shapefiles are created using Sentinel-1 dataset in Google Earth Engine (GEE). To generate water extent shapefiles, Sentinel-1 SAR imagery for Pookode Lake is selected for the study period. The SAR datasets are processed for noise filtering using speckle filtering algorithm. After noise correction, the images are classified for water using a binary thresholding technique, thereby generating a water mask raster. In this process, pixels with VV polarization backscatter less than −16 decibels are considered as containing water, and the corresponding pixel values are set to one, and the remaining pixels are set to zero (Bonnema and Hossain, 2019). This classified raster files are used to generate water extent shapefile. The shapefile created by GEE is pre-processed with Quantum Geographic Information System (QGIS, 2020) to add reference WSE as an attribute. Further section summarizes some of the most important steps in configuring and running the SWOT simulator in this study.

(a) In addition to WSE, the input shapefile is added with attributes such as river flag (RIV_FLAG) that is set to 0 for lakes and 1 for rivers. Two water extent shapefiles (pre-flood, 2015 and post-flood, 2020) are prepared and preprocessed for compatibility with SWOT simulator.

(b) Following the generation of input water extent files, a SWOT satellite pass plan is generated based on its orbit configuration. Synthetic SWOT data are produced using the SWOT simulator based on this SWOT plan. The SWOT simulator creates WSE in the form of a point cloud, taking into account all pixel size inconsistencies and errors in the SWOT swath. The point cloud offers high-quality observations for study.

EXTRACTION OF WATER LEVEL USING INTERFEROMETRIC SAR

In this study, we used SAR from Sentinel-1 data to compute difference in lake volume. The interferometric large swath mode collects data over a 250-km swath with a spatial resolution of 5 m by 20 m. To estimate variation in water level using SAR imagery, it is required to obtain short temporal baselines and broad perpendicular baselines. Sentinel-1 was introduced primarily to detect deformations (DInSAR), not to extract water levels.

For this study, two sets of imageries that met the criteria over the study region were separated into master and slave images and co-registered to extract information of phase differences. The details of SAR dataset are provided in **Table 1**. The first series of SAR images shows the water level in the lake in 2015 before two flood occurrences (2019, 2018), while the second set shows the water level after the flood in 2020. The main steps adopted for generating water level change using interferometry

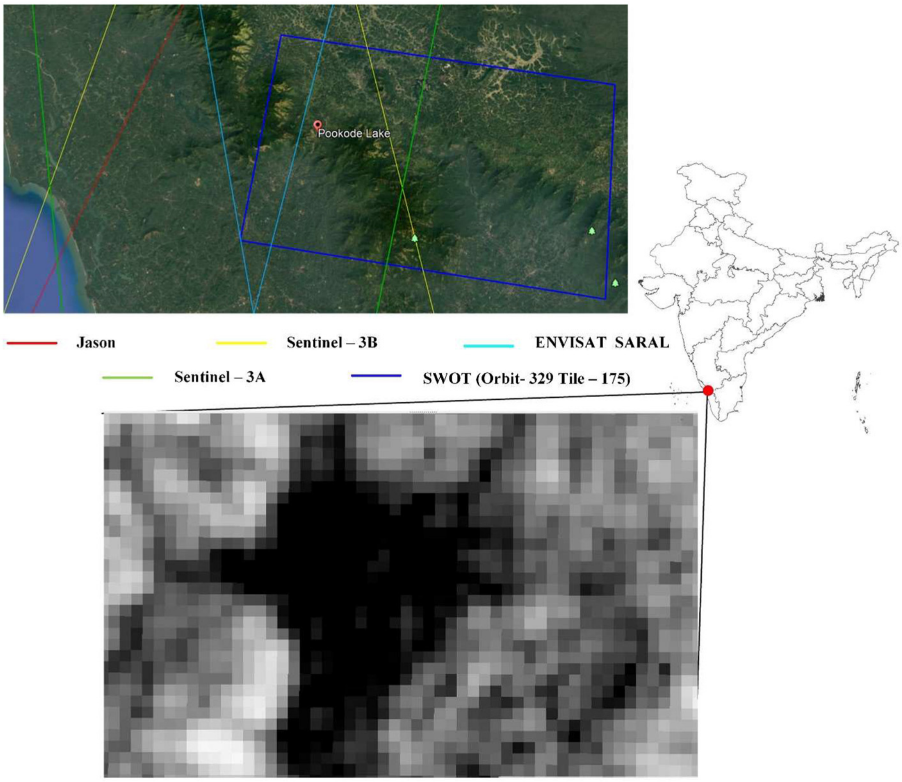


FIGURE 1 | Location of Pookode Lake with backscatter image of lake observed by Sentinel-1. Along with orbit tracks of altimetry missions and SWOT footprint over lake.

TABLE 1 | Summary of SAR image characteristics.

Master image acquisition	Slave image acquisition	Perpendicular baseline distance
July 20, 2015	July 8, 2015	151 m
October 21, 2020	October 9, 2020	90 m

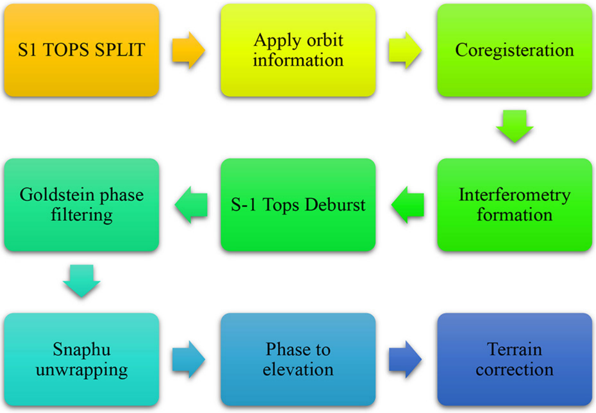


FIGURE 2 | Flowchart of methodology adopted to extract water level using interferometry.

is described. The key steps of process are shown in the flowchart in **Figure 2**.

Initial interferogram was developed with phase difference between two images using Equation 1.

$$\varnothing = \varnothing_{DEM} + \varnothing_{flat} + \varnothing_{disp} + \varnothing_{atm} + \varnothing_{noise} \ldots \quad (1)$$

Where $\varnothing_{DEM}$ refers to topographic effects, $\varnothing_{disp}$ denotes displacement, $\varnothing_{flat}$ indicates earth curvature error, $\varnothing_{atm}$ refers to atmospheric effects, and $\varnothing_{noise}$ denotes noise effects (Ferretti et al., 2001).

After this, de-burst algorithm (Esposito et al., 2020) was applied to interferogram to process the individual bursts from one sub swath together into a single sub-swath (Moran, 2006). Later, Goldstein filtering was applied to the image to remove noise and speckles and to enhance the signal-to-noise ratio of the image. In the interferogram, the interferometric phase is ambiguous and only known within the scale of 2π. To be able to relate the interferometric phase to the topographic height, the phase must first be unwrapped using Equation 2.

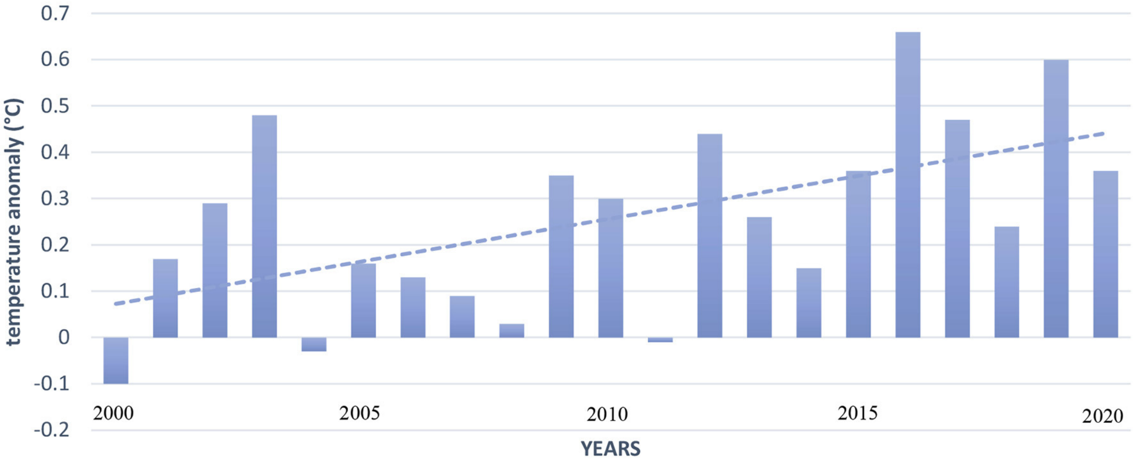


FIGURE 3 | Variation in temperature anomaly over a lake for a period of two decades from 2000 to 2020. The trend line indicates an increase in temperature of 0.018°C per year.

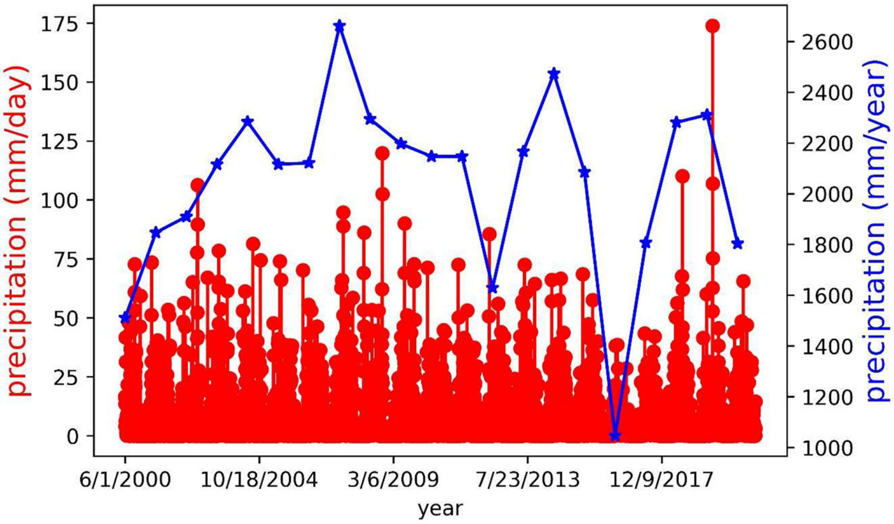


FIGURE 4 | Variation in daily precipitation from GPM IMERG over Pookode Lake over a period of 20 years from 2000 to 2020 along with variation in annual precipitation.

$$\Delta Z = \frac{-\phi\lambda}{4\pi \cos\theta} + n \ldots \tag{2}$$

Where λ is the wavelength of the SAR signal, ϕ is the unwrapped interferometric phase (surface displacement), θ is the incidence angle, n is the noise, and ΔZ is displacement.

After phase unwrapping, the data are processed for terrain correction. It is to be noted that terrain correction shall geocode the image by correcting SAR geometric distortions using the SRTM digital elevation model (DEM). Geocoding converts an image from slant range or ground range geometry into a map coordinate system. Terrain geocoding is used for correction of inherent geometric distortions, such as foreshortening, layover, and shadow. In this study, the shape of lake was assumed as pyramidal frustrum to calculate change in volume from year 2015 to 2020. For master images, surface area extent was computed from masked water bodies of sentinel photos. Using this information, the change in volume (Δv) for Pookode Lake after flood events was computed using Equation 3.

$$\Delta v = \frac{(h_1 - h_2)\left(A_1 + A_2 + \sqrt{A_1 A_2}\right)}{3} \ldots \tag{3}$$

Where h_1 is the high water level in the lake (during the study period, i.e., 2020), h_2 is the low water level (2015), A_1 is lake surface area during 2020, and A_2 is lake surface area during 2015.

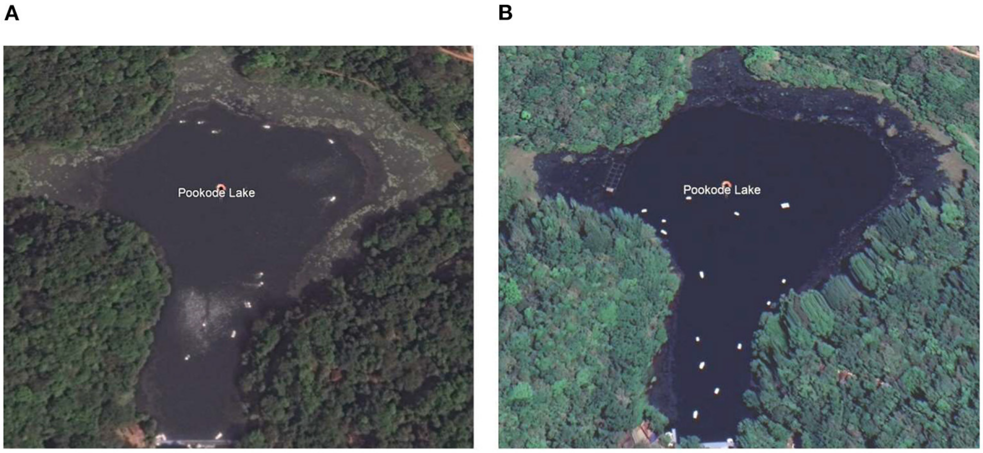


FIGURE 5 | Variation in water extent in Pookode Lake for **(A)** February 2016 and **(B)** February 2021.

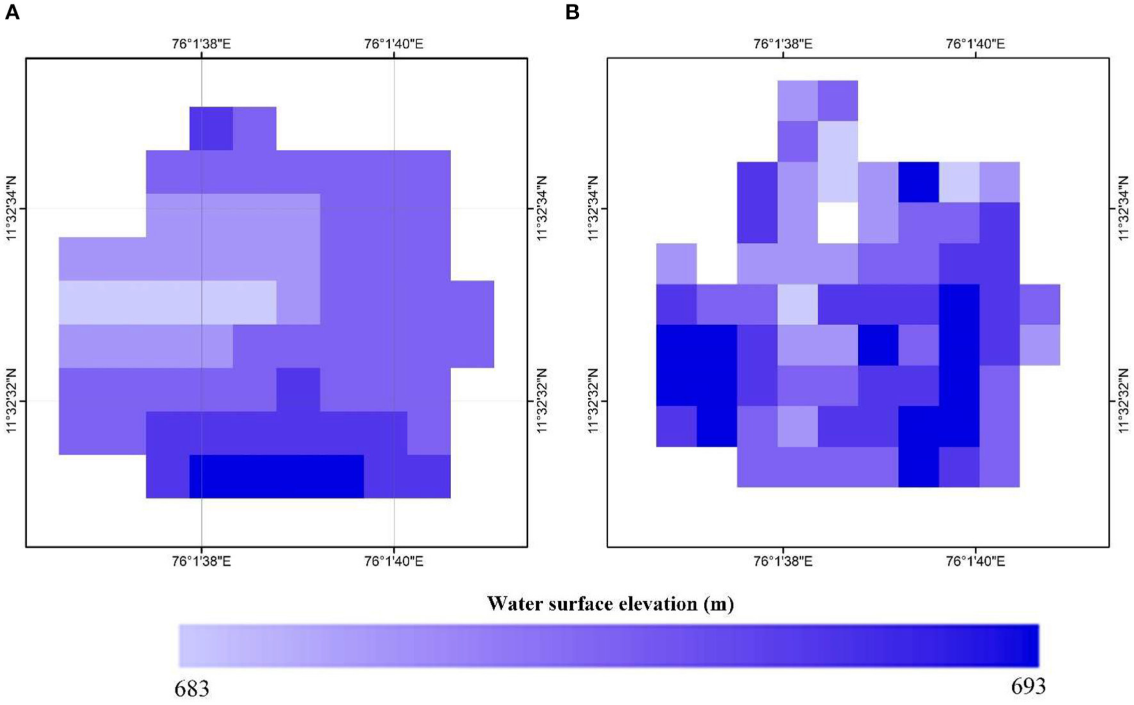


FIGURE 6 | Variation in water level in Pookode Lake extracted using interferometric SAR before and after flood in **(A)** 2015 and **(B)** 2020.

RESULTS

Storage of lake water is highly susceptible to temperature changes. As a first step, we use NOAA temperature measurements (https://www.climate.gov/maps-data/datasets) to assess the temperature anomaly (with respect to 1981 to 2010 average) over Pookode Lake. The lake temperatures were evaluated for two decades, from 2000 to 2020. The annual temperature anomaly is depicted in **Figure 3**. In this study, the increasing surface temperature is clearly evident. Over the lake, temperatures are found to be rising at a rate of 0.018°C per year. To compute this, we applied Mann-Kendall test on temperature data and results indicated a significant trend at 95% confidence interval with a *p*-value of 0.0217. The trend in temperature time series was computed based on Sen's slope indicator. If the temperature rises,

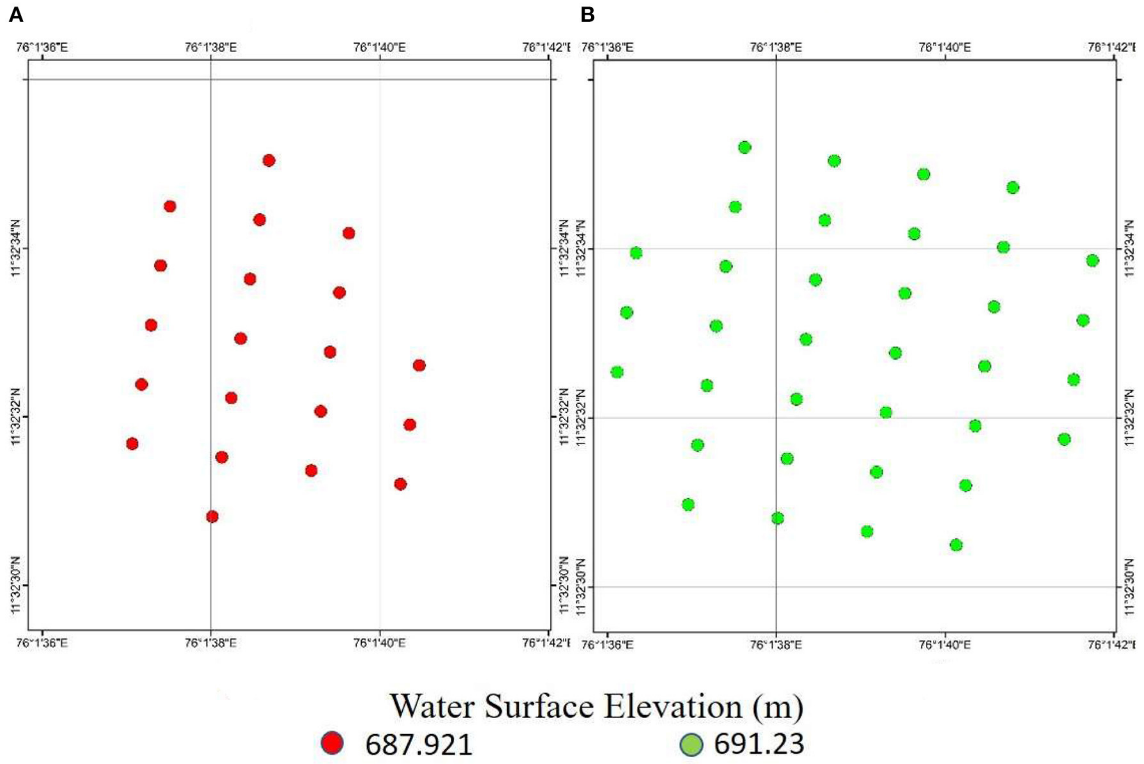


FIGURE 7 | Variation in water level in Pookode Lake as observed by synthetic SWOT data before and after flood in **(A)** 2015 and **(B)** 2020.

evapotranspiration increases, resulting in low runoff into the lakes.

Precipitation is a significant element that underpins the maintenance of lake's water balance. Precipitation has varied considerably in the study area. To evaluate this, we examined precipitation data from the Global Precipitation Mission (GPM) IMERG pixel over the study region. **Figure 4** represents the variation in daily precipitation over a 20-year period, from 2000 to 2020. In 2018 and 2019, the study area experienced significant floods. In 2019, the region received a rainfall of more than 170 mm in a single day. **Figure 4** also represents the annual precipitation variation across the study area. During 2018 and 2019, the area received a high amount of rain.

We compared high-resolution optical images of the lake from Google Earth to visually examine lake water storage after two floods. **Figure 5** depicts changes in the lake that happened prior to the flood (February, 2016) (**Figure 5A**) and after the flood (February, 2021) (**Figure 5B**). This figure shows the rise in lake water storage after successive flood events. This is because precipitation is the lake's primary source of fresh water. To measure this rise in lake water, we used an interferometric method (explained in section 4) to calculate water level increases. We chose interferometry for this challenge because there was no overpass of altimetry missions around the lake.

Figure 6 depicts the interferometry-derived water level estimates. **Figure 6A** illustrates the water level in 2015 before the flood, while **Figure 6B** illustrates the water level after the flood. The increase in lake water storage is distinctly evident, as the mean water elevation in **Figure 6A** is 687.9 m, while the lake water level after flooding is 691.2 m (**Figure 6B**). During this time, the lake's water level increased by about 3 m. We then used the technique described in section 4 to quantify the increase in lake volume. After analysis, the change in volume was observed as 63,739 m^3.

Figure 7 represents synthetic SWOT point cloud over the Pookode Lake during pre-flood (2015) and post-flood (2020). Each point observation in point cloud comprises information on location, water surface elevation, and other attributes. **Figure 7A** represents point cloud observed by future SWOT mission when water storage was low in lake. The crucial thing to remember is that over Pookode Lake, there was no measurement error from SWOT. As there are no overpasses of altimetry missions over Pookode Lake (as shown in **Figure 1**), the SWOT mission will be a promising source to monitor its water balance. The altimetry missions checked for overpass include Jason, Sentinel-3A, Sentinel-3B, and Envisat SARAL. These dense point cloud observations will have high potential in calculating change in lake volume. For the current study, reference height used for the SWOT simulator is obtained from DInSAR. The volume change exactly matches the volume computed from a previous step as mentioned above.

CONCLUSION

This study presents firsthand the results of the SWOT hydrology simulator to small lakes in India. The SWOT simulator used in this study does not account for layover impacts; however, given the location of Pookode Lake, the layover impact can be neglected. The synthetic observations demonstrated high accuracy in monitoring water level for the lake. The lake is subjected to climate change with a rise in temperature of 0.018°C per year and frequent heavy rainfall. Monitoring lake storage is crucial for water balance analysis and mitigates for future sustainable water for society and recreational purpose. Our study indicated an increase of 63,739 m^3 in water in 2020 as compared to 2015. The heavy precipitation in 2018 and 2019 contributed to the increase in lake volume. However, our analysis shows that the temperature in Pookode Lake is rising at a pace of 0.018°C each year. This rise in temperature will result in a large amount of evaporation from the lake. As a result, it is critical to focus on such lakes in order to reduce the risk of climate change.

The SWOT mission has the potential to provide an unparalleled capacity to monitor geographically diverse lakes and wetland WSE like Pookode Lake; therefore, continued scientific advancement is required. The capacity to address water balance components that have previously depended on limited ground-based monitoring would be enhanced by a revolutionary SWOT satellite-based surface water monitoring methodology. Following the launch of SWOT, constructing hypsometric curves for all lakes using SWOT will be tremendously helpful in building a lake inventory. This will also improve our understanding on impact of climate and climate change on lake water balance on a global scale.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: https://scihub.copernicus.eu/.

AUTHOR CONTRIBUTIONS

AN and NK: conducting the study, reporting results, and drafting manuscript. JI: conceptualizing study and drafting manuscript. BV: editing manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors acknowledge the support by the DST CNRS project through IFC/4126/DST-CNRS/2018-19/1/1726 and the DST-Center of Excellence in Climate studies, IIT Bombay under project DST/CCP/CoE/140/2018 (G).

ACKNOWLEDGMENTS

The authors are grateful to Prof. Faisal Hossain, University of Washington through the NASA SWOT Early Adopter project from IIT Bombay. The authors acknowledge the collaboration through CWRDM, Kerala through the NASA SWOT Early Adopter project. The authors wish to thank the support of DST Centre of Excellence in climate studies, IIT Bombay under project DST/CCP/CoE/140/2018 (G). The authors thank the SWOT Early Adopter Program supported by NASA Applied Science Program.

REFERENCES

Adrian, R., O'Reilly, C. M., and Zagarese, H. (2009). Lakes as sentinels of climate change. *Limnol. Oceanogr.* 54, 2283–2297. doi: 10.4319/lo.2009.54.6_part_2.2283

Bergeron, J., Siles, G., Leconte, R., Trudel, M., Desroches, D., and Peters, D. L. (2020). Assessing the capabilities of the Surface Water and Ocean Topography (SWOT) mission for large lake water surface elevation monitoring under different wind conditions. *Hydrol. Earth Syst. Sci.* 24, 5985–6000. doi: 10.5194/hess-24-5985-2020

Biancamaria, S., Lettenmaier, D. P., and Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. *Surv. Geophys.* 37, 307–337. doi: 10.1007/s10712- 015-9346-y

Bioresita, F., Puissant, A., Stumpf, A., and Malet, J-. P. (2019). Fusion of Sentinel-1 and Sentinel-2 image time series for permanent and temporary surface water mapping. *Int. J. Remote Sens.* 40, 1–24. doi: 10.1080/01431161.2019.1624869

Bonnema, M., and Hossain, F. (2019). Assessing the potential of the surface water and ocean topography mission for reservoir monitoring in the Mekong River Basin. *Water Resour. Res.* 55. doi: 10.1029/2018WR023743

Bonsal, B. R., Peters, D. L., Seglenieks, F., Rivera, A., and Berg, A. (2019). "Changes in freshwater availability across Canada," in E. Bush and D. S. Lemmen (eds.) Canada's changing climate report. Ottawa, ON: Government of Canada, Ottawa, pp. 261–342.

CNES (2020) *SWOT Hydrology Toolbox*. Available online at: https://github.com/CNES/swot-hydrology-toolbox

Crétaux, J. F., and Birkett, C. (2006). Lake studies from satellite radar altimetry. *Comptes Rendus-Geosci.* 338, 1098–1112. doi: 10.1016/j.crte.2006.08.002

Crétaux, J. F. Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G., and Maisongrande, P. (2016). Lake volume monitoring from space. *Surv. Geophys.* 37, 269–305. doi: 10.1007/s10712-016-9362-6

da Silva, D. A., Calmant, J. S., Seyler, S., Moreira, F., Oliveira, D. M. D., and Monteiro, A. (2014). Radar altimetry aids managing gauge networks. *Water Resour. Manag.* 28, 587–603. doi: 10.1007/s11269-013-0484-z

Duan, Z., and Bastiaanssen, W. G. M. (2013). Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sens. Environ.* 134, 403–416. doi: 10.1016/j.rse.2013.03.010

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., et al. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol. Rev.* 81, 163–182. doi: 10.1017/S1464793105006950

Elmer, N. J., Hain, C., Hossain, F., Desroches, D., and Pottier, C. (2020). Generating proxy SWOT water surface elevations using WRF-Hydro and the CNES SWOT Hydrology Simulator. *Water Resour. Res.* 56:e2020WR. https://doi.org/10.1029/2020WR027464[doi: 10.1029/2020WR027464

Esposito, G., Marchesini, I., Mondini, A., Reichenbach, P., Rossi, M., and Sterlacchini, S. (2020). A spaceborne SAR-based procedure to support the detection of landslides. *Nat. Hazards Earth Syst. Sci.* 20, 2379–2395. doi: 10.5194/nhess-20-2379-2020

Ferretti, A., Prati, C., Rocca, F. (2001). Permanent scatterers in SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* 39, 8–20. doi: 10.1109/36.898661

Frasson, R. P. D. M., Wei, R., Durand, M., Minear, J. T., Domeneghetti, A., Schumann, G., et al. (2017). Automated river reach definition strategies: applications for the surface water and ocean topography mission. *Water Resour. Res*. 53, 8164–8186. doi: 10.1002/2017WR020887

Gleason, C. J., and Hamdan, A. N. (2017). Crossing the (watershed) divide: satellite data and the changing politics of international river basins. *Geogr. J*. 183, 2–15. doi: 10.1111/geoj.12155

Grippa, M., Rouzies, C., Biancamaria, S., Blumstein, D., Cretaux, J. F., Gal, L., et al. (2019). Potential of SWOT for monitoring water volumes in sahelian ponds and lakes. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*. 12, 2541–2549. doi: 10.1109/JSTARS.2019.2901434

Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., and Xiao, T. (2014). An automated method for extracting rivers and lakes from landsat imagery. *Remote Sens*. 6, 5067–5089. doi: 10.3390/rs6065067

Lee, H., Durand, M., Jung, H. C., Alsdorf, D., Shum, C. K., and Sheng, Y. (2010). Characterization of surface water storage changes in Arctic lakes using simulated SWOT measurements. *Int. J. Remote Sens*. 31, 3931–3953. doi: 10.1080/01431161.2010.483494

Moran, M. S., McElroy, S., Watts, J. M., and Peters-Lidard, C. D. (2006). Radar remote sensing for estimation of surface soil moisture at the watershed scale. *Model Remote Sens. Appl. Agric. Model Remote Sens. Appl. Agric*. 91–106.

Nair, A. S., and Indu, J. (2020). "Reservoir water surface area detection using satellite observations for synthetic SWOT data simulation," in *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)*, 2020, 160–163.

Oubanas, H., Gejadze, I., Malaterre, P-. O., Durand, M., Wei, R., Frasson, R. P. M., et al. (2018). Discharge estimation in ungauged basins through variational data assimilation: the potential of the SWOT mission. *Water Resour. Res*. 54, 2405–2423. doi: 10.1002/2017WR021735

Palmer, S. C. J., Kutser, T., and Hunter, P. D. (2015). Remote sensing of inland waters: challenges, progress and future directions. *Remote Sens. Environ*. 157, 1–8. doi: 10.1016/j.rse.2014.09.021

Pavelsky, T. M., Durand, M. T., Andreadis, K. M., Beighley, R. E., Paiva, R. C. D., Allen, G. H., et al. (2014). Assessing the potential global extent of SWOT river discharge observations. *J. Hydrol*. 519, 1516–1525. doi: 10.1016/j.jhydrol.2014.08.044

QGIS (2020). Available online at: https://qgis.org/en/site

Rahman, M. S., and Di, L. (2017). The state of the art of spaceborne remote sensing in flood management. *Nat. Hazards* 85, 1223–1248. doi: 10.1007/s11069-016-2601-9

Raymond, P. A., Hartmann, J., and Lauerwald, R. (2013). Global carbon dioxide emissions from inland waters. *Nature* 503, 355–359. doi: 10.1038/nature12760

Rodriguez, E., and Moller, D. (2004). "Measuring surface water from space," In *AGU Fall Meeting Abstracts*, Vol. 2004, pp. H22C-08.

Sandeep, K., Warrier, A., Harshavardhana, B., and Shankar, R. (2012). Rock Magnetic Investigations of Surface and sub-surface soil samples from five Lake Catchments in Tropical Southern India. *Int. J. Environ. Res*. 6, 1–18. doi: 10.22059/ijer.2011.467

Schindler, D. W. (2009). Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnol. Oceanogr*. 54, 2349–2358. doi: 10.4319/lo.2009.54.6_part_2.2349

Seekell, D. A., Carr, J. A., Gudasz, C., and Karlsson, J. (2014). Upscaling carbon dioxide emissions from lakes. *Geophys. Res. Lett*. 41:7555. doi: 10.1002/2014GL061824

Siles, G., Trudel, M., Peters, D. L., and Leconte, R. (2020). Hydrological monitoring of high-latitude shallow water bodies from highresolution spaceborne D-InSAR. *Remote Sens. Environ*. 236:111444. doi: 10.1016/j.rse.2019.111444

Solander, K. C., Reager, J. T., and Famiglietti, J. S. (2016). How well will the Surface Water and Ocean Topography (SWOT) mission observe global reservoirs? *Water Resour. Res*. 52, 2123–2140. doi: 10.1002/ 2015WR017952

Wit, D. E., and Stankiewicz, M. (2006). Changes in surface water supply across Africa with predicted climate change. *Science* 311, 1917–1921. doi: 10.1126/science.1119929

Yoon, Y., Durand, M., Merry, C. J., Clark, E. A., Andreadis, K. M., and Alsdorf, D. E. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *J. Hydrol*. 464/465, 363–375. doi: 10.1016/j.jhydrol.2012.07.028

Zeng, C., Richardson, M., and King, D. J. (2017). The impacts of environmental variables on water reflectance measured using a lightweight unmanned aerial vehicle (UAV)-based spectrometer system. *ISPRS J. Photogramm. Remote Sens*. 130, 217–230. doi: 10.1016/j.isprsjprs.2017.06.004

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.

frontiers in Water

ORIGINAL RESEARCH
published: 06 January 2022
doi: 10.3389/frwa.2021.734518

The Effect of Pore-Scale Two-Phase Flow on Mineral Reaction Rates

Pei Li, Hang Deng and Sergi Molins*

Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

In various natural and engineered systems, mineral–fluid interactions take place in the presence of multiple fluid phases. While there is evidence that the interplay between multiphase flow processes and reactions controls the evolution of these systems, investigation of the dynamics that shape this interplay at the pore scale has received little attention. Specifically, continuum scale models rarely consider the effect of multiphase flow parameters on mineral reaction rates or apply simple corrections as a function of the reactive surface area or saturation of the aqueous phase, without developing a mechanistic understanding of the pore-scale dynamics. In this study, we developed a framework that couples the two-phase flow simulator of OpenFOAM (open field operation and manipulation) with the geochemical reaction capability of CrunchTope to examine pore-scale dynamics of two phase flow and their impacts on mineral reaction rates. For our investigations, flat 2D channels and single sine wave channels were used to represent smooth and rough geometries. Calcite dissolution in these channels was quantified with single phase flow and two phase flow at a range of velocities. We observed that the bulk calcite dissolution rates were not only affected by the loss of reactive surface area as it becomes occupied by the non-reactive non-aqueous phase, but also largely influenced by the changes in local velocity profiles, e.g., recirculation zones, due to the presence of the non-aqueous phase. The extent of the changes in reaction rates in the two-phase systems compared to the corresponding single phase system is dependent on the flow rate (i.e., capillary number) and channel geometry, and follows a non-monotonic relationship with respect to aqueous saturation. The pore-scale simulation results highlight the importance of interfacial dynamics in controlling mineral reactions and can be used to better constrain reaction rate descriptions in multiphase continuum scale models. These results also emphasize the need for experimental studies that underpin the development of mechanistic models for multiphase flow in reactive systems.

Keywords: multiphase reactive transport, pore-scale, reaction rate, gas bubble, roughness

Edited by:
Harrie-Jan Hendricks Franssen,
Helmholtz Association of German Research Centres (HZ), Germany

Reviewed by:
Julien Maes,
Heriot-Watt University,
United Kingdom
James E. McClure,
Virginia Tech, United States

****Correspondence:***
Hang Deng
hangdeng@lbl.gov;
hangdeng310@gmail.com

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *01 July 2021*
Accepted: *29 November 2021*
Published: *06 January 2022*

Citation:
Li P, Deng H and Molins S (2022) The Effect of Pore-Scale Two-Phase Flow on Mineral Reaction Rates. Front. Water 3:734518. doi: 10.3389/frwa.2021.734518

INTRODUCTION

Interactions between multiphase flow, geochemical reactions, and solute transport in fractured porous media are ubiquitous in Earth's critical zone and subsurface systems, and affect the dynamics of many environmental and energy engineering applications. Examples include supercritical CO_2 ($scCO_2$) injection, light non-aqueous phase fluid (LNAPL) contamination, and Enhanced Oil Recovery (EOR). In geologic carbon storage systems, the injected $scCO_2$ displaces the native brine and can be trapped in small pores or can dissolve into the brine and react with the host rocks.

The resulting mineral dissolution and precipitation can modify the porosity and permeability of the reservoirs and caprocks, affecting injectivity of the reservoirs and long-term storage security (Johnson et al., 2001, 2004; Xu et al., 2011b). In EOR, low-salinity water is injected to the oil reservoir to improve sweeping efficiency by modifying wettability via surface complexation reactions (Zhang et al., 2006; Kumar et al., 2011). In the shallow subsurface, LNAPL contamination of groundwater is a widespread environmental problem. Light non-aqueous phase fluid is insoluble in water and typically fluctuates with the water table. As a result, LNAPL is spread vertically following local drainage-imbibition cycles, which in turn affects the degradation and thus the fate of the contaminants (Ngien et al., 2012; Pan et al., 2016; Govindarajan et al., 2018). The interactions between two-phase flow and electrochemical reactions have also been identified as a research priority for the design of effective fuel cells, as gas bubbles (of e.g., H_2 and O_2) can be generated by side reactions during charging and affect power generation (Chen et al., 2017; Grunewald et al., 2021).

Continuum-scale numerical models are widely used to investigate and predict the evolution of fractured porous media caused by multiphase reactive transport processes. Models such as TOUGHREACT, OpenGeoSys, and PFLOTRAN are versatile tools that have been used to investigate $scCO_2$ migration in fractured formations (Xu et al., 2011a, 2019; Xiao et al., 2020), dissolution, transport and biodegradation of non-aqueous phase fluid (NAPL) in shallow aquifers (Pruess, 2004; Popp et al., 2015; Sookhak Lari et al., 2019), and heat extraction using CO_2 as a working fluid (Lichtner and Karra, 2014). These models however rely heavily on constitutive relations such as the Brooks and Corey equation and the van Genuchten equation that relates the water saturation with capillary pressure and relative permeability.

Pore-scale structural heterogeneity and processes that are not considered explicitly in the continuum description, however, could be important. For instance, microfluidic experiments from Karadimitriou et al. (2016, 2017) examined non-Fickian transport in a water-Fluorinert immiscible fluid system. The results demonstrated that the mass transfer rate between mobile–immobile zones under the two-phase flow condition is not constant, indicating that the conventional treatment of the mass transfer rate in the continuum scale Mobile–Immobile (MIM) model could introduce significant errors in the simulation. Pore-scale experiments from Jiménez-Martínez et al. (2015) using 2D microfluidic cells with homogeneous pore structures have also demonstrated that solute mixing can be significantly enhanced under multiphase flow conditions because of ramified finger flow structure, non-Fickian dispersion, and non-wetting phase clusters that limit the finger flow merging. Recent advancement in experimental techniques have also enabled direct observations of reactive transport in multiphase systems at the pore scale. Using biogenically calcite-functionalized micromodels, Song et al. (2018) observed a new microscale mechanism that affects reactive transport in the CO_2-brine-calcite system. The CO_2 gas phase can accumulate on the mineral surfaces following calcite dissolution, which protects calcite from further dissolution. Jiménez-Martínez et al. (2020) developed a high-pressure geomaterial microfluidic device and investigated mineral reactions in etched channels during co-injection of CO_2-saturated brine and $scCO_2$. They observed that the presence of $scCO_2$ bubbles significantly changed both the flow dynamics and the reaction patterns, compared to the single phase flow experiments. In the competing channels with different widths, mineral dissolution was more homogenized and carbonate precipitation was enhanced in the low-velocity regions formed as a result of the presence of the bubbles, which was also confirmed by pore-scale Lattice Boltzmann Methods (LBM) simulations.

Pore-scale models provide an invaluable tool to further our understanding of pore-scale dynamics. Pore-network models (PNMs) are a computationally efficient option for simulating the reactive transport processes under two-phase flow conditions, but require simplifications of the geometric solid-fluid interface and assumptions about uniform aqueous concentrations within a pore (Xiong et al., 2016). Lattice Boltzmann Methods (Chen et al., 2013, 2015) and direct numerical simulations (DNS) (Haroun et al., 2010b; Marschall et al., 2012; Maes and Soulaine, 2018; Soulaine et al., 2018, 2021) provide alternatives that relax these restrictions. The LBM was able to reproduce experimental observations on wettability alteration during Low Salinity Water Flooding (Akai et al., 2020). It was also shown by Maes and Geiger (2018) using a DNS approach that the alternation is driven by the concentration of the potential determining ions (PDI) instead of pH. Moreover, a series of micro-continuum simulations showed that the production of CO_2 bubbles resulting from carbonate dissolution may limit the subsequent dissolution and prevent the emergence of wormholes (Soulaine et al., 2018).

In spite of the growing use of fully-resolved pore-scale models (LBM or DNS) in multiphase reactive applications, they have not been used to understand how reaction rates are influenced by flow processes in the same way as it has been done in single-phase systems (e.g., Molins et al., 2012, 2014; Deng et al., 2018). In single phase systems, reactive surface area has been typically varied to account for flow dynamics and transport limitations. Pore-scale simulations have highlighted the complex dependence of the correction factor on pore-scale geometry and flow regimes (Deng et al., 2018). In multiphase systems, however, this information does not exist. As a result, multiphase continuum scale models rarely account for the effect of the flow dynamics on reaction rates. Further, reaction rates are commonly assumed to be independent of phase saturations (e.g., Xu et al., 2011a; Lichtner et al., 2015; Águila et al., 2020; Wu et al., 2021). Only in some rare instances, the reactive surface area is allowed to vary with liquid saturation. For example, in the active fracture module implemented in TOUGHREACT, the reactive surface area and thus reaction rate follows a power law relation with respect to the water saturation and the exponent is an empirical variable (Sonnenthal et al., 2005). However, there is still a lack of pore-scale studies and mechanistic understanding that support the development of this type of constitutive relationship for considering the impacts of multiphase flow on reaction rates.

Our study aims to bridge this gap, by performing a series of well-designed pore-scale multiphase reactive transport simulations. To this end, a pore-scale multiphase reactive transport modeling framework was developed by coupling

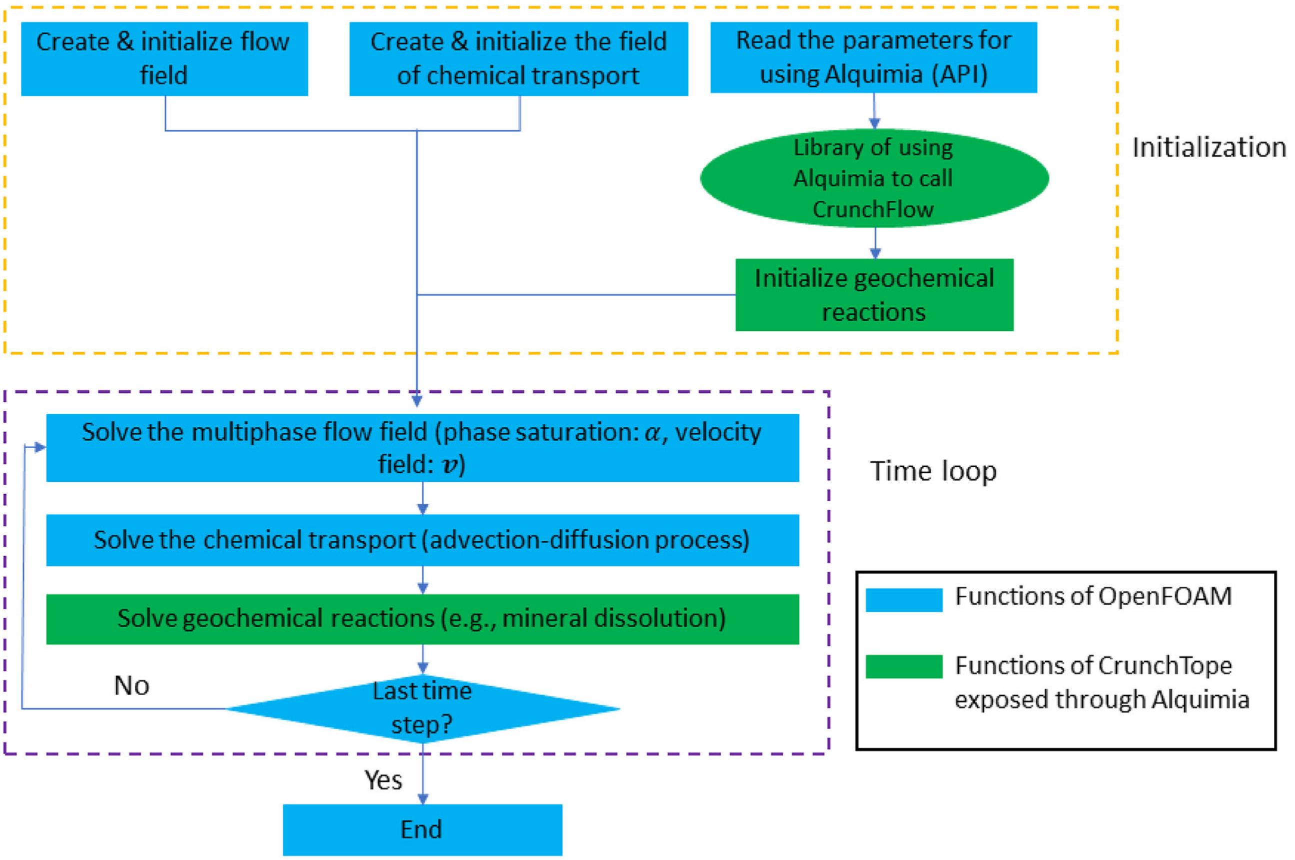


FIGURE 1 | Workflow of CrunchFOAM.

OpenFOAM and CrunchTope. A set of simulations were performed with co-injection of air and CO_2-acidified water into 2D calcite channels to examine calcite dissolution rate under a range of flow conditions. A sine wave geometry was used to introduce different levels of roughness in the channels. This simple geometry has been widely used to provide an idealized representation of surface roughness to investigate its impacts on fluid flow and chemical transport (Kitanidis and Dykaar, 1997; Bolster et al., 2009; Sund et al., 2015; Deng et al., 2018). Section Methodology details the modeling framework, the mathematical principles, and the simulation setups. The results of the numerical simulations, including analyses of the flow field and reaction rates, and observed relations between reaction rate and parameters such as saturation are presented in section Results. We discuss the broader implications of our study in section Discussion and conclude in section Summary and Conclusions.

METHODOLOGY

Our approach entails the simulation of two-phase flow and reactive transport in a series of synthetic geometries. For this purpose, we develop a modeling framework by coupling two widely used and thoroughly validated codes. In this section, we present this modeling framework and we describe the setup of the simulations.

Modeling Framework

This pore-scale multiphase reactive transport modeling framework couples the geochemical reaction solver CrunchTope (Steefel et al., 2015) with the open source software package, OpenFOAM (open field operation and manipulation, OpenFOAM-v1812) and the related open source libraries, using Alquimia. Alquimia is a generic interface, which allows any flow and transport simulator to access geochemical reaction functionalities of existing, thoroughly validated codes such as CrunchTope (Andre et al., 2013). The modeling framework is referred to as CrunchFOAM for short.

Figure 1 illustrates the workflow of the modeling framework. The geochemical conditions and reaction kinetics are specified in the CrunchTope input files. The initial and boundary conditions for the flow and the transport of the primary chemical species are specified in OpenFOAM. Two-phase flow, transport, and geochemical reactions are solved sequentially following the operator splitting approach. The time stepping is controlled by the flow solver in OpenFOAM. Within each time step, the flow field from the two-phase flow solver is passed to the transport solver to calculate the concentrations of the primary species by solving the advection-diffusion equation. Afterwards, CrunchTope is called in each cell to calculate aqueous speciation and mineral reactions, and to update the concentration fields for transport in the next time step.

Flow

Two phase flow is solved using interFoam, the standard OpenFOAM solver for transient incompressible isothermal flow of two immiscible fluids. The solver implements a modified version of the Volume of Fluid (VoF) method, and its performance has been confirmed for capillary numbers larger than 10^{-5} (Deshpande et al., 2012; Shuard et al., 2016).

The VoF approach treats the two fluid phases as an effective single phase. The velocity and pressure fields are solved by the single-field incompressible Navier Stokes Equation and continuity equation (Hirt and Nichols, 1981).

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}) = -\nabla p + [\nabla \cdot (\mu(\nabla \mathbf{u} + \nabla \mathbf{u}^T))] + \rho g + \mathbf{F}_{st} \tag{1}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{2}$$

where $\mathbf{u}$ is the velocity. Fluid density ρ, and viscosity μ are weighted averages of the two fluid phases based on the volume fraction (α) of a designated fluid, which is usually the wetting fluid

$$\rho = \rho_1 \alpha + \rho_2 (1 - \alpha) \tag{3}$$

$$\mu = \mu_1 \alpha + \mu_2 (1 - \alpha) \tag{4}$$

$\mathbf{F}_{\mathbf{st}}$ is the surface tension force and defined as

$$\mathbf{F}_{st} = \gamma \kappa \mathbf{n} \hat{\delta} \tag{5}$$

where γ is the interfacial tension, $\kappa = \nabla \cdot \mathbf{n}$ is the interface curvature, $\mathbf{n}$ is the unit vector normal to the interface given by $\frac{\nabla\alpha}{||\alpha||}$, $\hat{\delta}$ is a Dirac function located on the interface.

The phase volume fraction α is solved by the following transport equation

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{u}) + \nabla \cdot (\alpha(1 - \alpha)\mathbf{u}_r) = 0 \tag{6}$$

where $\mathbf{u}_r$ is the relative velocity between the two fluids/phases. It is typically defined as the compression velocity $\mathbf{u}_c$ to ensure a sharp interface, and its amplitude is determined by the maximum of the single-field velocity

$$\mathbf{u}_r \equiv \mathbf{u}_c \equiv \mathbf{n} \left[\min(c_\alpha \frac{|\Phi|}{A_f}, \max\left(\frac{|\Phi|}{A_f}\right)) \right] \tag{7}$$

where Φ is the volumetric flux, A_f is the cell surface area, $0 \leq c_\alpha \leq 1$ limits the compression velocity below the maximum face flux velocity $\frac{|\Phi|}{A_f}$ and is a user-specified coefficient ($c_\alpha = 1$ in our simulations). This formulation helps minimize numerical diffusion (Rusche, 2002).

The contact angle (θ) is defined at the solid boundary and the following equation needs to be satisfied (Aziz et al., 2018):

$$\mathbf{n} \cdot \mathbf{n}_s = cos\theta \tag{8}$$

where $\mathbf{n}_s$ is the normal vector to the solid wall.

Transport

For transport under multiphase flow conditions, the continuous species transfer (CST) method has been developed and implemented as a third-party solver in OpenFOAM (Haroun et al., 2010a; Marschall et al., 2012; Deising et al., 2016). The C-CST (compressive-CST) algorithm developed by (Maes and Soulaine, 2018) was implemented in this work as it minimizes numerical diffusion near the interface and ensures that the description of advection is fully consistent with the phase evolution equation (VoF equation). In this method, the transport of a species j dissolved in both phases is described by

$$\frac{\partial C_j}{\partial t} + \nabla \cdot (\mathbf{u} C_j) = -\nabla \cdot \left(\frac{(1 - H_j)\, C_j}{\alpha + (1 - \alpha) H_j} \alpha (1 - \alpha) \mathbf{u}_r \right) + \nabla \cdot \left(\hat{D}_j \nabla C_j + \Psi_j \right) + R_j \tag{9}$$

where $\hat{D}_j$ is the interpolation of the diffusion coefficient of the chemical species in the two phases

$$\hat{D}_j = \alpha D_{j,w} + (1 - \alpha) D_{j,nw} \tag{10}$$

where $D_{j,w}$ and $D_{j,nw}$ is the diffusion coefficient of chemical j in the wetting and non-wetting fluid, respectively. Ψ_i describes the concentration jump at the interface

$$\Psi_j = -\hat{D}_j \frac{1 - H_J}{\alpha + (1 - \alpha) H_j} C_j \nabla \alpha \tag{11}$$

where H_j is Henry's law constant. In this study, mass transfer across the interface is not considered, and H_j is set to be a small value (1×10^{-12}) to avoid zero denominators in Equation (9) when $\alpha = 0$. Because only reactions in the aqueous phase are considered (see below), neglecting mass transfer across the interface implies that concentrations in the non-aqueous phase remain equal to the initial condition (which is zero).

Geochemical Reactions

R_j in Equation (9) accounts for the contribution of mineral reactions to the changes in the mass of chemical species j, and is described by the transition state theory rate law

$$R_j = k_{rxn} A (1 - IAP/K_{eq}) \tag{12}$$

where k_{rxn} is the kinetic coefficient ($mol/m^2 \cdot s$), A is the surface area of the mineral phase (m^2), and is determined directly from the geometry in OpenFOAM, and the chemical affinity term is calculated from the ion activity product (IAP) and the equilibrium constant of the mineral reaction (K_{eq}). The aqueous reactions are assumed to reach equilibrium instantaneously and speciation is calculated based on the law of mass action and the concentrations of the primary species.

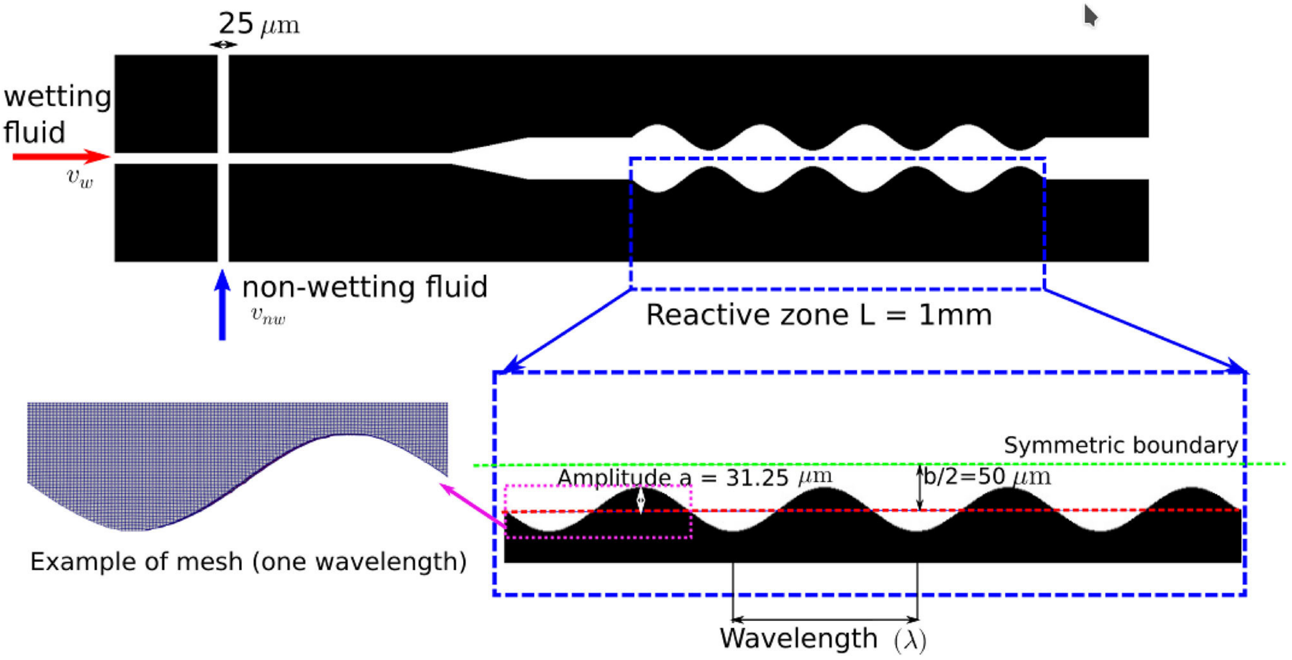


FIGURE 2 | An illustration of the geometry and mesh used in our numerical simulations. Here, the reactive zone has 4 full sine waves and the mesh resolution is 2 μm.

TABLE 1 | Summary of the geometries used in the simulations.

	Geometry a	**Geometry b**	**Geometry c**	**Geometry d**	**Geometry e**	**Geometry f**
Wavelength (λ)	N/A	1L	1/4L	1/8L	1/16L	1/32L
SRF	1	1.08	1.14	1.46	2.31	4.19

Simulation Setup

We investigate multiphase flow and reactive transport in a domain that seeks to represent a microcrack with an arbitrarily rough geometry, which can also be conceptualized as a sequence of pores. **Figure 2** provides an illustration of the geometry used in our simulations. To save computational time, we assumed a symmetric geometry and simulated half of the domain. A T-junction structure is used at the inlet for the injection of the wetting and non-wetting phase. The width of the inlets is 25 μm. The reactive zone, where the reactive mineral is located and highlighted in the blue box in **Figure 2**, is 1mm long and has an average width (b) of 100 μm. The dimensions are comparable to previous modeling and micromodel studies of geomaterials (Deng et al., 2018; Song et al., 2018; Jiménez-Martínez et al., 2020) and to the fiber diameter in batteries (Chen et al., 2017). In addition to the reference flat channel, a single sine wave was used to represent pore scale roughness in the reactive zone, following (Deng et al., 2018).

$$b(x) = \frac{\overline{b}}{2} + a \cdot sin\left(\frac{2\pi x}{\lambda}\right) \tag{13}$$

where λ is the wavelength, and a is the amplitude. In order to explore different levels of roughness, a number of simulations were performed each with a different wavelength (**Table 1**). In all cases, an amplitude of 31.25 μm was used.

The roughness in these sine wave geometries is measured by the surface roughness factor (SRF). It is defined as the ratio between the total surface area (A_{total}), which is calculated by summing the patch area defined as the mineral wall in the mesh generated by OpenFOAM, and the nominal surface area, which is equivalent of the surface area of the flat geometry (A_{flat}).

$$SRF = \frac{A_{total}}{A_{flat}} \tag{14}$$

The geometries were first generated by a Python script and Blender, and the STL files were then imported into OpenFOAM to generate the meshes using snappyHexMesh. The average mesh size was set to 2 μm. Although interFoam does not show convergence with decreasing mesh size (Pavuluri et al., 2018), this mesh size is comparable with the resolution used in previous studies that showed good results using interFoam in complicated pore structures (Yin et al., 2019; Carrillo et al., 2020). This mesh size also ensures that the single phase simulation results are not affected by further refinement.

Initially, the simulation domain is fully saturated with water except for the vertical branch of the T-junction, which is occupied

TABLE 2 | Physical properties of fluids (*ρw, water density; ρa, air density; μw, dynamic viscosity of water; μa, dynamic viscosity of air; γwa, interfacial tension between water and air; θ,contact angle*).

ρ_w (kg/m³)	ρ_a (kg/m³)	μ_w (Pa · s)	μ_a (Pa · s)	γ_{wa} (N/m)	θ
1000.0	1.0	1.05×10^{-3}	1.55×10^{-5}	0.072	30°

TABLE 3 | Boundary conditions for the two-phase flow simulations.

	Injected wetting fluid—inlet	Injected non-wetting fluid—inlet	Outlet
Flow	$v_{inw} = 0.1m/s$		Zero-gradient
	$v_{inw} = 0.04m/s$	1.$v_{innw} = 0.25v_{inw}$	Zero-gradient
	$v_{inw} = 0.4m/s$	2.$v_{innw} = 0.5v_{inw}$	Zero-gradient
Transport	Constant concentration	Constant concentration	Zero-gradient

TABLE 4 | Calcite dissolution reactions, the equilibrium constant, and the three reaction pathways with the kinetic coefficients.

Dissolution reaction	LogK$_{eq}$
$CaCO_3(s) + 2H^+ = Ca^{2+} + H_2CO_3^*(aq)$	8.16
Elementary reaction	**Logk(mol/m²·s)**
$CaCO_3(s) + H^+ \Leftrightarrow Ca^{2+} + HCO_3^-$	$k_1 = -0.05$
$CaCO_3(s) + H_2CO_3^* \Leftrightarrow Ca^{2+} + 2HCO_3^-$	$k_2 = -3.3$
$CaCO_3(s) \Leftrightarrow CO_3^{2-} + Ca^{2+}$	$k_3 = -6.19$

TABLE 5 | Aqueous reactions and the equilibrium constants.

Aqueous complexation reaction	LogK_{eq}
$OH^- + H^+ \rightleftharpoons H_2O$	13.99
$CO_3^{2-} + 2H^+ \leftrightharpoons H_2CO_3^*(aq)$	16.67
$HCO_3^- + H^+ \rightleftharpoons H_2CO_3^*(aq)$	6.34
$CaCO_3(aq) + 2H^+ \rightleftharpoons Ca^{2+} + H_2CO_3^*(aq)$	13.35
$CaHCO_3^- + H^+ \leftrightharpoons H_2CO_3^*\ (aq) + Ca^{2+}$	5.30
$CaOH^+ + H^+ \leftrightharpoons Ca^{2+} + H_2O$	12.85
$CaCl^+ \leftrightharpoons Ca^{2+} + Cl^-$	0.7
$CaCl_2(aq) \leftrightharpoons Ca^{2+} + 2Cl^-$	0.65
$NaCO_3^- + 2H^+ \leftrightharpoons H_2CO_3^*(aq) + Na^+$	16.16
$NaCl(aq) \leftrightharpoons Na^+ + Cl^-$	0.78
$NaHCO_3^-(aq) + H^+ \leftrightharpoons H_2CO_3^*(aq) + Na^+$	6.18

by the non-wetting phase (i.e., air). The physical properties of the fluids are summarized in **Table 2**.

A constant velocity boundary condition is applied at the inlets. For each geometry, three velocities for the wetting phase were simulated, which are 0.04, 0.1, and 0.4 m/s. The velocity of the non-wetting phase is a fraction of the wetting phase, and two ratios, 0.25 and 0.5, were used to control the frequency of the bubbles generated by the co-injection. This results in a Reynolds number (Re) of 1-10 and a capillary number (Ca) of $\sim 10^{-4}$-10^{-3}. The Ca-values are within the range that is relevant for typical reservoirs (Satter and Iqbal, 2016) and battery systems (Grunewald et al., 2021). At the outlet, the zero gradient boundary condition is applied for both flow and transport (**Table 3**).

Geochemical reactions are assumed to take place in the aqueous phase only. The solid phase in the reactive zone is composed of a single mineral, calcite. It dissolves in water, which has a NaCl concentration of 0.01 M/L and a pH of 5 due to dissolution of atmospheric CO_2. The kinetic coefficients for the three elementary reaction pathways reported in Chou et al. (1989) for calcite dissolution are summarized in **Table 4** and used in the simulations. The aqueous reactions and their equilibrium constants are summarized in **Table 5**. The activity coefficients used to convert concentrations to activities ($a_{species}$) are calculated using the extended Debye-Hückle equation.

$$k_{rxn} = k_1 a_{H^+} + k_2 a_{H_2CO_3} + k_3 \tag{15}$$

In this study, the geometry is not updated and we focus on steady state behavior. Given the time scale that is needed to reach steady state (within seconds as shown in section Results), the mineral reaction is not expected to cause any geometric change. The simulations were run until both the flow and reaction rate reached a steady state. For the analyses of calcite dissolution rate, the absolute average instantaneous reaction rate at a time point t ($\left|\overline{R_m^t}\right|$) was calculated for both single-phase ($m = s$) and two-phase ($m = t$) flow systems as

$$\left|\overline{R_m^t}\right| = \frac{\Sigma_{n=1}|R_{m,n}^t| V_n {\alpha_n}^t}{\Sigma_{n=1} A_n {\beta_n}^t} \tag{16}$$

$|R_{m,n}^t|$ is the instantaneous reaction rate in the wall grid cell n with a volume of V_n and wall surface area of A_n at time point t, ${\alpha_n}^t$ is the volume fraction of the wetting phase, i.e., saturation, in the grid cell, and ${\beta_n}^t$ is the wetted surface area ratio within the local wall grid cell, which is approximated by ${\alpha_n}^t$.

In addition to the multiphase simulations described above, a single-phase simulation is performed for each roughness level (**Table 1**) and flow condition (**Table 3**), with the same total flux for comparison. These single-phase results are used to elucidate the multiphase effects on reaction rates. All the simulations were performed with the reactive transport solver described in Section Methodology and the saturation (α) is set to one for single-phase simulations.

RESULTS

Reaction Rates

The co-injection of the wetting and non-wetting phases produces a series of gas bubbles that migrate through the reactive zone (**Figure 3**). The size and frequency of the gas bubbles are primarily controlled by the injection rate and the ratio between the wetting and non-wetting phase. Consistent with previous studies (van Steijn et al., 2010; Malekzadeh and Roohi, 2015; Mi et al., 2019), we observed that the bubble size is controlled by the ratio (v_{nw}/v_w) and the frequency is determined by magnitudes

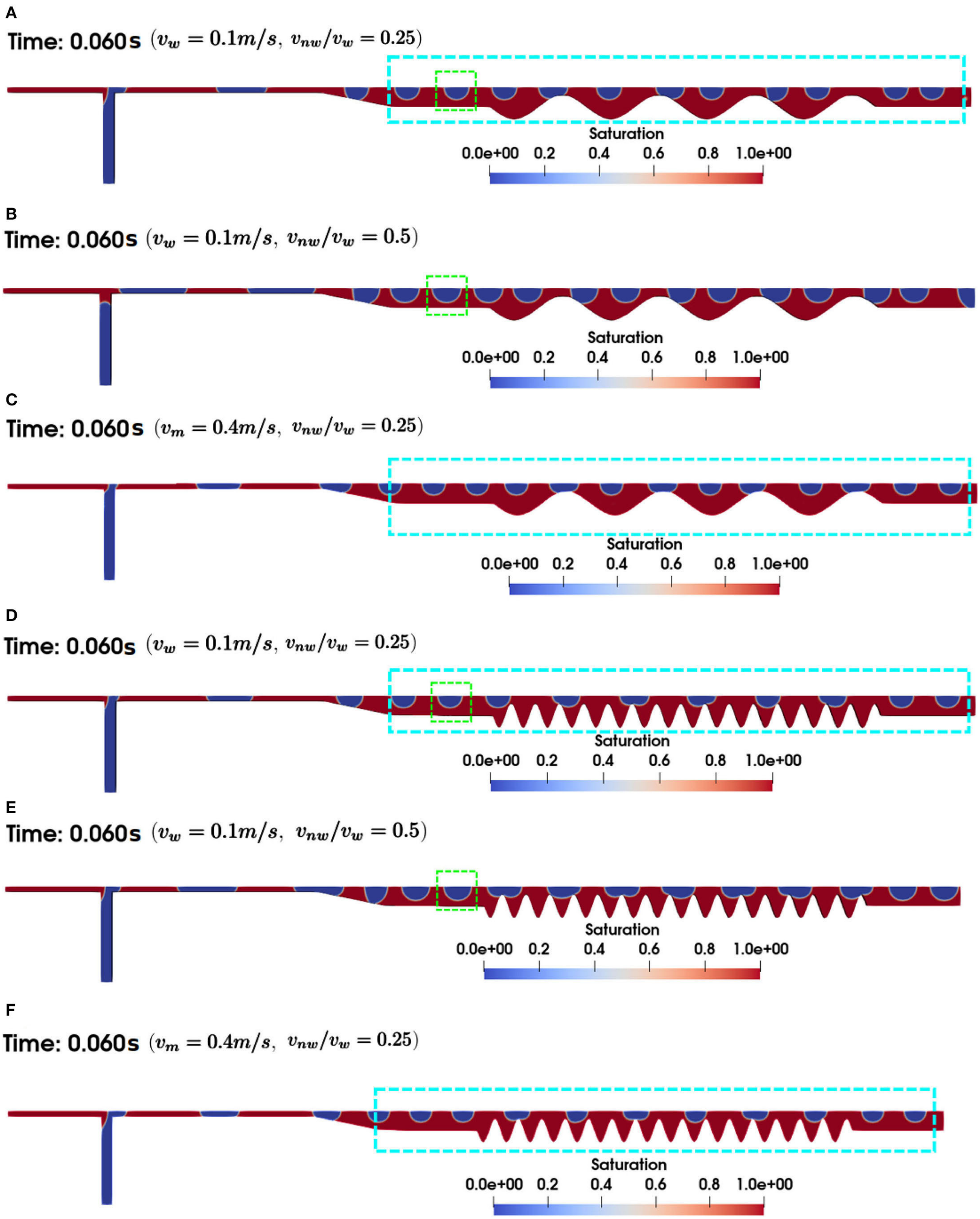


FIGURE 3 | Established flow of gas bubbles for **(A)** geometry c, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.25$; **(B)** geometry c, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.5$; **(C)** geometry c, $v_{inw} = 0.4m/s$, $v_{inw}/v_{innw} = 0.25$; **(D)** geometry e, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.25$; **(E)** geometry e, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.5$; **(F)** geometry e, $v_{inw} = 0.4m/s$, $v_{inw}/v_{innw} = 0.25$. The blue box is used to guide the comparison of the number of bubbles, i.e., bubble frequency. The green box highlights a single bubble to illustrate the dependence of bubble size on the injection ratio.

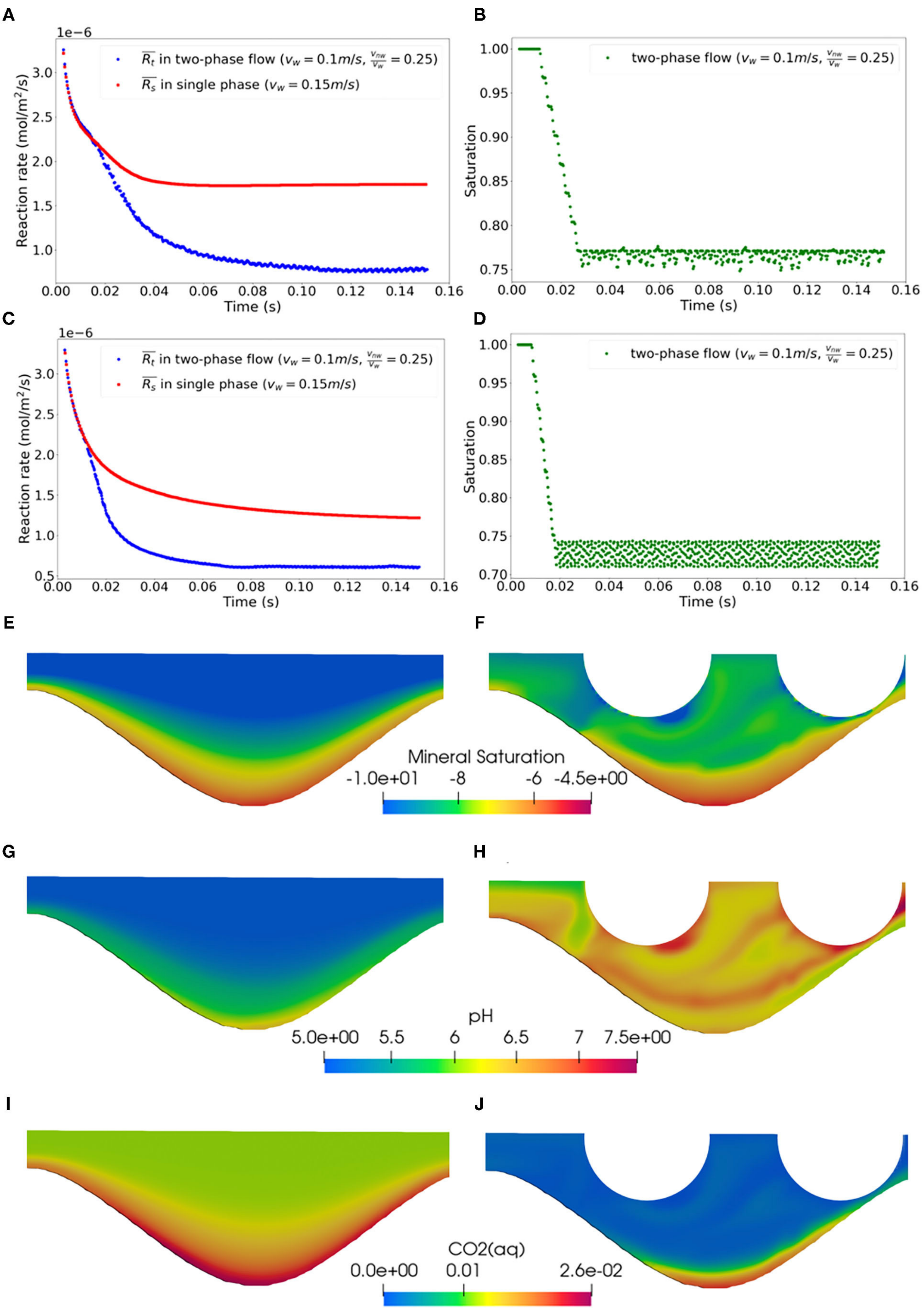


FIGURE 4 | Temporal profiles of the average instantaneous reaction rates **(A,C)** and average saturation **(B,D)** in the reactive zone for geometry c **(A,B)** and e **(C,D)**, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.25$. Close-up spatial profiles of calcite saturation index ($\log(IAP/K_{eq})$) **(E,F)**, pH **(G,H)**, and total CO_2(aq) concentration **(I,J)** in geometry c for the **(E,G,I)** single and **(F,H,J)** two phase simulations at time 0.06s.

of v_{inw} and v_{innw}. The diameter of the bubbles is 64 and 74 µm for v_{nw}/v_w of 0.25 and 0.5, respectively. In cases of large roughness, the shape of the gas bubbles changes temporarily at the narrow throats and can be recovered after entering the wide channel locations, i.e., sine wave troughs. Once the gas bubble flow is established, the hydrodynamics of the system reaches the steady state, as indicated by the average saturation, which is the mean of the phase field (α) in the reactive zone. The average saturation displays small oscillations around the steady state value when the bubbles enter or exit the reactive domain (**Figures 4B,D**). **Figure 4A** shows the temporal profiles of the average instantaneous reaction rates for the single-phase and two-phase cases for geometry **c** at $v_w = 0.1\text{m/s}$ with $v_w/v_{nw} = 0.25$. The reaction rate for the single phase flow simulation decreases initially as the calcite saturation state starts to increase following the dissolution reaction, and stabilizes at about 1.8×10^{-6} mol/m^{2}s after ∼0.04 s. This dissolution rate is comparable with the value reported in previous single phase reactive transport simulations in a similar system using the pore-scale code ChomboCrunch (Deng et al., 2018). Given the relatively high velocity and large Reynolds number, overall the system is far from equilibrium with respect to calcite and the reaction rate is relatively high. The concentration of the reactive solutes and thus calcite saturation index shows a thin boundary layer at the fluid-solid interface and the concentration gradient across the flow direction is large (**Figures 4E,F**). This is consistent with the observations in Deng et al. (2018) that the effect of transverse transport could be important in these cases.

The reaction rate for the two phase flow simulation decreases over time as well. The initial decreasing trend overlaps with the single phase simulation as the gas bubbles have not entered the reactive zone. The decreasing trend diverges as the gas bubbles enter the reactive zone. $\left|\overline{R_m^t}\right|$ reaches the steady state at ∼0.8 $\times$ 10^{-6} mol/m^{2}s after ∼0.12 s. The steady state instantaneous reaction rate of the two phase flow case is significantly lower than that of the single phase flow case. This observation is also consistent across all geometries and flow conditions simulated. **Figures 4C,D** provide another example with similar results for geometry **e** under the same flow condition.

Mechanisms for Reaction Rate Modification in Two Phase Systems

Analysis of local reaction rate confirms that the lower reaction rate in the two phase flow case is partially attributed to the changes in accessibility of rock surface area. **Figure 5A** shows the

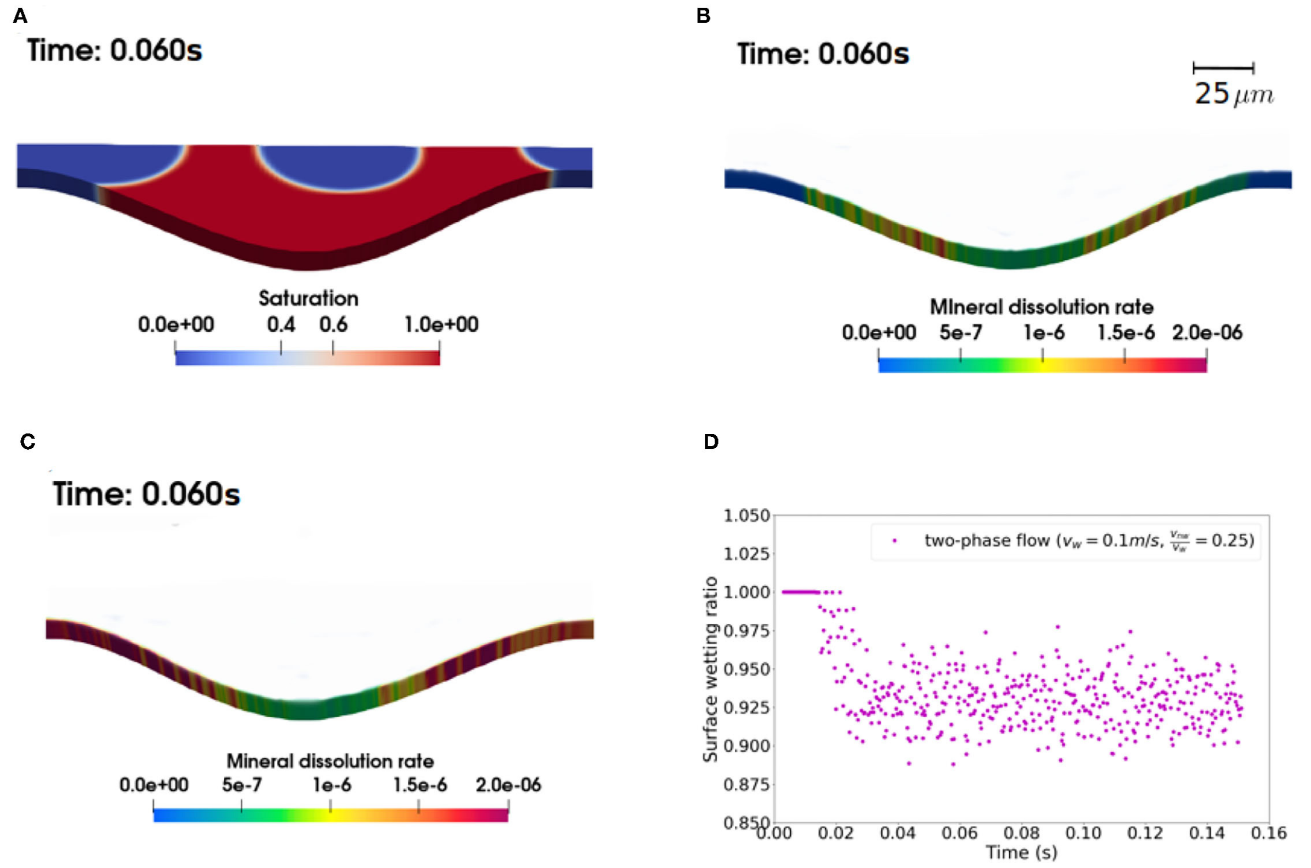


FIGURE 5 | (A) A snapshot of saturation at time 0.06s; local mineral dissolution rates in **(B)** the two-phase and **(C)** single phase simulation at time 0.06s; **(D)** average ratio of wetted surface area within the reactive zone over time, for geometry c, $v_{inw} = 0.1m/s$, $v_{inw}/v_{innw} = 0.25$.

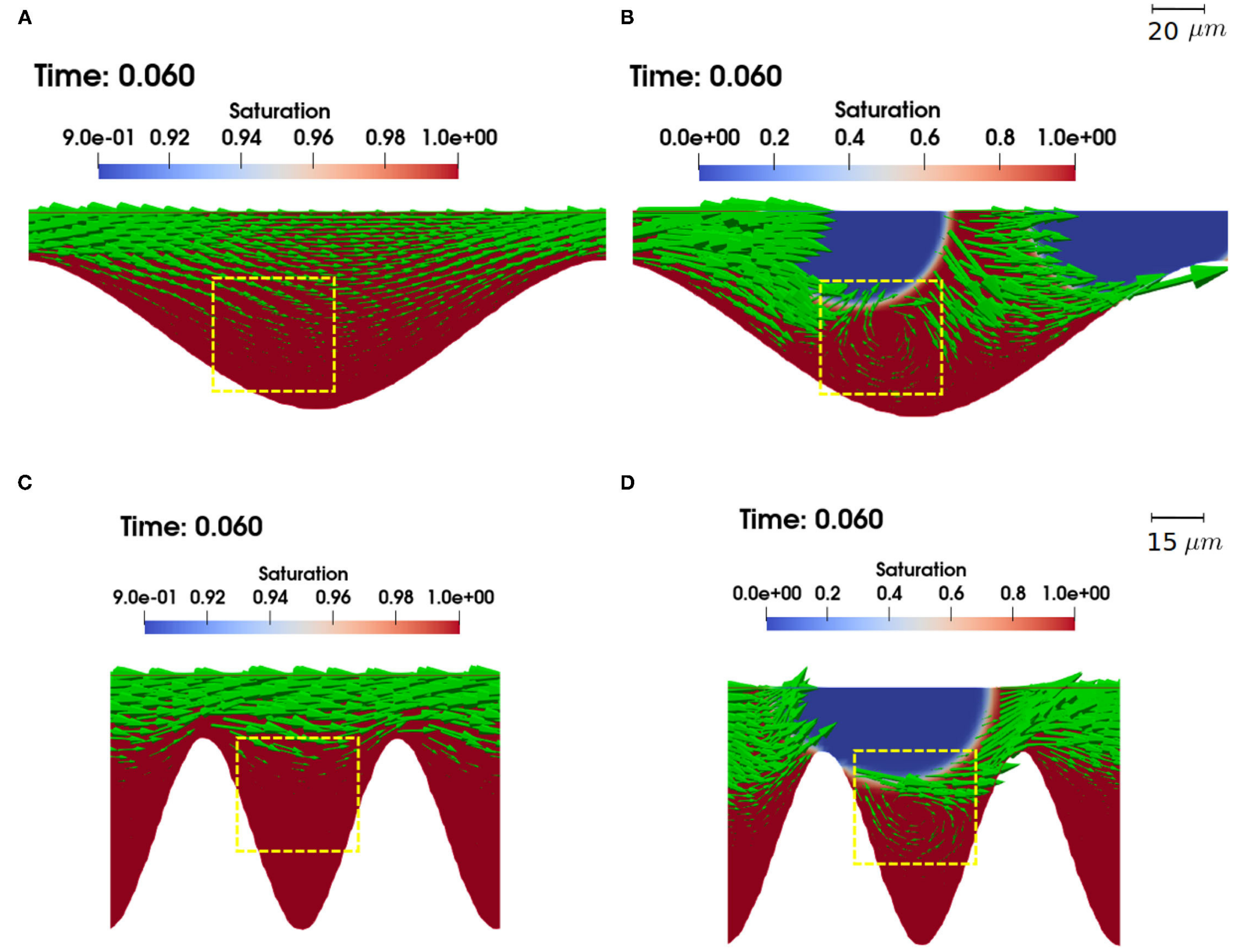


FIGURE 6 | Vector plots showing the flow field in single phase **(A,C)** and two phase flow simulations **(B,D)** of geometry c **(A,B)** and geometry e **(C,D)** with $v_w = 0.1 m/s$, $v_w / v_{nw} = 0.25$. The green/yellow boxes highlight the recirculation zones in the two flow simulations. The size of the arrows is scaled by velocity magnitude.

phase field for geometry **c** at $v_w = 0.1$m/s and $v_w/v_{nw} = 0.25$, and highlights that when the gas bubbles migrate through the narrow throats, the local reactive surface area becomes inaccessible to water-rock interactions (**Figure 5B**). In the experimental study of Song et al. (2018), CO_2 gas bubbles generated by calcite dissolution were observed to block access of the reactive fluid phase to the mineral grain surfaces and thus local calcite dissolution is suppressed. Furthermore, in the corresponding single phase simulation (**Figure 5C**), the local reaction rates on the walls in the narrow throats ($\sim 2 \times 10^{-6}$ mol/m^{2}s) tend to be higher than those in the troughs. As such, even though the wetted surface area in the reactive zone—which provides a measure of accessible mineral surface area—is only reduced by <10% (**Figure 5D**), the reduction in the average reaction rate is much more significant (>50%).

As shown in **Figures 5B,C**, the local reaction rate outside of the narrow throats that are occupied by the non-aqueous phase is also lower in the two-phase case than that in the single phase case. This indicates the presence of a stronger local transport limitation, which is also confirmed by the flow fields. **Figures 6A,B** compare the velocity vectors in the single and two-phase flow simulations for geometry **c** at $v_w = 0.1$m/s and $v_w/v_{nw} = 0.25$. In the two phase case, as the gas bubbles migrate through the troughs, the width of the wetting phase is compressed and recirculation zones formed locally. In contrast, for the same geometry and flow rate, no recirculation zones were observed in the single phase simulation. The recirculating phenomenon has been observed experimentally and numerically in two-phase systems across a wide range of flow conditions with Ca between 1×10^{-7}-1×10^{-2}, as a result of the shear stress exerted by the fluid phase that migrates faster on the other fluid phase that is less mobile or immobile (Blois et al., 2015; Roman et al., 2016; Heshmati and Piri, 2018; Maes and Soulaine, 2018; Mohammadi Alamooti et al., 2020). Previous studies have highlighted that

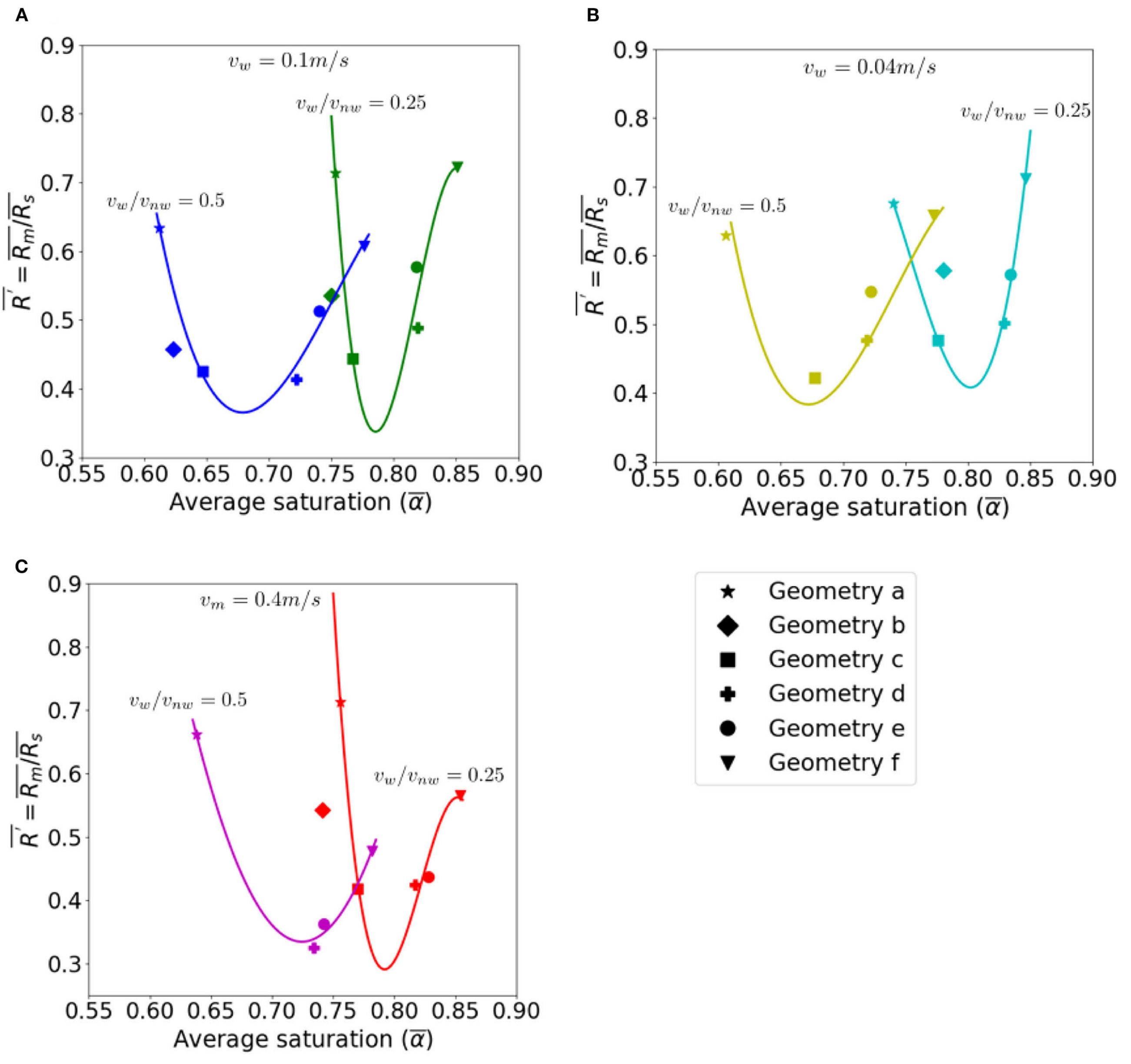


FIGURE 7 | Ratio of the average instantaneous reaction rate between the two-phase and the corresponding single-phase flow simulation (R') plotted against the average saturation (α). The data points are results from the pore-scale simulations and the solid lines are fitted curves using the third degree polynomials for **(A)** $v_w = 0.1m/s$, $v_{nw}/v_w = 0.25$ (green) and $v_{nw}/v_w = 0.5$ (blue); **(B)** $v_w = 0.04m/s$, $v_{nw}/v_w = 0.25$ (cyan) and $v_{nw}/v_w = 0.5$ (yellow); **(C)** $v_w = 0.4m/s$, $v_{nw}/v_w = 0.25$ (red) and $v_{nw}/v_w = 0.5$ (purple). A few data points for geometry b and c at the high v_{nw}/v_w are excluded from the analyses because of bubble merging, which creates hydrodynamics that are not directly comparable to the rest of the simulations.

recirculation zones can be an important mechanism that traps the solutes (Bolster et al., 2014; Sund et al., 2015; Deng et al., 2018; Yoon and Kang, 2021), reducing local thermodynamic driving force (**Figures 4E,F**), i.e., the chemical affinity term in Equation (12). In addition, the pH is higher and total concentration of CO_2(aq) is lower in the two phase case (**Figures 4G–J**), both would lead to a lower kinetic rate as given in Equation (15). **Figures 5C,D** show the flow fields for geometry **e** at the same flow conditions as **Figures 5A,B**. Similar patterns were observed. The recirculation zone is also more predominant in the troughs or the pore-body with the rougher geometry, accounting for a larger portion of the troughs. The dependence of the recirculation zone on pore morphology has also been observed in previous experimental studies (Heshmati and Piri, 2018).

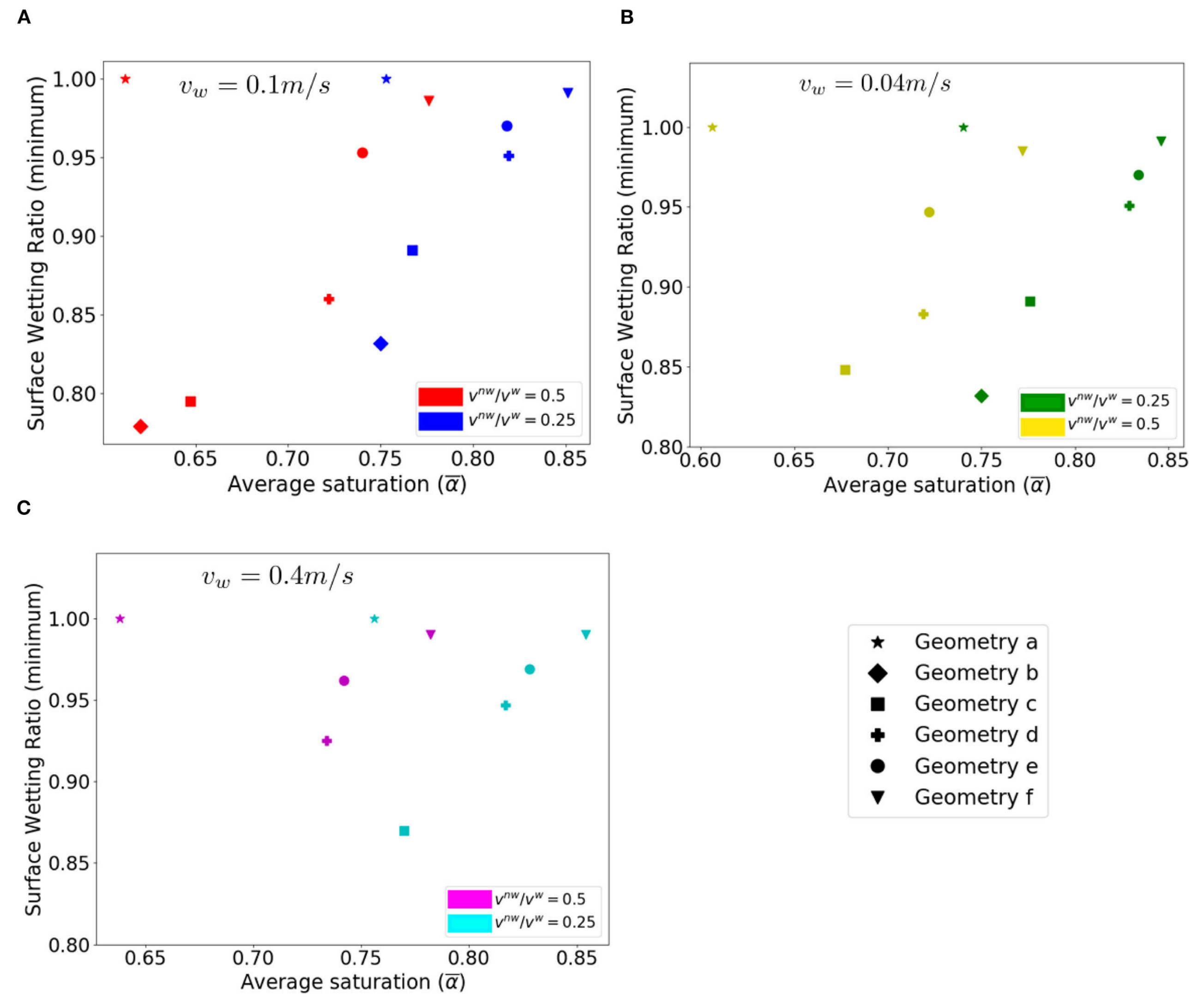


FIGURE 8 | The fraction of the wetted surface area in the reactive zone in relation to the average saturation for different geometries with **(A)** $v_w = 0.1m/s$, $v_{nw}/v_w = 0.25$ (blue) and $v_{nw}/v_w = 0.5$ (red); **(B)** $v_w = 0.04m/s$, $v_{nw}/v_w = 0.25$ (green) and $v_{nw}/v_w = 0.5$ (yellow); **(C)** $v_w = 0.4m/s$, $v_{nw}/v_w = 0.25$ (cyan) and $v_{nw}/v_w = 0.5$ (magenta).

Correlation Between Reaction Rate Modification and Liquid Saturation

In order to gain some insights regarding constitutive relations that can be used to upscale the impacts of two-phase flow dynamics on reaction rate, **Figure 7** summarizes the steady state reaction rates and saturations from the pore scale simulations. The ratio between the reaction rate of the two phase simulation and that of the corresponding single phase simulation is used to evaluate the effect of two-phase flow dynamics, specifically gas bubble migrations in 2D rough channels, on reaction rates. Liquid saturation is an indicator of the two-phase flow dynamics as is the case in many multiphase continuum models.

The impacts on reaction rate do not change monotonically with respect to the steady state saturation. For a given flow condition, the reaction rate ratio decreases as the steady state saturation increases for the less rough geometries (i.e., geometry **a–c**), whereas it increases with the saturation for the rougher geometries (i.e., geometry **d–f**). **Figure 8** shows that the steady state saturation increases with roughness, and so does the wetted surface area ratio (except for the flat reference geometry). This indicates that if the surface area accessibility is the dominant mechanism of the reaction rate reduction in the two phase case, a higher saturation corresponding to a rougher geometry should result in a lower reduction in reaction rate in the two phase case. However, as shown in **Figure 6**, the other mechanism, i.e., the transport limitation of the recirculation zones, is stronger in the rougher geometry. This implies that reaction rate reduction due to transport limitation arising from the presence of the gas bubbles is more significant in the rougher geometries which also have higher saturations. The tradeoff between the

TABLE 6 | Coefficients of polynomial equations and *R*-squared for the fitting curves in **Figure 7**.

	v_w/v_{nw}	A	B	C	D	R-squared
$v_w = 0.1m/s$	0.25	−2740.50	6726.12	−5493.92	1493.94	0.927
	0.5	−208.78	471.67	−351.75	87.10	0.976
$v_w = 0.04m/s$	0.25	845.11	−1911.40	1435.12	−357.09	0.993
	0.5	−256.41	569.40	−417.94	101.97	0.999
$v_w = 0.4m/s$	0.25	−2543.00	6272.58	−5149.03	−1407.47	0.969
	0.5	0	43.82	−63.48	23.33	0.994

two mechanisms can explain the inverted bell shape, which is observed for all flow conditions.

This non-monotonic trend indicates that a power-law relationship will not apply. The guiding lines in **Figure 7** were fitted using third degree polynomials, $\overline{R'} = A\alpha^3 + B\alpha^2 + C\alpha + D$ (**Table 6**). While the polynomial relationship provides a reasonable fit of the pore-scale modeling data statistically—the goodness of fitting as measured by R^2 is larger than 0.95 in most cases (**Table 6**)—they are not meant to be directly implemented or at least caution should be exercised.

Our observation is analogous to previously observed non-monotonic dependence on fluid saturation for other processes in the sense that competing mechanisms—because of their opposite "dependence" on saturation—are in play. For example, Jiménez-Martínez et al. (2017) reported that when saturation is above a threshold, mixing increases as saturation decreases because of stretching, whereas mixing decreases with saturation below the threshold as molecular diffusion becomes dominant. Our observations also reiterate the fact that saturation is a result of the multiphase dynamics and it alone does not provide a full description of the hydrodynamics in the system. For example, for a given flow condition in our simulations, saturation is determined by the roughness of the geometry, which can be expected for other more complex geometries. Given that both liquid saturation and reaction rate are dependent on the flow conditions and the geometries, future studies may focus on developing constitutive relations that are informed by the underlying physics or that explicitly integrate these controlling factors.

DISCUSSION

In our simulations, the recirculation zone is the dominant hydrodynamic feature affecting the mineral dissolution rate. The development of recirculation zones have been observed in single phase systems at high velocity when the contribution of inertial effect becomes significant; and roughness allows recirculation zones to form under lower velocities (Deng et al., 2018). Our results illustrate that in two phase systems, momentum transfer across the fluid–fluid interface further extends the conditions under which recirculation zones form as also reported in previous studies (Heshmati and Piri, 2018). Other multiphase flow dynamics are also observed in our results albeit the simplified setup considered. For instance, in geometry **b**, a few simulations at the high and low flow velocities especially with large v_w/v_{nw} showed coalescence of the gas bubbles, which resulted in significantly lower saturation in the reactive zone. Bubble coalescence and breakup are widely observed phenomena that are dependent on fluid properties, velocities, and geometries (Jo and Revankar, 2009; Paulsen et al., 2014; Chen et al., 2017; Mahabadi et al., 2018; Ren et al., 2020; Grunewald et al., 2021). These processes are accompanied by the re-organization of the fluid–fluid interface and can introduce perturbations in the velocity field. The simulations with gas bubble coalescence were excluded from the analyses in section Correlation Between Reaction Rate Modification and Liquid Saturation as the new hydrodynamics is not directly comparable with the other simulations. Nonetheless, this observation highlights that compared to single phase systems in which a lower velocity typically transfers to a stronger transport limitation, the two phase systems require consideration of additional hydrodynamics that may arise at a different velocity (Blois et al., 2015).

A variety of complex hydrodynamics can arise from the migration of fluid–fluid interfaces depending on the velocity (e.g., capillary numbers) and pore morphology (e.g., Berg et al., 2013; Spurin et al., 2019; Li et al., 2021; Wang et al., 2021).

From a macroscopic perspective, as capillary number (Ca) increases, interface migration transitions from the capillary fingering regime with more random local movement to the viscous fingering with more stable displacement (Toussaint et al., 2012; Li et al., 2019; Grunewald et al., 2021). However, Ca alone does not provide a good description of the fluid–fluid interface dynamics (Armstrong et al., 2015), which is more sensitive to pore-scale roughness/morphology at low capillary numbers (Toussaint et al., 2012).

From a microscopic perspective, both the magnitude and direction of flow were observed to fluctuate before the arrival of the invading front (Roman et al., 2016; Li et al., 2017). Strong instabilities of the interface can also lead to pore-scale burst events such as Haines jumps and can increase local velocities by one–two orders of magnitude (Blois et al., 2015; Li et al., 2017, 2021). These hydrodynamics are reported to promote mixing and thus reactions in the liquid phases (Jiménez-Martínez et al., 2015, 2016, 2017), and may also increase local mineral reaction rates. In the study of Jiménez-Martínez et al. (2020), calcite dissolution rate in the two phase experiment with Ca = $\sim 10^{-5}$ and $v_w/v_{nw} = 1.0$ is 68% of the dissolution rate in the single phase experiment. This is comparable to the simulation results in the smooth channel, e.g., the reaction rate ratio is ~64 and ~72% at $v_w = 0.1$ m/s (i.e., Ca = $\sim 3.5 \times 10^{-4}$) for v_w/v_{nw} of 0.5 and 0.25, respectively (**Figure 7**). Their reaction rate ratio is slightly higher than what would be expected based on our simulations, which is likely because of the oscillation of gas bubbles in the presence of a system of channels and thus higher local velocity at the fluid–fluid interface, in addition to the continuous supply of CO_2 from the co-injected $scCO_2$ phase. The competition of different fluid pathways can result in more dynamic bubble migration and flow field rearrangement, which influences self-organization of the system, which is also illustrated in the column experiment of Ott and Oedai (2015) and the pore-scale numerical simulations of Soulaine et al. (2018). Geometry units such as pore-doublets can be used to capture such dynamics and thus offer additional insights (Mohammadi Alamooti et al., 2020; Alizadeh and Fatemi, 2021). More realistic geometries may be needed to fully examine the complexity of real systems, as flow instabilities are primarily driven by geometry when morphological heterogeneity is large (Li et al., 2017; Heshmati and Piri, 2018).

Water film is also an important contributor to the fluid–fluid interface (Li et al., 2019). While for the given Ca and channel width used in our study, water film is not expected to be well developed and thus contribute to transport significantly or affect our analyses (Roman et al., 2017), well-developed water film may become important in maintaining water-rock contact and transport pathways between isolated water parcels. For instance, water film has been observed during the drainage process in various experiments (Rücker et al., 2015; Schlüter et al., 2016; Roman et al., 2017; Moura et al., 2019). It enhances the connectivity between residual water and may cause the "snap-off" phenomenon, affecting the contact between water and the solid phase (e.g., rock) and the transport processes.

Overall, this study represents an early-stage effort in highlighting the importance of a dynamic coupling between reaction rates and multiphase flow dynamics. As such, it focuses on a specific geometric setup and a relatively limited number of flow scenarios. However, compared to the experimental studies that typically report temporally and/or spatially integrated reaction rates, our model provides information on local and instantaneous reaction rate along with detailed velocity field and solute transport, which helps to better explore microscopic mechanisms. The two mechanisms identified in our pore-scale two-phase reactive transport simulations—the surface area effect and local hydrodynamics—support recent observations from micromodel experiments (Song et al., 2018; Jiménez-Martínez et al., 2020). Moreover, the pore-scale perspective in the model enables insights into the impacts of multiphase flow dynamics on mineral reaction rate that have broader implications. Namely, results indicate that in addition to the fluid-solid interface, a direct measure of the reactive surface area, the fluid–fluid interface also plays an important role in controlling solid-phase reactions by modifying local hydrodynamics. As discussed above, local hydrodynamics that can affect mineral reaction rate is largely influenced by fluid–fluid interfaces, our study suggests that adding interfacial area in the formulation of reactive surface area in two-phase flow conditions in addition to saturation may be a logical step. In fact, it has been proposed to use the fluid–fluid interface as a measure of flow topology and thus an additional parameter in constitutive relations of relative permeability (Picchi and Battiato, 2018). Such consideration may be of particular interest in e.g., scCO_2-brine systems where mass transfer across the fluid–fluid interface also affect fluid chemistry and thus mineral reaction rate.

In order to develop constitutive relations that faithfully reflect the coupling between mineral reaction rates and multiphase flow dynamics and are broadly applicable in natural and engineered fractured porous materials, systematic studies that explore a broader range of multiphase flow and geometric conditions are needed. While the modeling framework developed in this work can be adapted to consider these processes and thus provides a suitable modeling tool in a wide range of applications for pore-scale mechanistic investigations, it needs to be acknowledged that these investigations hinge upon further development of modeling capabilities for multiphase flow dynamics (Aboukhedr et al., 2018; Qin et al., 2020). For instance, there is still a lack of comprehensive comparison and benchmarking of pore-scale models for low Ca flow systems (Zhao et al., 2019). It has also been shown that because of the sensitivity of multiphase flow to local perturbations, a deterministic reproduction of the dynamics as observed in experiments with relatively large porous domains using numerical models is challenging, while statistical behaviors can still be captured (Ferrari et al., 2015). The consideration of reactions, however, may require higher spatial accuracy because the spatial variations (in mineral distribution and reaction rate) can be important. The impacts of multiphase hydrodynamics on aqueous reactions and mineral reactions may also need to be considered separately. Taking the recirculation zone as an example, it increases interfacial mass transfer (Maes and Soulaine, 2018), but reduces mineral reactions (Deng et al., 2018). These research needs further emphasize the need for experimental studies that underpin the development of mechanistic models for multiphase flow in reactive systems, and studies that expand our investigations from quasi-2D micromodels to 3D systems. The development of constitutive relations will also require the investigations be extended to larger scales, which calls for development of a complementary approach that leverages modeling and experiments at different scales (Blunt et al., 2013).

SUMMARY AND CONCLUSIONS

A pore-scale multiphase reactive transport model was established and used to investigate the dependence of mineral reaction rate on pore-scale two-phase flow dynamics. A series of numerical simulations were performed, in which CO_2-acidified water and air were co-injected at a range of velocities into 2D calcite channels with different levels of roughness as defined by a single sine wave. The simulation results showed that gas bubbles migrating through the reactive zone as a result of the two-phase co-injection caused the reaction rate to be lower than that of the single phase flow simulation with the same total injection rate. Our analyses revealed that the decrease of the mineral reaction rate is caused by a combination of two mechanisms: the reduction in the wetted surface area and the transport limitations that arise from the two phase flow. In the rough geometries, the narrow throats are reaction hotspots in the single phase flow simulations, whereas these locations are occupied by gas bubbles periodically and thus not accessible for reactions. Meanwhile, the flow field showed clear "vortices," i.e., recirculation zones, as the gas bubbles migrate through the trough of the sine wave. As a result, more reaction products are trapped in the trough and the local thermodynamic driving force and kinetic rates are reduced. Additionally, the correlations between fluid saturation and the extent of reaction rate reduction—which is measured by the ratio between the reaction rate from the two-phase flow simulation and that from the corresponding single phase simulation—were analyzed to provide insights for continuum-scale modeling. A non-monotonic relationship was observed, which is because the contribution from wetted surface area reduction and local transport limitation show opposite dependence on the roughness of the channel, which controls the saturation for a given flow condition. Through these numerical simulations, we highlighted the complexity of reactive transport in multiphase flow systems and identified two important mechanisms through which mineral reaction rates are affected. These results highlight the need for consideration of interfacial dynamics on mineral reaction rates in multiphase flow systems, and also emphasize the need for experimental studies that underpin the development of mechanistic models for multiphase flow in reactive systems.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

PL contributed to developing the code, designing the study, performing the simulations and analyses, and writing the manuscript. HD designed the study, contributed to code development, simulation results analyses, and writing the manuscript. SM contributed to code development and writing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by the Laboratory Directed Research and Development (LDRD) award from Berkeley Lab, provided by the Director, Office of Science of the U.S. Department of Energy (DOE), and by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, DOE, under contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11.

ACKNOWLEDGMENTS

The authors would like to thank Julien Maes and Vitalii Starchenko for helpful discussions on OpenFOAM, and thank the reviewers for the constructive comments and suggestions.

REFERENCES

Aboukhedr, M., Georgoulas, A., Marengo, M., Gavaises, M., and Vogiatzaki, K. (2018). Simulation of micro-flow dynamics at low capillary numbers using adaptive interface compression. *Comput. Fluids* 165, 13–32. doi: 10.1016/j.compfluid.2018.01.009

Águila, J. F., Samper, J., Mon, A., and Montenegro, L. (2020). Dynamic update of flow and transport parameters in reactive transport simulations of radioactive waste repositories. *Appl. Geochem.* 117, 104585. doi: 10.1016/j.apgeochem.2020.104585

Akai, T., Blunt, M. J., and Bijeljic, B. (2020). Pore-scale numerical simulation of low salinity water flooding using the Lattice Boltzmann Method. *J. Colloid Interface Sci.* 566, 444–453. doi: 10.1016/j.jcis.2020.01.065

Alizadeh, M., and Fatemi, M. (2021). Pore-doublet computational fluid dynamic simulation of the effects of dynamic contact angle and interfacial tension alterations on the displacement mechanisms of oil by low salinity water. *Int. J. Multiphase Flow* 143, 103771. doi: 10.1016/j.ijmultiphaseflow.2021.103771

Andre, B., Molins, S., Johnson, J., and Steefel, C. I. (2013). *Alquimia Computer Software*. Retreived from: https://github.com/LBL-EESA/alquimiadev (accessed August 01, 2013).

Armstrong, R. T., Evseev, N., Koroteev, D., and Berg, S. (2015). Modeling the velocity field during Haines jumps in porous media. *Adv. Water Resour.* 77, 57–68. doi: 10.1016/j.advwatres.2015.01.008

Aziz, R., Joekar-Niasar, V., and Martinez-Ferrer, P. (2018). Pore-scale insights into transport and mixing in steady-state two-phase flow in porous media. *Int. J. Multiphase Flow* 51–62. doi: 10.1016/j.ijmultiphaseflow.2018.07.006

Berg, S., Ott, H., Klapp, S. A., Schwing, A., Neiteler, R., Brussee, N., et al. (2013). Real-time 3D imaging of Haines jumps in porous media flow. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3755–3759. doi: 10.1073/pnas.1221373110

Blois, G., Barros, J. M., and Christensen, K. T. (2015). A microscopic particle image velocimetry method for studying the dynamics of immiscible liquid–liquid interactions in a porous micromodel. *Microfluid. Nanofluidics* 18, 1391–1406. doi: 10.1007/s10404-014-1537-1

Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., et al. (2013). Pore-scale imaging and modelling. *Adv. Water Resour.* 51, 197–216. doi: 10.1016/j.advwatres.2012.03.003

Bolster, D., Dentz, M., and Le Borgne, T. (2009). Solute dispersion in channels with periodically varying apertures. *Phys. Fluids* 21, 056601. doi: 10.1063/1.3131982

Bolster, D., Méheust, Y., Le Borgne, T., Bouquain, J., and Davy, P. (2014). Modeling preasymptotic transport in flows with significant inertial and trapping effects – the importance of velocity correlations and a spatial Markov model. *Adv. Water Resour.* 70, 89–103. doi: 10.1016/j.advwatres.2014.04.014

Carrillo, F. J., Bourg, I. C., and Soulaine, C. (2020). Multiphase flow modeling in multiscale porous media: an open-source micro-continuum approach. *J. Comput. Phys. X* 8, 100073. doi: 10.1016/j.jcpx.2020.100073

Chen, L., He, Y., Tao, W.-Q., Zelenay, P., Mukundan, R., and Kang, Q. (2017). Pore-scale study of multiphase reactive transport in fibrous electrodes of vanadium redox flow batteries. *Electrochim. Acta* 248, 425–439. doi: 10.1016/j.electacta.2017.07.086

Chen, L., Kang, Q., Robinson, B. A., He, Y.-L., and Tao, W.-Q. (2013). Pore-scale modeling of multiphase reactive transport with phase transitions and dissolution-precipitation processes in closed systems. *Phys. Rev. E* 87, 043306. doi: 10.1103/PhysRevE.87.043306

Chen, L., Kang, Q., Tang, Q., Robinson, B. A., He, Y.-L., and Tao, W.-Q. (2015). Pore-scale simulation of multicomponent multiphase reactive transport with dissolution and precipitation. *Int. J. Heat Mass. Transf.* 85, 935–949. doi: 10.1016/j.ijheatmasstransfer.2015.02.035

Chou, L., Garrels, R. M., and Wollast, R. (1989). Comparative study of the kinetics and mechanisms of dissolution of carbonate minerals. *Chem. Geol.* 78, 269–282. doi: 10.1016/0009-2541(89)90063-6

Deising, D., Marschall, H., and Bothe, D. (2016). A unified single-field model framework for Volume-Of-Fluid simulations of interfacial species transfer applied to bubbly flows. *Chem. Eng. Sci.* 139, 173–195. doi: 10.1016/j.ces.2015.06.021

Deng, H., Molins, S., Trebotich, D., Steefel, C., and DePaolo, D. (2018). Pore-scale numerical investigation of the impacts of surface roughness: upscaling of reaction rates in rough fractures. *Geochim. Cosmochim. Acta* 239, 374–389. doi: 10.1016/j.gca.2018.08.005

Deshpande, S. S., Anumolu, L., and Trujillo, M. F. (2012). Evaluating the performance of the two-phase flow solver interFoam. *Comput. Sci. Discov.* 5, 014016. doi: 10.1088/1749-4699/5/1/014016

Ferrari, A., Jimenez-Martinez, J., Le Borgne, T., Méheust, Y., and Lunati, I. (2015). Challenges in modeling unstable two-phase flow experiments in porous micromodels. *Water Resour. Res.* 51, 1381–1400. doi: 10.1002/2014wr016384

Govindarajan, D., Deshpande, A. P., and Raghunathan, R. (2018). Enhanced mobility of non aqueous phase liquid (NAPL) during drying of wet sand. *J. Contam. Hydrol.* 209, 1–13. doi: 10.1016/j.jconhyd.2017.12.005

Grunewald, J. B., Goswami, N., Mukherjee, P. P., and Fuller, T. F. (2021). Two-phase dynamics and hysteresis in the PEM fuel cell catalyst layer with the lattice-boltzmann method. *J. Electrochem. Soc.* 168, 024521. doi: 10.1149/1945-7111/abe5e8

Haroun, Y., Legendre, D., and Raynal, L. (2010a). Direct numerical simulation of reactive absorption in gas–liquid flow on structured packing using interface capturing method. *Chem. Eng. Sci.* 65, 351–356. doi: 10.1016/j.ces.2009.07.018

Haroun, Y., Legendre, D., and Raynal, L. (2010b). Volume of fluid method for interfacial reactive mass transfer: application to stable liquid film. *Chem. Eng. Sci.* 65, 2896–2909. doi: 10.1016/j.ces.2010.01.012

Heshmati, M., and Piri, M. (2018). Interfacial boundary conditions and residual trapping: a pore-scale investigation of the effects of wetting phase flow rate and viscosity using micro-particle image velocimetry. *Fuel* 224, 560–578. doi: 10.1016/j.fuel.2018.03.010

Hirt, C. W., and Nichols, B. D. (1981). Volume of fluid (VOF) method for the dynamics of free boundaries. *J. Comput. Phys.* 39, 201–225. doi: 10.1016/0021-9991(81)90145-5

Jiménez-Martínez, J., Anna, P., de, Tabuteau, H., Turuban, R., Borgne, T. L., and Méheust, Y. (2015). Pore-scale mechanisms for the enhancement of mixing in unsaturated porous media and implications for chemical reactions. *Geophys. Res. Lett.* 42, 5316–5324. doi: 10.1002/2015GL064513

Jiménez-Martínez, J., Hyman, J. D., Chen, Y., William Carey, J., Porter, M. L., Kang, Q., et al. (2020). Homogenization of dissolution and enhanced precipitation induced by bubbles in multiphase flow systems. *Geophys. Res. Lett.* 47:e2020GL087163. doi: 10.1029/2020GL087163

Jiménez-Martínez, J., Le Borgne, T., Tabuteau, H., and Méheust, Y. (2017). Impact of saturation on dispersion and mixing in porous media: photobleaching pulse injection experiments and shear-enhanced mixing model. *Water Resour. Res.* 53, 1457–1472. doi: 10.1002/2016WR019849

Jiménez-Martínez, J., Porter, M. L., Hyman, J. D., William Carey, J., and Viswanathan, H. S. (2016). Mixing in a three-phase system: enhanced production of oil-wet reservoirs by CO_2 injection. *Geophys. Res. Lett.* 43, 196–205. doi: 10.1002/2015GL066787

Jo, D., and Revankar, S. T. (2009). Bubble mechanisms and characteristics at pore scale in a packed-bed reactor. *Chem. Eng. Sci.* 64, 3179–3187. doi: 10.1016/j.ces.2009.04.006

Johnson, J. W., Nitao, J. J., and Knauss, K. G. (2004). Reactive transport modelling of CO_2 storage in saline aquifers to elucidate fundamental processes, trapping mechanisms and sequestration partitioning. *Geol. Soc. Lond. Spec. Publ.* 233, 107–128. doi: 10.1144/GSL.SP.2004.233.01.08

Johnson, J. W., Nitao, J. J., Steefel, C. I., and Knauss, K. G. (2001). "Reactive transport modeling of geologic CO_2 sequestration in saline aquifers: the influence of intra-aquifer shales and the relative effectiveness of structural, solubility, and mineral trapping during prograde and retrograde sequestration," in *First National Conference on Carbon Sequestration* (Washington, DC: National Energy and Technology Laboratory USA), 14–17.

Karadimitriou, N. K., Joekar-Niasar, V., Babaei, M., and Shore, C. A. (2016). Critical role of the immobile zone in non-Fickian two-phase transport: a new paradigm. *Environ. Sci. Technol.* 50, 4384–4392. doi: 10.1021/acs.est.5b05947

Karadimitriou, N. K., Joekar-Niasar, V., and Brizuela, O. G. (2017). Hydrodynamic solute transport under two-phase flow conditions. *Sci. Rep.* 7, 6624. doi: 10.1038/s41598-017-06748-1

Kitanidis, P. K., and Dykaar, B. B. (1997). Stokes flow in a slowly varying two-dimensional periodic pore. *Transp. Porous Media* 26, 89–98. doi: 10.1023/A:1006575028391

Kumar, R., Mohanty, K. K., and Others (2011). "Sweep efficiency of heavy oil recovery by chemical methods," in *SPE Annual Technical Conference and Exhibition* (Society of Petroleum Engineers). doi: 10.2118/146839-MS

Li, Y., Blois, G., Kazemifar, F., and Christensen, K. T. (2019). High-speed quantification of pore-scale multiphase flow of water and supercritical CO_2 in 2-D heterogeneous porous micromodels: flow regimes and interface dynamics. *Water Resour. Res.* 55, 3758–3779. doi: 10.1029/2018WR024635

Li, Y., Blois, G., Kazemifar, F., Molla, R. S., and Christensen, K. T. (2021). Pore-scale dynamics of liquid CO_2-water displacement in 2D axisymmetric porous micromodels under strong drainage and weak imbibition conditions: high-speed μPIV measurements. *Front. Water* 3:710370. doi: 10.3389/frwa.2021.710370

Li, Y., Kazemifar, F., Blois, G., and Christensen, K. T. (2017). Micro-PIV measurements of multiphase flow of water and liquid CO_2 in 2-D heterogeneous porous micromodels. *Water Resour. Res.* 53, 6178–6196. doi: 10.1002/2017WR020850

Lichtner, P., and Karra, S. (2014). "Modeling multiscale-multiphase-multicomponent reactive flows in porous media: application to CO_2 sequestration and enhanced geothermal energy using PFLOTRAN," in *Computational Models for CO_2 Geo-sequestration and Compressed Air Energy Storage*, eds R. Al-Khoury and J. Bundschuh (Boca Raton, FL: CRC Press), 81–136.

Lichtner, P. C., Hammond, G. E., Lu, C., Karra, S., Bisht, G., Andre, B., et al. (2015). *PFLOTRAN User Manual: A Massively Parallel Reactive Flow and Transport Model for Describing Surface and Subsurface Processes*. Technical Report, U.S. Department of Energy. doi: 10.2172/1168703

Maes, J., and Geiger, S. (2018). Direct pore-scale reactive transport modelling of dynamic wettability changes induced by surface complexation. *Adv. Water Resour.* 111, 6–19. doi: 10.1016/j.advwatres.2017.10.032

Maes, J., and Soulaine, C. (2018). A new compressive scheme to simulate species transfer across fluid interfaces using the volume-of-fluid method. *Chem. Eng. Sci.* 190, 405–418. doi: 10.1016/j.ces.2018.06.026

Mahabadi, N., Zheng, X., Yun, T. S., van Paassen, L., and Jang, J. (2018). Gas bubble migration and trapping in porous media: pore-scale simulation. *J. Geophys. Res. Solid. Earth* 123, 1060–1071. doi: 10.1002/2017JB015331

Malekzadeh, S., and Roohi, E. (2015). Investigation of different droplet formation regimes in a T-junction microchannel using the VOF technique in OpenFOAM. *Microgravity Sci. Technol.* 27, 231–243. doi: 10.1007/s12217-015-9440-2

Marschall, H., Hinterberger, K., Schüler, C., Habla, F., and Hinrichsen, O. (2012). Numerical simulation of species transfer across fluid interfaces in free-surface flows using OpenFOAM. *Chem. Eng. Sci.* 78, 111–127. doi: 10.1016/j.ces.2012.02.034

Mi, S., Weldetsadik, N. T., Hayat, Z., Fu, T., Zhu, C., Jiang, S., et al. (2019). Effects of the gas feed on bubble formation in a microfluidic T-junction: constant-pressure versus constant-flow-rate injection. *Ind. Eng. Chem. Res.* 58, 10092–10105. doi: 10.1021/acs.iecr.9b01262

Mohammadi Alamooti, A. H., Azizi, Q., and Davarzani, H. (2020). Direct numerical simulation of trapped-phase recirculation at low capillary number. *Adv. Water Resour.* 145, 103717. doi: 10.1016/j.advwatres.2020.103717

Molins, S., Trebotich, D., Steefel, C. I., and Shen, C. (2012). An investigation of the effect of pore scale flow on average geochemical reaction rates using direct numerical simulation. *Water Resour. Res.* 48, W03527. doi: 10.1029/2011WR011404

Molins, S., Trebotich, D., Yang, L., Ajo-Franklin, J. B., Ligocki, T. J., Shen, C., et al. (2014). Pore-scale controls on calcite dissolution rates from flow-through laboratory and numerical experiments. *Environ. Sci. Technol.* 48, 7453–7460. doi: 10.1021/es5013438

Moura, M., Flekkøy, E. G., Måløy, K. J., Schäfer, G., and Toussaint, R. (2019). Connectivity enhancement due to film flow in porous media. *Phys. Rev. Fluids* 4, 094102. doi: 10.1103/PhysRevFluids.4.094102

Ngien, S. K., Rahman, N. A., Bob, M. M., Ahmad, K., Sa'ari, R., and Lewis, R. W. (2012). Observation of light non-aqueous phase liquid migration in aggregated soil using image analysis. *Transp. Porous Media* 92, 83–100. doi: 10.1007/s11242-011-9892-9

Ott, H., and Oedai, S. (2015). Wormhole formation and compact dissolution in single- and two-phase CO_2-brine injections. *Geophys. Res. Lett.* 42, 2270–2276. doi: 10.1002/2015GL063582

Pan, Y., Yang, J., Jia, Y., and Xu, Z. (2016). Experimental study on non-aqueous phase liquid multiphase flow characteristics and controlling factors in heterogeneous porous media. *Environ. Earth Sci.* 75, 75. doi: 10.1007/s12665-015-4888-3

Paulsen, J. D., Carmigniani, R., Kannan, A., Burton, J. C., and Nagel, S. R. (2014). Coalescence of bubbles and drops in an outer fluid. *Nat. Commun.* 5, 1–7. doi: 10.1038/ncomms4182

Pavuluri, S., Maes, J., and Doster, F. (2018). Spontaneous imbibition in a microchannel: analytical solution and assessment of volume of fluid formulations. *Microfluid. Nanofluidics*. 22:90. doi: 10.007/s10404-018-2106-9

Picchi, D., and Battiato, I. (2018). The impact of pore-scale flow regimes on upscaling of immiscible two-phase flow in porous media. *Water Resour. Res.* 54, 6683–6707. doi: 10.1029/2018WR023172

Popp, S., Beyer, C., Dahmke, A., and Bauer, S. (2015). Model development and numerical simulation of a seasonal heat storage in a contaminated shallow aquifer. *Energy Proc.* 76, 361–370. doi: 10.1016/j.egypro.2015.07.842

Pruess, K. (2004). The TOUGH codes—a family of simulation tools for multiphase flow and transport processes in permeable media. *Vadose Zone J.* 3, 738–746. doi: 10.2113/3.3.738

Qin, Z., Esmaeilzadeh, S., Riaz, A., and Tchelepi, H. A. (2020). Two-phase multiscale numerical framework for modeling thin films on curved solid surfaces in porous media. *J. Comput. Phys.* 413, 109464. doi: 10.1016/j.jcp.2020.109464

Ren, T., Zhu, Z., Zhang, R., Shi, J., and Yan, C. (2020). Visualization experiment of bubble coalescence in a narrow vertical rectangular channel. *Front. Energy Res.* 8, 96. doi: 10.3389/fenrg.2020.00096

Roman, S., Abu-Al-Saud, M. O., Tokunaga, T., Wan, J., Kovscek, A. R., and Tchelepi, H. A. (2017). Measurements and simulation of liquid films during drainage displacements and snap-off in constricted capillary tubes. *J. Colloid Interface Sci.* 507, 279–289. doi: 10.1016/j.jcis.2017.07.092

Roman, S., Soulaine, C., AlSaud, M. A., Kovscek, A., and Tchelepi, H. (2016). Particle velocimetry analysis of immiscible two-phase flow in micromodels. *Adv. Water Resour.* 95, 199–211. doi: 10.1016/j.advwatres.2015.08.015

Rücker, M., Berg, S., Armstrong, R. T., Georgiadis, A., Ott, H., Schwing, A., et al. (2015). From connected pathway flow to ganglion dynamics. *Geophys. Res. Lett.* 42, 3888–3894. doi: 10.1002/2015GL064007

Rusche, H. (2002). *Computational Fluid Dynamics of Dispersed Two-Phase Flows at High Phase Fractions*. Ph.D., Imperial College London.

Satter, A., and Iqbal, G. M. (2016). "Reservoir rock properties," *Reservoir Engineering*, ed A. Satter and G. M. Iqbal (Boston, MA: Gulf Professional Press), 29–79.

Schlüter, S., Berg, S., Rücker, M., Armstrong, R. T., Vogel, H.-J., Hilfer, R., et al. (2016). Pore-scale displacement mechanisms as a source of hysteresis for two-phase flow in porous media. *Water Resour. Res.* 52, 2194–2205. doi: 10.1002/2015WR018254

Shuard, A. M., Mahmud, H. B., and King, A. J. (2016). Comparison of two-phase pipe flow in openfoam with a mechanistic model. *IOP Conf. Ser. Mater. Sci. Eng.* 121, 012018. doi: 10.1088/1757-899X/121/1/012018

Song, W., Ogunbanwo, F., Steinsb,ø, M., Fern,ø, M. A., and Kovscek, A. R. (2018). Mechanisms of multiphase reactive flow using biogenically calcite-functionalized micromodels. *Lab Chip* 18, 3881–3891. doi: 10.1039/C8LC00793D

Sonnenthal, E., Ito, A., Spycher, N., Yui, M., Apps, J., Sugita, Y., et al. (2005). Approaches to modeling coupled thermal, hydrological, and chemical processes in the drift scale heater test at Yucca Mountain. *Int. J. Rock Mech. Min. Sci.* 42, 698–719. doi: 10.1016/j.ijrmms.2005.03.009

Sookhak Lari, K., Davis, G. B., Rayner, J. L., Bastow, T. P., and Puzon, G. J. (2019). Natural source zone depletion of LNAPL: a critical review supporting modelling approaches. *Water Res.* 157, 630–646. doi: 10.1016/j.watres.2019.04.001

Soulaine, C., Maes, J., and Roman, S. (2021). Computational microfluidics for geosciences. *Front. Water* 3:643714. doi: 10.3389/frwa.2021.643714

Soulaine, C., Roman, S., Kovscek, A., and Tchelepi, H. A. (2018). Pore-scale modelling of multiphase reactive flow: application to mineral dissolution with production of. *J. Fluid Mech.* 855, 616–645. doi: 10.1017/jfm.2018.655

Spurin, C., Bultreys, T., Bijeljic, B., Blunt, M. J., and Krevor, S. (2019). Intermittent fluid connectivity during two-phase flow in a heterogeneous carbonate rock. *Phys. Rev. E* 100, 043103. doi: 10.1103/PhysRevE.100.043103

Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., et al. (2015). Reactive transport codes for subsurface environmental simulation. *Comput. Geosci.* 19, 445–478. doi: 10.1007/s10596-014-9443-x

Sund, N. L., Bolster, D., and Dawson, C. (2015). Upscaling transport of a reacting solute through a peridocially converging–diverging channel at pre-asymptotic times. *J. Contam. Hydrol.* 182, 1–15. doi: 10.1016/j.jconhyd.2015.08.003

Toussaint, R., Maloy, K. J., Meheust, Y., Lovoll, G., Jankov, M., Schaefer, G., et al. (2012). Two-phase flow: structure, upscaling, and consequences for macroscopic transport properties. *Vadose Zone J.* 11:vzj2011.0123. doi: 10.2136/vzj2011.0123

van Steijn, V., Kleijn, C. R., and Kreutzer, M. T. (2010). Predictive model for the size of bubbles and droplets created in microfluidic T-junctions. *Lab Chip* 10, 2513–2518. doi: 10.1039/c002625e

Wang, C., Mehmani, Y., and Xu, K. (2021). Capillary equilibrium of bubbles in porous media. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2024069118. doi: 10.1073/pnas.2024069118

Wu, R., Chen, X., Hammond, G., Bisht, G., Song, X., Huang, M., et al. (2021). Coupling surface flow with high-performance subsurface reactive flow and transport code PFLOTRAN. *Environ. Model. Softw.* 137, 104959. doi: 10.1016/j.envsoft.2021.104959

Xiao, T., Xu, H., Moodie, N., Esser, R., Jia, W., Zheng, L., et al. (2020). Chemical-mechanical impacts of CO_2 intrusion into heterogeneous caprock. *Water Resour. Res.* 56:e2020WR027193. doi: 10.1029/2020WR027193

Xiong, Q., Baychev, T. G., and Jivkov, A. P. (2016). Review of pore network modelling of porous media: experimental characterisations, network constructions and applications to reactive transport. *J. Contam. Hydrol.* 192, 101–117. doi: 10.1016/j.jconhyd.2016.07.002

Xu, T., Spycher, N., Sonnenthal, E., Zhang, G., Zheng, L., and Pruess, K. (2011a). TOUGHREACT Version 2.0: a simulator for subsurface reactive transport under non-isothermal multiphase flow conditions. *Comput. Geosci.* 37, 763–774. doi: 10.1016/j.cageo.2010.10.007

Xu, T., Zheng, L., and Tian, H. (2011b). Reactive transport modeling for CO_2 geological sequestration. *J. Pet. Sci. Eng.* 78, 765–777. doi: 10.1016/j.petrol.2011.09.005

Xu, T., Zhu, H., Feng, G., Yang, Z., and Tian, H. (2019). Numerical simulation of calcite vein formation and its impact on caprock sealing efficiency – case study of a natural CO_2 reservoir. *Int. J. Greenhouse Gas Control* 83, 29–42. doi: 10.1016/j.ijggc.2019.01.021

Yin, X., Zarikos, I., Karadimitriou, N. K., Raoof, A., and Hassanizadeh, S. M. (2019). Direct simulations of two-phase flow experiments of different geometry complexities using Volume-of-Fluid (VOF) method. *Chem. Eng. Sci.* 195, 820–827. doi: 10.1016/j.ces.2018.10.029

Yoon, S., and Kang, P. K. (2021). Roughness, inertia, and diffusion effects on anomalous transport in rough channel flows. *Phys. Rev. Fluids* 6, 014502. doi: 10.1103/PhysRevFluids.6.014502

Zhang, Y., Morrow, N. R., and Others (2006). "Comparison of secondary and tertiary recovery with change in injection brine composition for crude-oil/sandstone combinations," in *SPE/DOE Symposium on Improved Oil Recovery* (Society of Petroleum Engineers). doi: 10.2118/99757-MS

Zhao, B., MacMinn, C. W., Primkulov, B. K., Chen, Y., Valocchi, A. J., Zhao, J., et al. (2019). Comprehensive comparison of pore-scale models for multiphase flow in porous media. *Proc. Natl. Acad. Sci. U.S.A.* 116, 13799–13806. doi: 10.1073/pnas.1901619116

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.

frontiers in Water

ORIGINAL RESEARCH
published: 17 January 2022
doi: 10.3389/frwa.2021.797259

Unraveling Long-Term Flood Risk Dynamics Across the Murray-Darling Basin Using a Large-Scale Hydraulic Model and Satellite Data

Serena Ceola[1]*, Alessio Domeneghetti[1] and Guy J. P. Schumann[2,3]

[1] Department of Civil, Chemical, Environmental, and Materials Engineering, Alma Mater Studiorum Università di Bologna, Bologna, Italy, [2] School of Geographical Sciences, University of Bristol, Bristol, United Kingdom, [3] Research and Education Department, RSS-Hydro, Dudelange, Luxembourg

Edited by:
Mariele Evers,
University of Bonn, Germany

Reviewed by:
Maurizio Mazzoleni,
Uppsala University, Sweden
Weili Duan,
Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Serena Ceola
serena.ceola@unibo.it

Specialty section:
This article was submitted to Water and Human Systems, a section of the journal Frontiers in Water

Received: 18 October 2021
Accepted: 21 December 2021
Published: 17 January 2022

Citation:
Ceola S, Domeneghetti A and Schumann GJP (2022) Unraveling Long-Term Flood Risk Dynamics Across the Murray-Darling Basin Using a Large-Scale Hydraulic Model and Satellite Data.
Front. Water 3:797259.
doi: 10.3389/frwa.2021.797259

River floods are one of the most devastating extreme hydrological events, with oftentimes remarkably negative effects for human society and the environment. Economic losses and social consequences, in terms of affected people and human fatalities, are increasing worldwide due to climate change and urbanization processes. Long-term dynamics of flood risk are intimately driven by the temporal evolution of hazard, exposure and vulnerability. Although needed for effective flood risk management, a comprehensive long-term analysis of all these components is not straightforward, mostly due to a lack of hydrological data, exposure information, and large computational resources required for 2-D flood model simulations at adequately high resolution over large spatial scales. This study tries to overcome these limitations and attempts to investigate the dynamics of different flood risk components in the Murray-Darling basin (MDB, Australia) in the period 1973–2014. To this aim, the LISFLOOD-FP model, i.e., a large-scale 2-D hydrodynamic model, and satellite-derived built-up data are employed. Results show that the maximum extension of flooded areas decreases in time, without revealing any significant geographical transfer of inundated areas across the study period. Despite this, a remarkable increment of built-up areas characterizes MDB, with larger annual increments across not-flooded locations compared to flooded areas. When combining flood hazard and exposure, we find that the overall extension of areas exposed to high flood risk more than doubled within the study period, thus highlighting the need for improving flood risk awareness and flood mitigation strategies in the near future.

Keywords: flood hazard, flood exposure, flood memory, satellite and model data, Murray-Darling basin (MDB), long-term

INTRODUCTION

Economic losses and social consequences associated with riverine inundations appear to increase worldwide and the intensification of extreme hydrological events due to climate change is often pointed out as the main cause (de Moel et al., 2011; Barnes, 2017; IPCC, 2021). Yet urban growth, flood mitigation infrastructures and increasing human presence and activity in floodplains rule, in a similar way, flood risk, as they contribute to shape the origin and location of flood events as well as the elements at risk (Gupta et al., 2015; Duan et al., 2016; Merz et al., 2021; Tellman et al., 2021). River floods are driven by prolonged rainfall periods that induce high water levels overtopping river embankments, thus inundating nearby locations. Flood risk analyzes cascading impacts on society and the environment. More specifically, actual flood risk results from the superposition of three components (IPCC, 2012), such as hazard (i.e., frequency of occurrence of flood events), exposure (i.e., elements at risk, including direct and indirect damages, people, capital investment, and land or property value), and vulnerability (i.e., the capacity to deal with flood events), which separately can control and impact different dynamics of flood risk evolution (Merz et al., 2010, 2021; Ceola et al., 2014; Domeneghetti et al., 2015). Future projections of increasing population and economic activities on river floodplains as derived from socioeconomic growth scenarios, as well as increasing heavy rainfall estimates associated to climate variability and change, will likely result in increasing flood risk (Hirabayashi et al., 2013; Winsemius et al., 2016; Kam et al., 2021). As a consequence, it is crucial to unravel long-term dynamics of flood risk and its components.

Several methods have been developed both at local and global scales to allow for a detailed assessment of flood hazard, either based on traditional hydrological and hydraulic models (Bates et al., 2010; Yamazaki et al., 2011; Pappenberger et al., 2012; Winsemius et al., 2013; Rudari et al., 2015; Sampson et al., 2015; Dottori et al., 2016; Schumann et al., 2016) or innovative DEM-based (digital elevation model) techniques (Lee et al., 2017; Samela et al., 2017; Tavares da Costa et al., 2020). Typically, flood models simulate inundated areas based on the probability of exceedance of a particular discharge value (i.e., by considering a particular return period) or based on long-term time series of discharge, without accounting for detailed topographic features along floodplains. To overcome this issue, which may result in an approximate identification of the actual flood spatial extension, and thus exposed people and assets, Bates et al. (2010) and Schumann et al. (2016) proposed to assess flood hazard by accounting for detailed floodplain topography. Similarly, Tellman et al. (2021), contributed to estimate the actual flood extent and population exposure for several flood events occurred between 2000 and 2018 by employing high-resolution satellite imagery. Both approaches provide new standards and are expected to improve the accuracy of local and global flood models, our knowledge about how climate, human, land changes interact with flood dynamics, as well as the development of effective flood management strategies.

In order to assess flood exposure, several population and built-up datasets recently released are typically employed (Ceola et al., 2014; Leyk et al., 2019; Bernhofen et al., 2021). These include e.g., gridded population of the world (Center for International Earth Science Information Network (CIESIN), 2016), global human settlement layer (Corbane et al., 2018), global urban footprint (Palacios-Lopez et al., 2019), Facebook's High Resolution Settlement Layer (Facebook Connectivity Lab, and Center for International Earth Science Information Network (CIESIN), 2016), WorldPop (Lloyd et al., 2019), but also unconventional datasets such as nighttime lights (NOAA - Earth Observation Group, 2016). These datasets present different spatial and temporal resolutions and coverage, which are not always suited for an effective assessment of long-term flood exposure dynamics (Leyk et al., 2019; Bernhofen et al., 2021).

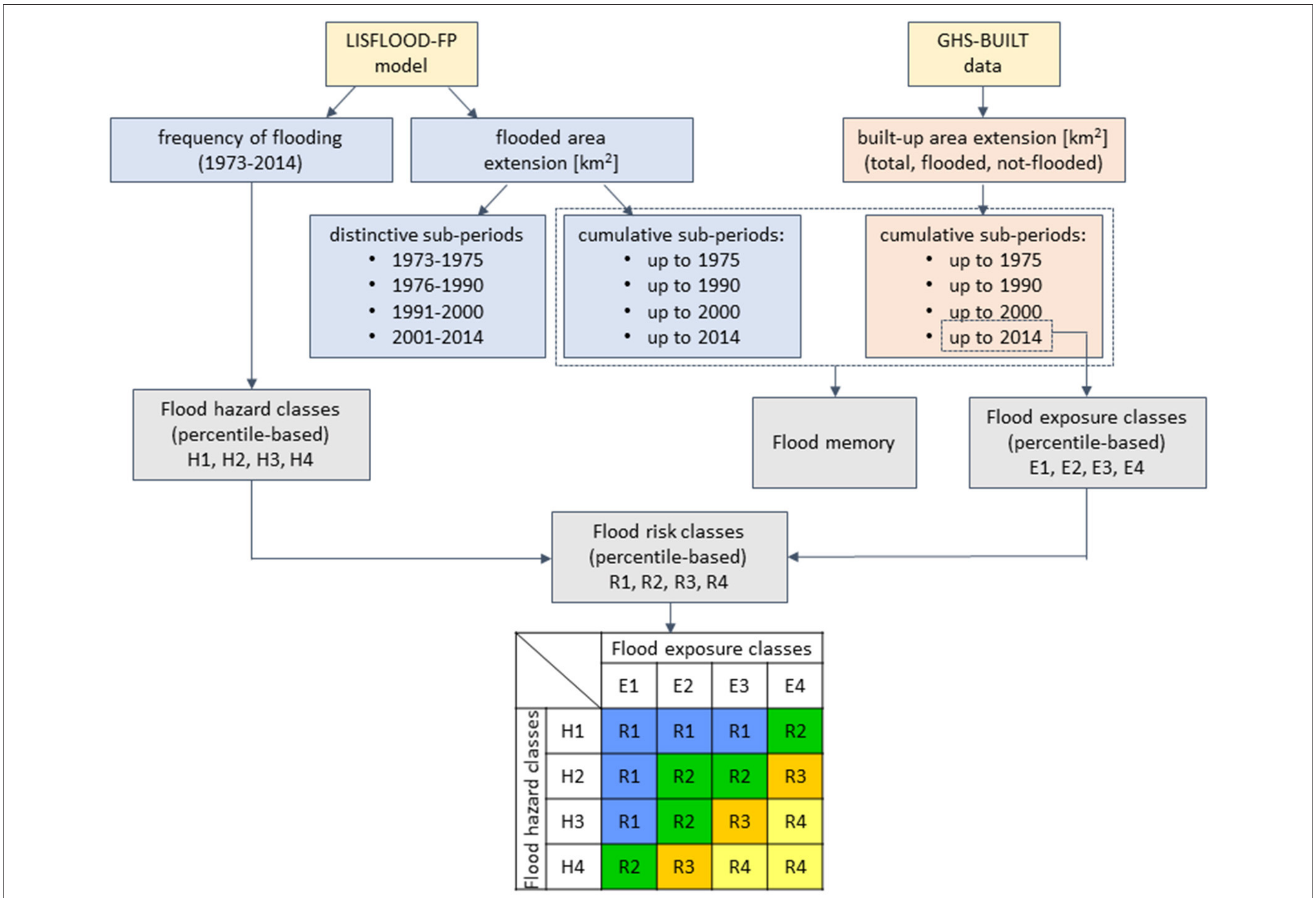
The assessment of flood vulnerability, which depends on economic, social, demographic, cultural, institutional and governance factors, is typically the most challenging task. People can alter the hydrologic regime, including extremes such as floods and droughts (Viglione et al., 2016), and in turn these extremes can also shape human society (Ridolfi et al., 2020). It is well-acknowledged that previous flooding experience, income and education control risk perception and influence flood memory and thus flood vulnerability (Jonkman and Kelman, 2005; EU Floods Directive, 2007; Kreibich et al., 2011; Garde-Hansen et al., 2017; Aerts et al., 2018), though it is difficult to quantify and model it. Some attempts are available from the scientific literature, e.g., FLOPROS, a global database of flood protection standards used as a proxy of flood vulnerability (Scussolini et al., 2016), or comparative estimations of observed vs. modeled data (Tanoue et al., 2016), yet further efforts are needed.

Therefore, unraveling long-term dynamics of both flood risk and its components is critical and crucial to allow for an effective flood management, by planning investments in adaptation strategies and improving people awareness of flood risk (Duan et al., 2016; Merz et al., 2021). However, a comprehensive long-term analysis of flood risk components is not straightforward, mostly due to a lack of hydrological data, exposure information, and large computational resources required for 2-D flood model simulations at adequately high resolution over large spatial scales. This study tries to overcome these limitations and attempts to investigate the dynamics of different flood-risk components in the Murray-Darling basin (MDB, Australia) in the period 1973–2014. In particular, we examined long-term dynamics of (i) flood hazard, (ii) flood exposure, (iii) flood memory and (iv) flood risk (see flowchart in **Figure 1**). More specifically, the spatio-temporal evolution of flood hazard is assessed by means of a large-scale 2-D hydrodynamic model, based on remotely-sensed STRM (Shuttle Radar Topography Mission) data, while for exposure, the Global Human Settlement Layer (GHSL), particularly the built-up area data, provided and distributed by the JRC, is used. As a preventive measure, flood vulnerability is not considered here.

MATERIALS AND METHODS

Study Area

The study area considered here is the MDB located in the southeastern part of Australia (**Figure 2**). The MDB is the largest river basin in Australia, with a drainage area of $\sim 10^6$ km^2, covering 14% of Australia's land area. The MDB is composed by 22 sub-basins (Murray-Darling Basin Authority, 2020). The MDB provides water to 2.2 million people, typically used for agriculture, drinking and recreation and it is also home to internationally significant wetlands and supports a huge range of endangered species



	E1	E2	E3	E4
H1	R1	R1	R1	R2
H2	R1	R2	R2	R3
H3	R1	R2	R3	R4
H4	R2	R3	R4	R4

FIGURE 1 | Analysis of long-term flood risk dynamics: research flowchart.

(Murray-Darling Basin Authority, 2020). Given its relevance, several research initiatives and projects focused on MDB water resource management, hydrological extremes and environmental and societal impacts have been recently developed (e.g., Gallant et al., 2012; Heimhuber et al., 2016; Bishop-Taylor et al., 2018).

The two main rivers in the MDB, the Murray (2,530 km) and the Darling (2,740 km), are two of Australia's longest rivers. Water in the northern part of MDB runs into the Darling River while water in the southern part of MDB runs into the Murray River. Most of the tributaries in the MDB start in the Great Dividing Range, a series of mountains along the east coast. Then, water flows on flat plains, which cover the majority of the basin area. A pronounced climate gradient characterizes the MDB: from the southeast to the northwest, average annual rainfall decreases (from 1,500 to 300 mm, approximately) and climate variability and evapotranspiration increase (Murray-Darling Basin Authority, 2020). As a consequence, flooding regimes differ significantly across the MDB, with severe flood outbreaks during summer in the northern part of MDB and long flood events driven by rainfall and snowmelt during winter and spring in the southern part of MDB (Bunn et al., 2006; Penton and Overton, 2007). Several catastrophic flood events occurred across the MDB in 1917, 1931, 1952, 1956, 1974, 1993 and 2020 (Murray-Darling Basin Authority, 2020). By chance, our study period includes one of the largest flood events ever occurred across the MDB (i.e., flood event in January 1974, also known as the "big wet"), which was the biggest event in the considered study period.

2-D Hydrodynamic Model for Flood Hazard Assessment

Long-term flood hazard dynamics over the period 1973–2013 has been reconstructed referring to 2-D hydrodynamic simulations performed across the overall MDB. In particular, in this study we retrieve the inundation maps obtained from the continental-scale version of the 2-D hydrodynamic model settled by Schumann et al. (2016) over Australia, specifically applied to the MDB by Grimaldi et al. (2019). The 2-D numerical model (LISFLOOD-FP; Bates et al., 2010) adopts a single explicit finite difference scheme to solve the inertial momentum equation over a regular grid (e.g., raster-based scheme). The model was built from freely available Shuttle Radar Topography Mission (SRTM) data, which

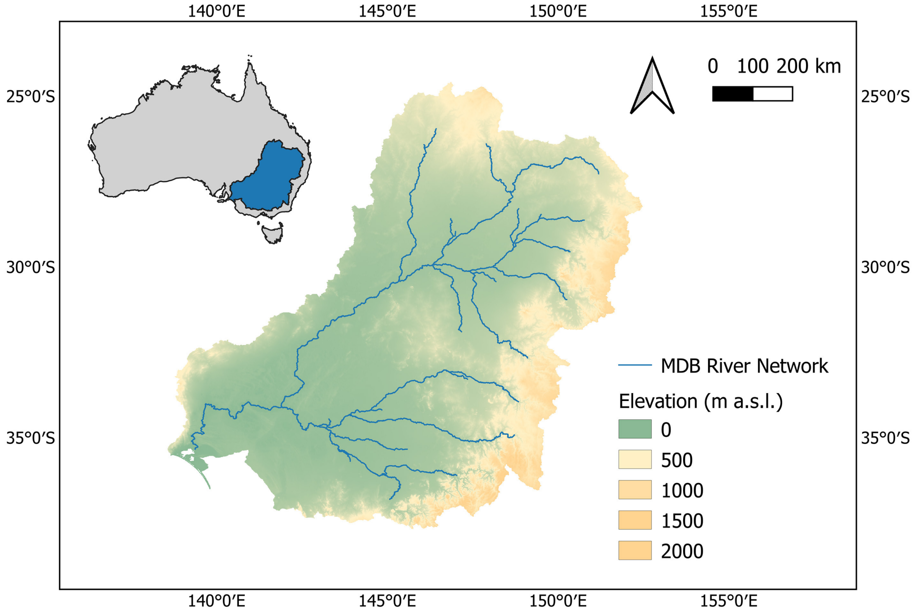


FIGURE 2 | Study area: Location of the Murray Darling Basin, main river network and elevation.

was corrected for vegetation canopy height using a global ICESat-1 canopy data set (Simard et al., 2011). Despite the nominal raster resolution, the subgrid channel formulation adopted in the modeling framework (Neal et al., 2012) ensures the capability to reproduce hydraulic dynamics of rivers and floodplains having widths smaller than the adopted grid cell.

All rivers that drain a catchment area $>$10,000 km^2 were explicitly represented in our model. Furthermore, significant flow contributions from smaller tributaries were accounted for as additional inflow points along those major rivers. The model also includes lakes and reservoirs from the Global Lake and Wetland Database (Lehner and Döll, 2004), as well as the evaporative water loss over the MDB based on satellite-sensed daily evaporations estimates (i.e., Global Land Evaporation Amsterdam Model; Martens et al., 2017). Reservoirs and lakes were filled before the simulation was run and were implicitly regulated by the hydrodynamics of the model during simulation. In light of the convincing calibration scores achieved in reproducing past inundations by previous investigations, this study adopts the same model parameters. The reader can refer to Schumann et al. (2016) and Grimaldi et al. (2019) for additional details on model settings, calibration events and modeling assumptions, including also a thorough assessment of uncertainty as influenced by model input, structure and parameters. Flood hazard maps employed in this study are produced at 1 km resolution from January 1973 until July 2013. Outputs from LISFLOOD-FP model simulations are 487 monthly (30 days long) maps of flooded area, flooded water depth, flooded water volume and frequency of flooding. In our analysis, water depth and water volume data are not considered since we are interested only in the identification of flooded locations and their frequency of flooding.

Satellite Data for Flood Exposure Assessment

In order to assess the spatio-temporal evolution of human exposure to floods, we employ the Global Human Settlement Layer (GHSL), a project of the European Commission's Joint Research Centre, which freely provides spatially detailed information on population and settlements (Corbane et al., 2018). In this research, we employ the GHS-BUILT data (at 30 m resolution) that contain a multitemporal information layer on built-up presence as derived from Landsat image collections (GLS1975, GLS1990, GLS2000, and *ad-hoc* Landsat eight collection 2013/2014), whose values range from 0 (no built-up surface) to 1 (whole built-up surface). In total, four snapshot information on built-up density and distribution are available within the considered study period. Only for comparative purposes, we use the GHS-POP data (at 250 m resolution, Schiavina et al., 2019) that provides the distribution of population, expressed as the number of people per cell for years 1975, 1990, 2000 and 2015. GHS-POP estimates are derived from a spatial disaggregation of CIESIN GPWv4.10 data (Center for International Earth Science Information Network (CIESIN), 2016) based on GHS-BUILT data. In order to perform a coherent superimposition of inundation and built-up data, we aggregate both GHS-BUILT and GHS-POP data at the same resolution as the hydrodynamic model outputs (i.e., 1 km resolution).

Analysis of Long-Term Dynamics of Flood Risk Variables

The temporal evolution of flood risk dynamics from 1973 to 2014 across the MDB is investigated across four sub-periods, whose subdivision depends on the intrinsic temporal availability of built-up data (i.e., 1975, 1990, 2000 and 2014). Inundation maps, as derived from LISFLOOD-FP simulations, are available from January 1973 to July 2013. Since no modeled inundation data is available afterwards (from August 2013 to December 2014), we assume that no flood event occurred during this short time window (i.e., flooded area and frequency of flooding are equal to zero in each grid cell and month), thus slightly underestimating flood hazard.

We assess long-term dynamics of (i) flood hazard, (ii) flood exposure, (iii) flood memory, and (iv) flood risk, as detailed in what follows and schematically shown in **Figure 1**.

Flood Hazard

Regarding flood hazard (H), we compute the frequency of occurrence of inundation states for each grid cell within the whole study period (1973–2014), by looking at grid cells that were flooded at least one month, as derived from LISFLOOD-FP model simulations (Schumann et al., 2016; Grimaldi et al., 2019). The frequency of occurrence of inundation states is defined as the ratio between the total number of flooded months and the total length of the study period (i.e., 504 months). We categorize these values into four classes, based on the 25th, 50th, 75th and 100th percentiles, where values within the first ($0 < H \leq H_{25}$), second ($H_{25} < H \leq H_{50}$), third ($H_{50} < H \leq H_{75}$) and fourth ($H_{75} < H \leq H_{100}$) quartile belong to class H1, H2, H3, and H4, respectively. We also quantify the maximum extension of flooded areas (in km^2) and analyze its temporal evolution to identify increasing or decreasing trends. In particular, the maximum extension of flooded area is computed by considering distinct sub-periods (i.e., flooded areas in 1973–1975, 1976–1990, 1991–2000, 2001–2014) and cumulative sub-periods (i.e., flooded areas in 1973–1975, 1973–1990, 1973–2000, 1973–2014). While the first approach allows to detect singularities within the study period (i.e., "flood rich" vs. "flood poor" sub-periods, larger vs. smaller flooded areas), the second one allows to identify possible geographical transfers of flood hazard in time. Finally, for each flood hazard class and for each distinct sub-period we quantify the maximum extension of flooded areas to be employed for the assessment of flood risk dynamics (see section Flood Risk).

Flood Exposure

Concerning flood exposure (E), we quantify the maximum extension of built-up areas (i.e., an indicator of human presence), as derived from GHS-BUILT data (Corbane et al., 2018), and then analyze its temporal trend. Since GHS-BUILT provide built-up data based on cumulative sub-periods (i.e., built-up values until 1975, until 1990, until 2000, and until 2014), we assess the spatial extension at the end of each sub-period. GHS-BUILT values range from 0 to 1, i.e., from no built-up to whole built-up surface, respectively. Based on these features, we compute the maximum extension of built-up area, defined as the sum of built-up values times the grid cell size (in km^2). We then disaggregate this result and distinguish between flooded and not flooded built-up locations, by superimposing the geographical location of inundated areas as derived from LISFLOOD-FP model simulations based on cumulative sub-periods. For the sake of completeness, we also evaluate the total number of built-up grid cells, regardless of their built-up value (i.e., by considering any grid cell with GHS-BUILT>0). Similarly to the analysis performed for flood hazard, we define four classes of exposure based on quartile values associated to the most recent built-up data (up to 2014), which embeds all built-up areas in any epoch, i.e., E1 where $0 < E \leq E_{25}$, E2 where $E_{25} < E \leq E_{50}$, E3 where $E_{50} < E \leq E_{75}$, and E4 where $E_{75} < E \leq E_{100}$. We then compute the maximum extension of built-up areas for each exposure class and sub-period, which is functional for the analysis of flood risk dynamics (see section Flood Risk).

Flood Memory

We then analyze flood memory to estimate the temporal evolution of the extension of built-up areas based on antecedent flood hazard conditions (Aerts et al., 2018; Merz et al., 2021). Flood memory is known to influence human resilience to floods, as prolonged periods without significant flood events may increase flood exposure and lead to low risk awareness, which may potentially result in an inadequate response to flood disasters (Garde-Hansen et al., 2017). Here, the goal is to verify if previous flood events may have discouraged the subsequent evolution of built-up areas, thus revealing a flood memory behavior. More specifically, given the overall flooded area in each cumulative sub-period, we quantify the maximum extension of subsequent built-up areas and compute the ratio between flooded built-up areas and cumulative flooded areas (as %). In particular, when considering the flooded area from 1973 to 1975, we look at the built-up evolution in 1975, 1990, 2000 and 2014; when considering the flooded area from 1973 to 1990, we analyze built-up in 1990, 2000 and 2014; and finally, when considering the flooded area from 1973 to 2000, we examine the built-up area in 2000 and 2014.

Flood Risk

We finally analyze the temporal evolution of flood risk (R), here computed by considering concurrent hazard and exposure, whose values are evaluated based on distinct sub-periods. As mentioned before, flood vulnerability is not considered due to lack of reliable data. Starting from the classification of flood hazard and exposure (see section Flood Exposure and Flood Memory), we define four categories of flood risk, as shown in **Figure 1**. In a given risk category, we consider only those grid cells where concurrent hazard and exposure values are both larger than zero and we quantify flood risk as the maximum extension of built-up areas in each risk category and distinct sub-period.

RESULTS AND DISCUSSION

The frequency of occurrence of inundation states across the MDB from 1973 to 2014, as derived from LISFLOOD-FP model simulations, is shown in **Figure 3A**. Percentile values of the

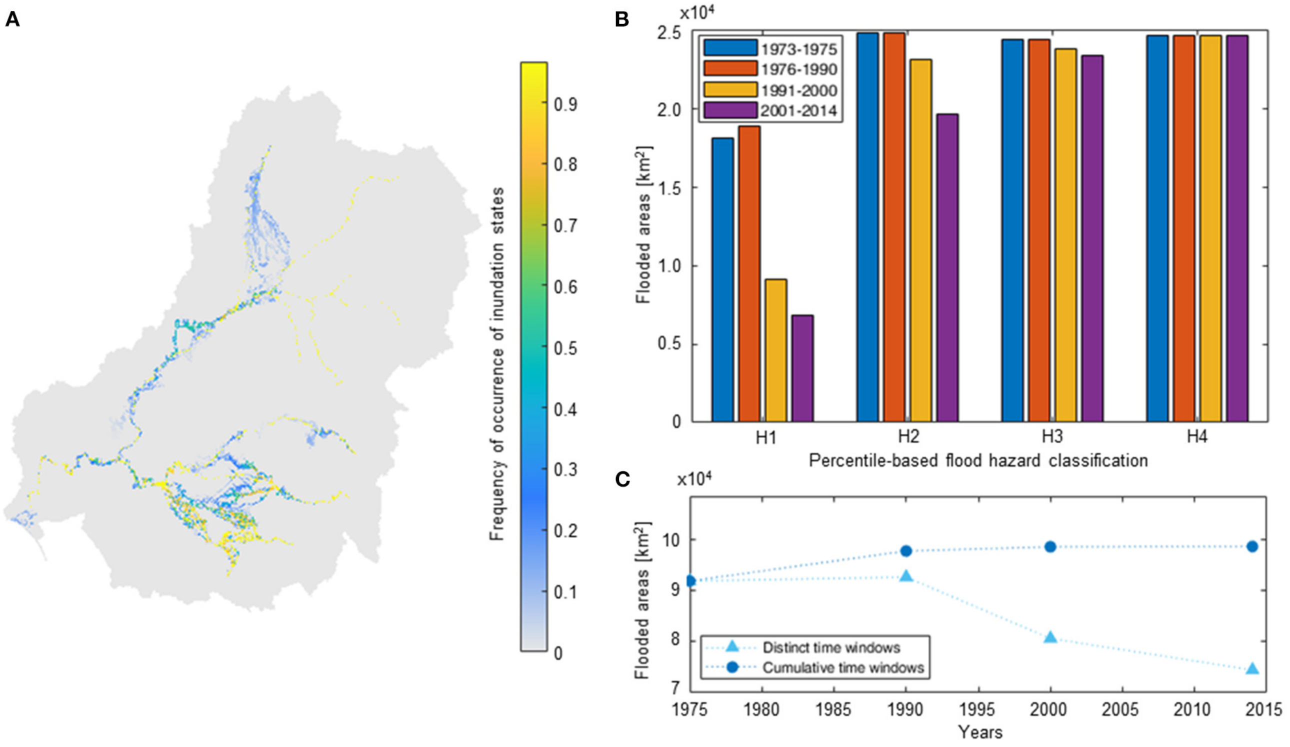


FIGURE 3 | Flood hazard assessment across the MDB from 1973 to 2014. **(A)** Geographical representation of the frequency of occurrence of inundation states. **(B)** Temporal evolution of the maximum extension of flooded areas (in km^2) in each flood hazard category (percentile-based classification). **(C)** Temporal evolution of the maximum extension of flooded areas (in km^2), as derived from distinct (light blue triangles) and cumulative (dark blue circles) sub-periods.

frequency of occurrence of inundation states that characterize flood hazard classes are $H_{25} = 0.03$, $H_{50} = 0.12$, $H_{75} = 0.50$ and $H_{100} = 0.97$, showing, among all, that half of the flooded grid cells was inundated on average for ∼5 years within the whole study period. The maximum extension of inundated areas based on distinct sub-periods decreases in time (**Figure 3C**, light blue line), from more than 90,000 km^2 (nearly 6.3% of MDB area) to a∼75,000 km^2 (5% of MDB area). Given the selected percentile-based classification, the maximum extension of flooded areas shows a remarkable reduction in H1 and H2 classes, which identify unfrequently inundated locations likely associated to major flood events (**Figure 3B**). Conversely, H3 and H4 classes, corresponding to recurring inundation states, are characterized by a slightly decreasing or even invariant pattern in time. In particular, our results suggest that 1973–1975 and 1976–1990 are the most devastating sub-periods, as also confirmed by the absence of increasing trends in the number of flood peaks and flooded areas afterwards (Hu et al., 2018; Merz et al., 2021). When considering cumulative sub-periods (**Figure 3C**, dark blue line), we do not find any remarkable geographical transfer of flood hazard in time, since the maximum extension of inundated areas does not significantly change in time (i.e., only a feeble increase is detected). This outcome might suggest that inundations occur more frequently in the same locations compared to the past, as influenced for instance by improved risk awareness and structural defenses.

The maximum extension of built-up areas (i.e., the sum of built-up values times the grid cell size, in km^2), as shown in **Figure 4A** (black line), increases in time from approximately 770 km^2 in 1975 to nearly 1,780 km^2 in 2014, (i.e., 2.3 times more). Percentile values from the most recent built-up extension that define flood exposure classes are $E_{25} = 9.2 \cdot 10^{-4}$ km^2, $E_{50} = 3 \cdot 10^{-3}$km^2, $E_{75} = 0.01$ km^2 and $E_{100} = 0.99$. A highly skewed distribution characterizes built-up values. Indeed, the majority of built-up grid cells presents extremely low built-up surfaces and <2.5% of grid cells has at least half of its surface built-up. We also analyze the total number of built-up grid cells, regardless of their built-up values (**Supplementary Figure 1**), and find that the 2014 value is three times more than the initial one in 1975, meaning that MDB experienced a remarkable conversion from natural land to built-up land. The growth rate of built-up areas is lower than that of the total number of built-up grid cells (i.e., 2.3 vs. 3), proving that new built-up locations are characterized by low built-up values, which are typically found within an urban sprawl context. In particular, when looking at the temporal evolution across flood exposure classes, different

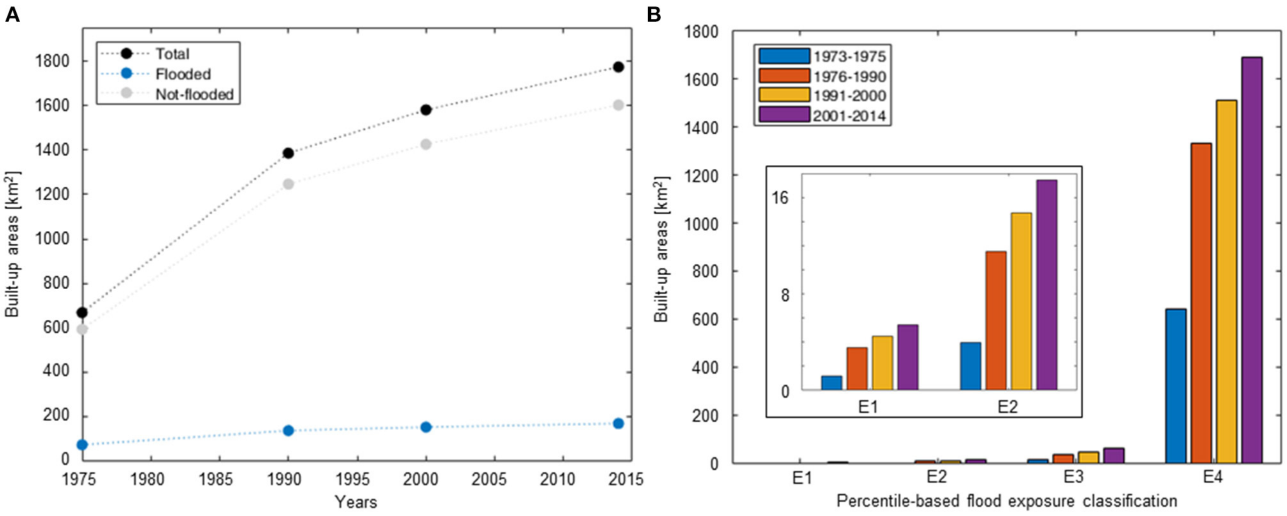


FIGURE 4 | Flood exposure assessment across the MDB from 1973 to 2014. **(A)** Temporal evolution of the actual maximum extension of built-up areas (in km^2), distinguishing between total (black), flooded (blue) and not-flooded (gray) locations. **(B)** Temporal evolution of the maximum extension of built-up areas (in km^2) in each flood exposure category (percentile-based classification). The inset shows a zoom for classes E1 and E2.

relative increments emerge. The relative increment of built-up surfaces is larger in E1 and E2 classes (i.e., 4.5 times), where built-up values are very small, rather than in E3 and E4 classes, characterized by 3.3 and 2.6 relative increments, respectively (**Figure 4B**).

We then disaggregate the maximum extension of built-up areas by distinguishing between flooded vs. not flooded built-up locations, as derived from inundation states based on cumulative sub-periods. The actual maximum extension of built-up area that was flooded at least once feebly increases in time from 1975 to 1990, then remains stable until 2014 (**Figure 4A**, blue line). The maximum extension of built-up area that did not experience any flood from 1973 to 2014 increases in time from 590 km^2 in 1975 to 1636 km^2 in 2014 (**Figure 4A**, gray line), showing the same pattern as the maximum extension of total built-up area. We then quantify the temporal evolution of built-up areas (total, flooded and not-flooded) by using a linear regression model. Average annual increments of built-up areas (in relative terms) are equal to 4.17%/year (R^2 = 0.91), 3.2%/year (R^2 = 0.90) and 4.29%/year (R^2 = 0.91) for total, flooded and not-flooded built-up areas, respectively. Our results confirm that flood exposure increased over the past four decades (Hu et al., 2018; Merz et al., 2021), despite the relative temporal invariance of flood hazard (see **Figure 3B**, cumulative sub-periods). Not-flooded built-up areas show larger relative annual increments compared to flooded built-up areas, which could possibly drive an increase in flood risk in the near future.

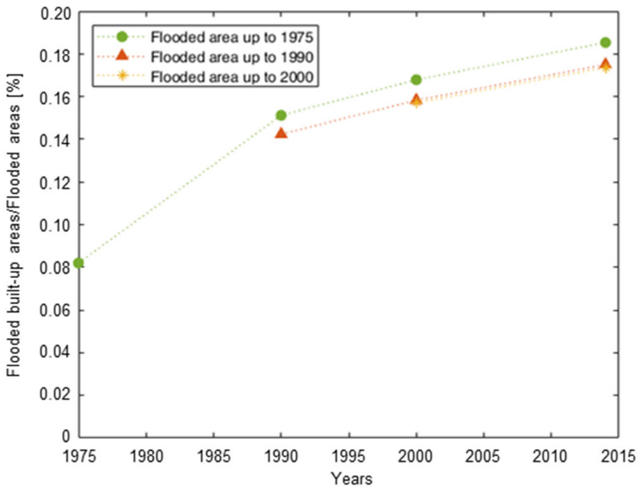


FIGURE 5 | Flood memory assessment across the MDB from 1973 to 2014: temporal evolution of the ratio between built-up areas in previously flooded areas and flooded areas (as %) up to 1975 (green circles), up to 1990 (orange triangles), up to 2014 (yellow stars).

To further support our results, we analyze the total number of people in the MDB, by using GHS-POP data from the GHSL dataset (Schiavina et al., 2019), also differentiating between flooded and not-flooded locations (**Supplementary Figure 2**). A similar overall pattern is found, thus supporting the use of built-up data as a proxy of human presence. Yet, minor differences between built-up and population datasets emerge. For instance, the marked increment of built-up areas from 1975 to 1990 (higher than subsequent increments) is not observed in population data. This could be explained by a large land conversion up to 1990, not accompanied by a comparable demographic trend. Also, this difference may depend on the intrinsic characteristics of the population dataset, where population data in 1975, 1990 and 2014 was derived from the NASA-CIENSIN Gridded Population of the World dataset (Center for International Earth Science Information Network (CIESIN), 2016) in 2010, by applying an exponential model and by using built-up data.

Since we observe a lower relative increment of built-up areas in previously flooded locations compared to not-flooded built-up areas (i.e., nearly 43% difference within the whole study period), we test if a flood memory behavior can be detected across the MDB. Despite the huge flood event in 1974, a large increment of built-up areas across locations affected by flood events was observed afterwards, thus revealing a very low flood memory behavior (green line in **Figure 5**). Indeed, historically several flood mitigation structures and water supply dams were built across MDB, which induced people to over rely on structural measures rather than flood risk awareness to reduce human exposure to floods (Cook, 2017). Conversely, when considering the flooded area up to 1990 and 2000, our results seem to suggest a stronger flood memory type behavior (orange and yellow lines in **Figure 5**), as shown by the lower increase of built-up areas.

We finally analyze long-term dynamics of flood risk (**Figure 6**), as categorized values following the percentile-based classification of flood hazard and exposure (**Figure 1**), and quantify flood risk as the overall extension of built-up areas flooded in a given sub-period. Regardless of flood risk classes, the areas at risk increase in time from 75.34 km^2 in 1975 to 136.77 km^2 in 2014. Despite the differences in terms of absolute values across risk classes, due to increasing percentages of built-up surface from R1 to R4 (see **Figure 1**), we find that R1, R2 and R3 classes are characterized by a hump-shaped temporal trend, with a peak in 2000 for R1 and in 1990 for R2 and R3. This outcome might suggest a reduction or a stabilization within the whole study period of the built-up area at risk, for low risk classes. Conversely, class R4 more than doubled its overall built-up area from 1975 to 2014, clearly indicating a marked increase of the areas at high flood risk. This increasing pattern is mainly influenced by increasing built-up areas, since flooded areas decreased in time.

CONCLUSIONS

In this study, we coupled a large-scale 2-D hydrodynamic model with remotely sensed built-up data to assess high-resolution long-term dynamics of flood risk and its components in the Murray-Darling Basin (MDB) within a 42-year long time period (from 1973 to 2014). In particular, we analyzed the frequency of occurrence of flood events and the temporal evolution of the maximum extension of flooded areas and built-up areas, the combination of which allows for a detailed assessment of flood exposure, flood memory and flood risk.

The following findings are of importance:

1. The maximum extension of flooded areas decreases in time, without revealing any significant geographical transfer of inundated areas across the study period. In particular, recurrently inundated gridded areas present an invariant pattern in time, whereas a remarkable reduction is typical of infrequently inundated locations.

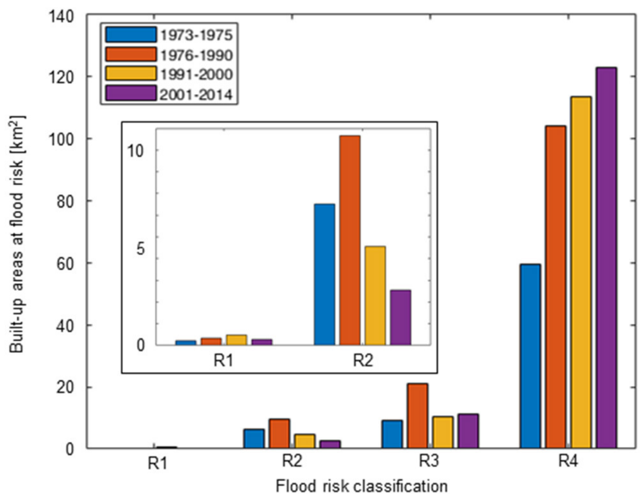


FIGURE 6 | Flood risk assessment, as categorized values derived from the percentile-based classification of flood hazard and exposure, across the MDB from 1973 to 2014: temporal evolution of the overall extension of built-up areas at risk in each flood risk class (in km^2), as defined in **Figure 1**. The inset shows a zoom for classes R1 and R2.

2. Our analysis shows that the maximum extension of built-up areas increased over the past four decades, even though only <2.5% of built-up grid cells shows at least half of their surface built-up. Larger annual increments of built-up surfaces (in relative terms) are typical of not-flooded built-up locations compared to flooded built-up areas.
3. When testing if previous flood events could have influenced the subsequent built-up expansion, we find a lower increase of built-up areas in previously flooded locations after 1990, which might suggest a flood memory pattern.
4. We finally combine flood hazard and exposure to assess categories of flood risk and find that areas classified at high risk for flood more than doubled their spatial extension during the past 40 years. Conversely, areas at lower risk for flood present an invariant or even decreasing trend in time.

Despite a few limitations (e.g., reservoir dynamics not included in the modeling framework, a simple method for assessing flood memory), our study supports the analysis of long-term dynamics of flood risk and its components at a large scale by combining different perspectives in an innovative way. More specifically, our results clearly point out that long-term dynamics of flood risk across the MDB are mainly driven by the temporal evolution of built-up areas, since flood hazard appears to be more or less invariant in time. Thus, adopting an urban development plan that explicitly incorporates strategies and actions for improving flood risk awareness would seem to be highly beneficial for reducing flood risk in the near future.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

GS performed hydraulic simulations. SC performed the analysis and wrote the first draft of the manuscript. AD wrote section 2-D hydrodynamic model for flood hazard assessment of the manuscript. All authors discussed the results, contributed to manuscript revision, conception and design of the study, read, and approved the submitted version.

ACKNOWLEDGMENTS

The authors gratefully thank Piera Colanzi for preliminary analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frwa.2021.797259/full#supplementary-material

REFERENCES

Aerts, J. C. J. H., Botzen, W. J., Clarke, K. C., Cutter, S. L., Hall, J. W., Merz, B., et al. (2018). Integrating human behaviour dynamics into flood disaster risk assessment. *Nat. Clim. Change* 8, 193–199. doi: 10.1038/s41558-018-0085-1

Barnes, J. (2017). The future of the Nile: climate change, land use, infrastructure management, and treaty negotiations in a transboundary river basin. *Wiley Interdiscip. Rev.* 8:e449. doi: 10.1002/wcc.449

Bates, P. D., Horritt, M. S., and Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two dimensional flood inundation modelling. *J. Hydrol.* 387, 33–45. doi: 10.1016/j.jhydrol.2010.03.027

Bernhofen, M. V., Trigg, M. A., Sleigh, P. A., Sampson, C. C., and Smith, A. M. (2021). Global flood exposure from different sized rivers. *Nat. Hazards Earth Syst. Sci.* 21, 2829–2847. doi: 10.5194/nhess-21-2829-2021

Bishop-Taylor, R., Tulbure, M. G., and Broich, M. (2018). Evaluating static and dynamic landscape connectivity modelling using a 25-year remote sensing time series. *Landsc. Ecol.* 33, 625–640. doi: 10.1007/s10980-018-0624-1

Bunn, S. E., Thoms, M. C., Hamilton, S. K., and Capon, S. J. (2006). Flow variability in dryland rivers: Boom, bust and the bits in between. *River Res. Appl.* 22, 179–186. doi: 10.1002/rra.904

Center for International Earth Science Information Network (CIESIN) (2016). Gridded Population of the World, Version 4 (GPWv4): Population Count, NASA Socioeconomic Data and Applications Center (SEDAC) [Dataset]. New York, NY: Columbia University.

Ceola, S., Laio, F., and Montanari, A. (2014). Satellite nighttime lights reveal increasing human exposure to floods worldwide. *Geophys. Res. Lett.* 41, 7184–7190. doi: 10.1002/2014GL061859

Cook, M. (2017). *The Australia Day Floods, January 1974. Environment & Society Portal, Arcadia, 15.* Munich: Rachel Carson Center for Environment and Society.

Corbane, C., Florczyk, A., Pesaresi, M., Politis, P., and Syrris, V. (2018). *GHS Built-Up Grid, Derived From Landsat, Multitemporal (1975-1990-2000-2014), R2018A.* Ispra: European Commission, Joint Research Centre (JRC). Available online at: http://data.europa.eu/89h/jrc-ghsl-10007 (accessed February 06, 2020).

de Moel, H., Aerts, J. C. J. H., and Koomen, E. (2011). Development of flood exposure in the Netherlands during the 20th and 21st century. *Global Environ. Change* 21, 620–627. doi: 10.1016/j.gloenvcha.2010.12.005

Domeneghetti, A., Carisi, F., Castellarin, A., and Brath, A. (2015). Evolution of flood risk over large areas: quantitative assessment for the Po river. *J. Hydrol.* 527, 809–823, doi: 10.1016/j.jhydrol.2015.05.043

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., and Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. *Adv. Water Resour.* 94, 87–102. doi: 10.1016/j.advwatres.2016.05.002

Duan, W., He, B., Nover, D., Fan, J., Yang, G., Chen, W., et al. (2016). Floods and associated socioeconomic damages in China over the last century. *Nat. Hazards* 82, 401–413. doi: 10.1007/s11069-016-2207-2

EU Floods Directive (2007). *Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the Assessment and Management of Flood Risks*. Copenhagen: European Environment Agency.

Facebook Connectivity Lab, and Center for International Earth Science Information Network (CIESIN) (2016). *High Resolution Settlement Layer (HRSL)*. Palisades, NY: Center for International Earth Science Information Network.

Gallant, A. J. E., Kiem, A. S., Verdon-Kidd, D. C., Stone, R. C., and Karoly, D. J. (2012). Understanding hydroclimate processes in the Murray-Darling Basin for natural resources management. *Hydrol. Earth Syst. Sci.* 16, 2049–2068. doi: 10.5194/hess-16-2049-2012

Garde-Hansen, J., McEwen, L., Holmes, A., and Jones, O. (2017). Sustainable flood memory: remembering as resilience. *Mem. Stud.* 10, 384–405, doi: 10.1177/1750698016667453

Grimaldi, S., Schumann, G. J. P., Shokri, A., Walker, J. P., and Pauwels, V. R. N. (2019). Challenges, opportunities and pitfalls for global coupled hydrologic-hydraulic modeling of floods. *Water Resour. Res.* 55, 5277–5300. doi: 10.1029/2018WR024289

Gupta, S. C., Kessler, A. C., Brown, M. K., and Zvomuya, F. (2015). Climate and agricultural land use change impacts on streamflow in the upper midwestern United States. *Water Resour. Res.* 51, 5301–5317. doi: 10.1002/2015WR017323

Heimhuber, V., Tulbure, M. G., and Broich, M. (2016). Modeling 25 years of spatio-temporal surface water and inundation dynamics on large river basin scale using time series of earth observation data, *Hydrol. Earth Syst. Sci.* 20, 2227–2250, doi: 10.5194/hess-20-2227-2016

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nat. Clim. Change* 3, 816–821. doi: 10.1038/nclimate1911

Hu, P., Zhang, Q., Shi, P. J., Chen, B., and Fang, J. Y. (2018). Flood-induced mortality across the globe: spatiotemporal pattern and influencing factors. *Sci. Total Environ.* 643, 171–182. doi: 10.1016/j.scitotenv.2018.06.197

IPCC (2012). "Glossary of terms," in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom; New York, NY, United States.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom; New York, NY, United States (in press).

Jonkman, S. N., and Kelman, I. (2005). An analysis of the causes and circumstances of flood disaster deaths. *Disasters* 29, 75–97. doi: 10.1111/j.0361-3666.2005.00275.x

Kam, P. M., Aznar-Siguan, G., Schewe, J., Milano, L., Ginnetti, J., Willner, S., et al. (2021). Global warming and population change both heighten future risk of human displacement due to river floods. *Environ. Res. Lett.* 16:044026. doi: 10.1088/1748-9326/abd26c

Kreibich, H., Seifert, I., Thieken, A. H., Lindquist, E., Wagner, K., and Merz, B. (2011). Recent changes in flood preparedness of private households and businesses in Germany. *Reg. Environ. Change* 11, 59–71. doi: 10.1007/s10113-010-0119-3

Lee, S., Kim, J.-C., Jung, H.-S., Lee, M. J., and Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea, Geomatics. *Nat. Hazards Risk* 8, 1185–1203, doi: 10.1080/19475705.2017.1308971

Lehner, B., and Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* 296, 1–22, doi: 10.1016/j.jhydrol.2004.03.028

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., et al. (2019). The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth Syst. Sci. Data* 11, 1385–1409. doi: 10.5194/essd-11-1385-2019

Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F. R., et al. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data* 3, 108–139. doi: 10.1080/20964471.2019.1625151

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geosci. Model Dev.* 10, 1903–1925. doi: 10.5194/gmd-10-1903-2017

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nat. Rev. Earth Environ.* 2, 592–609. doi: 10.1038/s43017-021-00195-3

Merz, B., Hall, J., Disse, M., and Schumann, A. (2010). Fluvial flood risk management in a changing world. *Nat. Hazards Earth Syst. Sci.* 10, 509–527, doi: 10.5194/nhess-10-509-2010

Murray-Darling Basin Authority (2020). *Basin Plan Annual Report*. Canberra, ACT: Australian Government.

Neal, J., Schumann, G. J.-P., and Bates, P. D. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resour. Res.* 48, 1–16,. doi: 10.1029/2012W. R.0 12514

NOAA - Earth Observation Group (2016). *Version 4 DMSP-OLS Nighttime Lights Time Series*. Available online at: http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed March 09, 2017).

Palacios-Lopez, D., Bachofer, F., Esch, T., Heldens, W., Hirner, A., Marconcini, M., et al. (2019). New perspectives for mapping global population distribution using world settlement footprint products. *Sustainability* 11:6056. doi: 10.3390/su11216056

Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrol. Earth Syst. Sci.* 16, 4143–4156. doi: 10.5194/hess-16-4143-2012

Penton, D., and Overton, I. (2007). "Spatial modelling of floodplain inundation combining satellite imagery and elevation models," in *Paper presented at the MODSIM 2007 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand CSIRO* (Clayton South, VIC).

Ridolfi, E., Albrecht, F., and Di Baldassarre, G. (2020). Exploring the role of risk perception in influencing flood losses over time. *Hydrol. Sci. J.* 65, 12–20. doi: 10.1080/02626667.2019.1677907

Rudari, R., Silvestro, F., Campo, L., Rebora, N., Boni, G., and Herold, C. (2015). *Improvement of the Global Flood Model for the GAR 2015*. Geneva: United Nations.

Samela, C., Troy, T. J., and Manfreda, S. (2017). Geomorphic classifiers for flood-prone areas delineation for data-scarce environments. *Adv. Water Resour.* 102, 13–28. doi: 10.1016/j.advwatres.2017.01.007

Sampson, C. C., Smith, A. M., Bates, P. B., Neal, J. C., Alfieri, L., and Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resour. Res.* 51, 7358–7381. doi: 10.1002/2015WR016954

Schiavina, M., Freire, S., and MacManus, K. (2019). *GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A*. Ispra: European Commission, Joint Research Centre (JRC). Availablee online at: http://data.europa.eu/89h/0c6b9751-a71f-4062-830b-43c9f432370f (accessed January 14, 2019).

Schumann, G. J.-P., Stampoulis, D., Smith, A. M., Sampson, C. C., Andreadis, K. M., Neal, J. C., et al. (2016). Rethinking flood hazard at the global scale. *Geophys. Res. Lett.* 43, 249–256, doi: 10.1002/2016GL070260

Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., de Moel, H., et al. (2016). FLOPROS: an evolving global database of flood protection standards. *Nat. Hazards Earth Syst. Sci.* 16, 1049–1061. doi: 10.5194/nhess-16-1049-2016

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res.* 116:G04021. doi: 10.1029/2011JG001708

Tanoue, M., Hirabayashi, Y., and Ikeuchi, H. (2016). Global-scale river flood vulnerability in the last 50 years. *Sci. Rep.* 6:36021. doi: 10.1038/srep36021

Tavares da Costa, R., Zanardo, S., Bagli, S., Hilberts, A. G. J., Manfreda, S., Samela, C., et al. (2020). Predictive modeling of envelope flood extents using geomorphic and climatic-hydrologic catchment characteristics. *Water Resour. Res.* 56:e2019WR026453. doi: 10.1029/2019WR026453

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., et al. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature* 596, 80–86, doi: 10.1038/s41586-021-03695-w

Viglione, A., Merz, B., Viet Dung, N., Parajka, J., Nester, T., and Blöschl, G. (2016). Attribution of regional flood changes based on scaling fingerprints. *Water Resour. Res.* 52, 5322–5340, doi: 10.1002/2016WR019036

Winsemius, H., Aerts, J., van Beek, L., Bierkens, M. F. P., Bouwman, A., Jongman, B., et al. (2016). Global drivers of future river flood risk. *Nat. Clim. Change* 6, 381–385, doi: 10.1038/nclimate2893

Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., and Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrol. Earth Syst. Sci.* 17, 1871–1892. doi: 10.5194/hess-17-1871-2013

Yamazaki, D., Kanae, S., Kim, H., and Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resour. Res.* 47:W04501. doi: 10.1029/2010WR009726

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.

ORIGINAL RESEARCH
published: 24 January 2022
doi: 10.3389/frwa.2021.742780

Exploring Dissolved Organic Carbon Variations in a High Elevation Tropical Peatland Ecosystem: Cerro de la Muerte, Costa Rica

Sánchez-Murillo Ricardo[1,2]*, Gastezzi-Arias Paola[3], Sánchez-Gutiérrez Rolando[2], Esquivel-Hernández Germain[2], Pérez-Salazar Roy[4] and Poca María[5]

[1] Department of Earth and Environmental Sciences, University of Texas at Arlington, Arlington, TX, United States, [2] Stable Isotopes Research Group and Water Resources Management Laboratory, Chemistry School, Universidad Nacional, Heredia, Costa Rica, [3] Doctorado en Ciencias Naturales Para el Desarrollo Costa Rica (DOCINADE), Instituto Tecnológico de Costa Rica and Laboratorio de Vida Silvestre y Salud, Vicerrectoría de Investigación de la Universidad Estatal a Distancia, San José, Costa Rica, [4] Laboratorio de Gestión de Desechos y Aguas Residuales, Escuela de Química, Universidad Nacional, Heredia, Costa Rica, [5] Instituto de Matemática Aplicada San Luis, Universidad Nacional de San Luis, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Grupo de Estudios Ambientales, San Luis, Argentina

Edited by:
François Birgand,
North Carolina State University,
United States

Reviewed by:
Jonghun Kam,
Pohang University of Science and Technology, South Korea
Christopher Osburn,
North Carolina State University,
United States

***Correspondence:**
Sánchez-Murillo Ricardo
ricardo.sanchez.murillo@una.cr

Specialty section:
This article was submitted to
Water and Built Environment,
a section of the journal
Frontiers in Water

Received: *16 July 2021*
Accepted: *14 December 2021*
Published: *24 January 2022*

Citation:
Ricardo S-M, Paola G-A, Rolando S-G, Germain E-H, Roy P-S and María P (2022) Exploring Dissolved Organic Carbon Variations in a High Elevation Tropical Peatland Ecosystem: Cerro de la Muerte, Costa Rica. Front. Water 3:742780. doi: 10.3389/frwa.2021.742780

Tropical peatlands are distributed mainly in coastal lowlands; however high elevation regions exhibit a large prevalence of small and fragmented peatlands that are mostly understudied. Artificial drainage of peatlands to expand the area of cattle farming, horticulture, and urbanization is increasing carbon losses to the atmosphere and streams worldwide. Here, we present an exploratory characterization of dissolved carbon optical properties in ombrotrophic peat bogs of the Talamanca range of Costa Rica, across an altitudinal gradient (2,400–3,100 m a.s.l.) during the rainy season. Dissolved organic matter (DOM) sources and decomposition processes were evaluated in the light of dissolved organic and inorganic carbon (DOC and DIC), optical properties, and major water chemistry. DOC concentrations ranged from 0.2 up to 47.0 mg/L. DIC concentrations were below 2 mg/L and $\delta^{13}C_{DIC}$ values indicated a mixture between soil organic matter, CO_2 in soil water, and to a lesser degree DIC derived from bacterial CO_2. Absolute fluorescence intensity of humic-like peaks was 6–7 times greater than fresh-like peaks across all sites. Fluorescence peak ratios coupled with the biological and humification indexes point to a greater relative contribution of recalcitrant soil-derived DOM. Excitation/Emission matrices denoted a high prevalence of humic and fulvic acids in the peat bogs, with moderate intensities in soluble microbial by-products-like and aromatic protein regions at three sites. Our data provides a baseline to underpin tropical carbon dynamics across high elevation peatlands.

Keywords: tropical high elevation peatlands, dissolved organic carbon, optical properties, carbon storage, climate variability

INTRODUCTION

Peatlands are a type of wetlands which are defined as having saturated soils, dense vegetation, anoxic conditions, and large deposits of partially decomposed organic plant material or peat (i.e., soil organic matter 30–50% in a 20–40 cm profile) (Yu, 2011; Page and Baird, 2016; Bourgeau-Chavez et al., 2018; Villa et al., 2019; Ribeiro et al., 2021). Peat accumulates as a result of the long-term imbalance between carbon production and decomposition (Hapsari et al., 2017; Bourgeau-Chavez et al., 2018). Tropical peatlands play an important role in the global carbon cycle due to their immense storage capacity, preserving between 469 and 694 Gt C in a relatively small area (90–170 Mha), equivalent to one third of the total global carbon pool (Müller et al., 2015; Gumbricht et al., 2017; Hapsari et al., 2017).

Tropical peatlands are distributed mainly in the river deltas and coastal plains, although mountainous areas also contain many small and fragmented peatlands (Hribljan et al., 2016; Hapsari et al., 2017; Lourençato et al., 2017; Silva et al., 2019). Mountainous peatlands offer a broad spectrum of ecosystem services, such as regulating stream discharge and lowland flooding, sediment and nutrient retention, carbon storage, and are home of a large number of endemic species (Kimmel and Mander, 2010; Ribeiro et al., 2021). Artificial peatland drainage may result in the net loss of the C storage function and in a large release of greenhouse gases to the atmosphere (e.g., CO, CO_2, and CH_4; Page and Baird, 2016).

Despite the recent research interest and new data available about the key role of peatland worldwide, the carbon quality (i.e., humic and fulvic acids, protein and microbial-like materials) and ecohydrological functioning of high elevation peatlands remain understudied in the tropics mainly due to sampling sites accessibility during the rainy season (Gandois et al., 2014; Hribljan et al., 2016; Pesántez et al., 2018).

Tropical peatlands, especially those in freshwater swamp forests, are mostly fed by partially decayed organic matter, which is often composed of recalcitrant plant litter that decomposes under acidic and anaerobic conditions at particularly slow rates (Bourgeau-Chavez et al., 2018; Cooper et al., 2019; Leng et al., 2019). In general, the decomposition process is regulated by local precipitation amount and seasonal perched water tables, which determine the shape and extent of the peatland surface and specify the ultimate stable morphology, and hence the carbon storage, within the adjacent network of rivers or canals (Villa et al., 2019). Therefore, a substantial reduction in rainfall may alter the peat surface dynamics, which in combination to warm temperatures may lead to higher rates of mineralization and degassing of the partly decomposed organic matter within peatlands (Li et al., 2007; Fenner and Freeman, 2011; Lourençato et al., 2017; Leng et al., 2019).

Dissolved organic carbon (DOC) in peatlands can originate from several sources, including recent photosynthetic by-products, decomposition, and dissolution products within the peat column (Moore et al., 2013). DOC in peatlands reflects the carbon source characteristics and transformations, namely vegetation, peat/soil respiration, and microbial activity (Cooper et al., 2010; Marx et al., 2017). The advent of cost-effective laboratory techniques, such as absorbance and fluorescence, has promoted the use of dissolved organic matter (DOM) as an environmental tracer in many ecosystems (Cook et al., 2017; Dalmagro et al., 2017; Park and Snyder, 2018; Chaves et al., 2020).

Contrary to boreal and temperate peatlands where seasons control freezing and thawing, as well as microbial activity, DOM composition in wet tropical peatlands is mainly shaped by the large rainfall variability (Bispo et al., 2016). For example, during drought periods perched water tables are lowered, which in turn produces oxic conditions and stimulates phenol oxidase enzymes, therefore reducing the overall concentration of polyphenolics and their inhibitory effect on DOC decomposition *via* hydrolase enzymes (Ritson et al., 2017 and references therein).

Natural drainage of these systems are then important pathways for terrestrial carbon export in the forms of DOC, particulate organic carbon (POC), and dissolved inorganic carbon (DIC) (Kiew et al., 2018; Waldron et al., 2019). In some regions such as Southeast Asia, however, increasing human disturbances of tropical peatlands by fire and artificial drainage for agriculture have resulted in emissions of stored carbon at very fast rates (Page et al., 2002; Couwenberg et al., 2010; Itoh et al., 2017; Cooper et al., 2019). Therefore, an assessment of the stability of these tropical carbon stocks across different ecosystems is necessary to estimate the potential effects of increasing anthropogenic disturbances (Drake et al., 2019; Ribeiro et al., 2021).

In Central America, examples of these high mountain peatlands or bogs, known as "Turberas de Altura" (Gómez, 1986) are found between 2,400 and 3,100 m a.s.l. along the continental divide of the Talamanca range in Costa Rica (Brak et al., 2005; Corrales Ulate, 2018; **Figure 1**). These peat-forming bogs are restricted to small and poorly drained depressions subject to seasonal flooding (during the rainy season from May to November; Amador, 1998), and with an area just over 235 ha (Jiménez, 2016). In this region, soil organic carbon ranged from 31 up to 43% in the first 50 cm of the soil profile (Corrales Ulate, 2018). Given the limited information and current anthropogenic pressure across mountainous ecosystems of Costa Rica (Corrales Ulate, 2018; Stan and Sanchez-Azofeifa, 2019), there is an urgent need to understand the high elevation peatlands potential for carbon storage and export in a changing climate and land use scenario.

Here we present an exploratory characterization of DOM optical properties in peat bogs of the Talamanca range across an altitudinal transect (from 2,400 to 3,100 m a.s.l.) during a complete wet season. DOM sources and decomposition processes were evaluated with respect to DOC concentrations, optical properties, and major water chemistry composition. We hypothesize that water depth, temperature seasonality, and solar radiation exert a large control on DOC properties within the high elevation peat bogs.

Our data provides a baseline to underpin carbon dynamics of high mountain peatlands in the Central American region, which in turn may help to allocate efforts and resources for conservation strategies under the Payment for Environmental Services (i.e., payment for forest conservation and ecosystem services) scheme

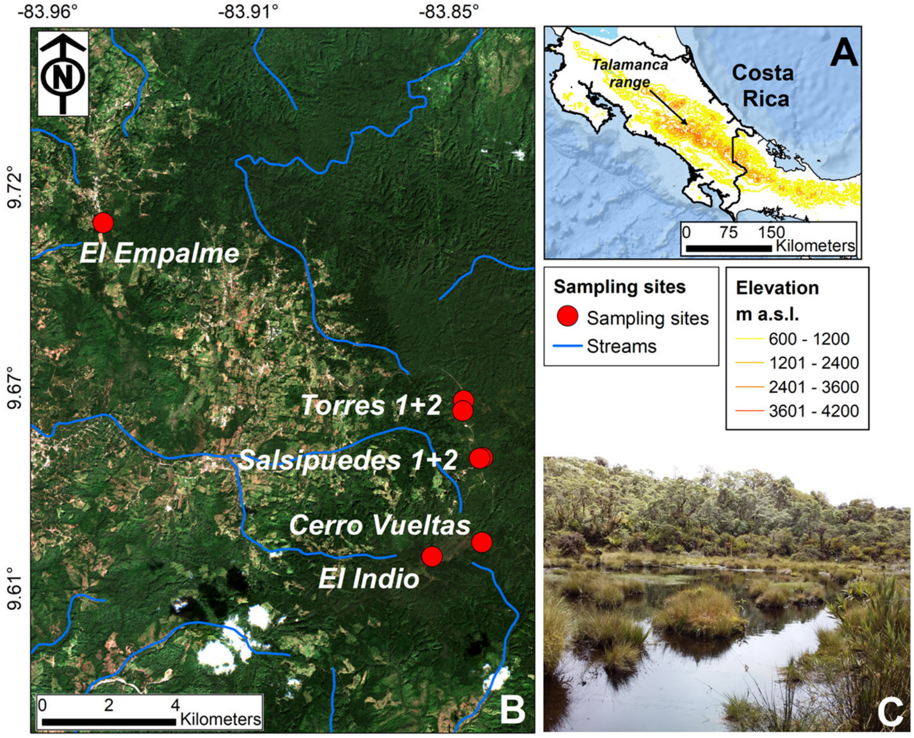


FIGURE 1 | (A) Talamanca range overview and elevation gradient of Costa Rica. **(B)** Landsat image of the peatland study area (also known as *Turberas de Altura*) within the high elevation Páramo ecosystem, including site locations and main streams. **(C)** Typical peatland condition throughout the wet season. El Empalme y El Indio represent the lowest and highest sampling sites, respectively.

(Arriagada et al., 2015; Berbés-Blázquez et al., 2017; Wallbott et al., 2019) across these sensitive ecosystems.

STUDY AREA

Study sites are located in the Pacific domain of the Talamanca range, Costa Rica (**Figure 1A**). During the last glacial maximum, this region was glaciated with an ice cap of roughly 5 km^2; the latter is evidenced by smoothed, grooved, and channeled andesite bedrock surfaces (Lachniet and Seltzer, 2002). Long-term weathering has facilitated the formation of concave landforms, which in turn facilitates water accumulation in such topographic depressions. These peat-forming bogs are found between 2,400 and 3,100 m a.s.l. and cover small areas (up to 0.5 ha), whereby surface drainage is relatively poor.

Meteorological conditions across this gradient are fairly similar throughout the wet season. Water table fluctuation is mostly controlled by the seasonal rainfall input with ~80% of the annual rainfall falling between May and November (**Figure 2**). The greatest variation in temperature and relative humidity was observed during the dry season (December–April, **Figure 2**). During the wet season, rainfall input leads to saturated soils and anoxic conditions, peat accumulation, and water acidification. In general, solar radiation effective hours range from 4 to 8 per day during the wet and dry seasons, respectively. More than 100 species of forbs, grasses, and sedges have been reported from the Páramo peat bogs of Costa Rica (Gómez, 1986). **Table 1** summarizes the main vegetation and morphological characteristics of the study sites obtained from field observations.

METHODS

Sample Collection

Seven highland peat bogs were studied during the 2019 wet season (**Figure 1B**) (7 monthly sampling campaigns). Peat bog samples for the analysis of DOC/DIC, optical properties, stable isotopes, and major ions were collected during the wet season, from the beginning of May through the end of November, 2019. Site accessibility and water storage within the peatlands were key factors to select the sites (**Figure 1C**).

Carbon and major ions samples were collected on monthly basis in pre-cleaned 250 mL glass amber bottles (including oven-heated treatment) covered with aluminum foil to prevent samples from solar/light radiation. Stable isotopes samples were collected in 50 mL pre-cleaned HDPE bottles with plastic inserts to prevent evaporation. Sampling bottles were rinsed at least three times with the peatland water before collection. Rainfall monitoring for stable isotopes analysis was conducted using a passive Palmex collector (daily collection; $N = 103$) (Gröning et al., 2012; Esquivel-Hernández et al., 2021). During transportation and after filtering, all samples were stored at 5°C until analysis. Electrical conductivity (EC, µS/cm), pH, turbidity (NTU), and *f*DOM

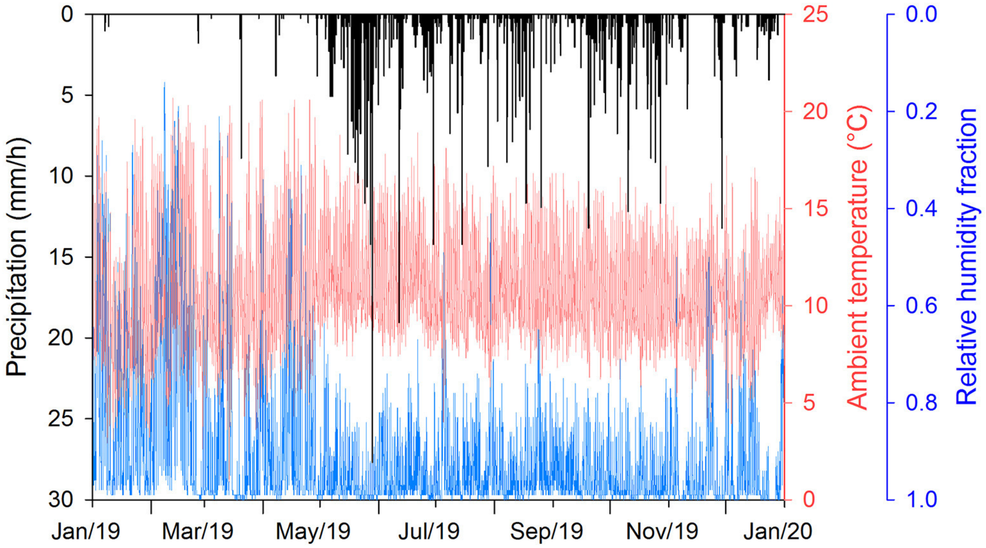


FIGURE 2 | Study area meteorological conditions during 2019. Black bars denote rainfall in mm/h. Red and blue lines correspond to hourly ambient temperature and relative humidity fraction, respectively.

TABLE 1 | Summary of main vegetation and morphological characteristics of the study sites.

Study sites	Elevation (m a.s.l.)	Main characteristics
El Empalme	2,400	This site presents heterogeneous vegetation of secondary forest mixed with pasture and shrubs, due to human activities (i.e., cattle farms, crops). In the rainy season, this peat bog forms small intermittent ponds of little depth (water column depth <60 cm). In the dry season, the soils remain between field capacity and saturation. Tadpoles and aquatic insects have been observed in these ponds.
Torres 1-2	2,600	The vegetation is represented by a continuous strip of montane forest with scattered arborescent ferns, Iridaceae, Ciperaceae, and mosses covering the ground. These peat bogs are permanent and shallow (water column depth <60 cm) with visible draining channels covered by herbaceous vegetation. In the dry season, the soils remain between field capacity and saturation.
Salsipuedes 1-2	2,800	Vegetation is characterized by montane forest, with tree ferns, mosses, Ciperaceae (grass-like), and bromeliads. These peat bogs are comprised of small, shallow lagoons during the rainy season. In the dry season, the soils remain between field capacity and saturation.
Cerro Vueltas and Indio	3,100	The vegetation is represented by secondary subalpine forest, and Páramo grasslands, with the presence of arborescent ferns, mosses, bromeliads, and Chusquea. Both peat bogs are mostly seasonal. In the dry season, the soils remain between field capacity and saturation.

(QSU; quinine sulfate units) were recorded using a pre-calibrated multi-parameter sonde EXO1 (YSI Inc., USA) immediately after sample arrival at the laboratory.

Water Stable Isotopes Analysis

Samples were analyzed at the Stable Isotopes Research Group laboratory at the Universidad Nacional (Heredia, Costa Rica) using an IWA-45EP water analyzer (Los Gatos Research, Inc., California, USA) with a precision of ±0.5‰ for δ^2H and ±0.1‰ for δ^{18}O (1σ; 8 injections). Stable isotope compositions are expressed as $\delta^{18}O$ or $\delta^2H = (Rs/Rstd - 1)\cdot 1{,}000$, where R is the $^{18}O/^{16}O$ or $^2H/^1H$ ratio in a sample (s) or standard (std) and reported in the delta-notation (‰) relative to V-SMOW/SLAP scale. The instrument accuracy was assessed with a combination of in-house and primary international water standards (SMOW and SLAP). Deuterium excess was calculated as $d\text{-excess} = \delta^2H - 8\cdot\delta^{18}O$ (Dansgaard, 1964).

DOC Analysis

DOC contents were measured using a TOC analyzer (Aurora 1030W, OI Analytical) following the Heated-Persulfate Oxidation Method 5310C (APHA AWWA, 2005). Each water sample was filtered using 0.45 µm PTFE filter into 100 mL pre-cleaned vials. A standard curve was conducted using six calibration solutions (i.e., stock solution of anhydrous primary-standard-grade potassium biphthalate, $C_8H_5KO_4$) ranging from

0 to 100 mg C L^{-1} (Method SM5310 C; APHA AWWA, 2005). The quantification and detection limits were 0.05 and 0.03 mg C L^{-1} (Sánchez-Murillo et al., 2019).

$\delta^{13}C_{DIC}$ Analysis and Total Alkalinity

Filtered aliquots (0.45 µm cellulose filter) (6 mL) were pipetted into Exetainer® vials. An automated DIC sample preparation system (Picarro AutoMate FX system, USA) was utilized to inject a 10% phosphoric acid solution into each sealed vial to liberate CO_2 from the sample. Additionally, a stream of dry and ultra-pure nitrogen was bubbled through the acidified solution to flush the released CO_2 from the vial headspace. The CO_2 was captured into gas sampling bags of a Picarro Liaison^{TM} Universal Interface before being analyzed using Cavity Ring Down Spectroscopy (CRDS, Picarro G2201-*i*). After each measurement, the instrument and the gas sampling bag were purged with fully dry and ultra-pure nitrogen between successive DIC/$\delta^{13}C$ measurements. The corresponding uncertainty for $\delta^{13}C$ in DIC is ±0.1‰ (1σ). Calibration was done using the following standards: University of McGill, Canada: CO_2 mixing ratio: 1,553 ppmv, $\delta^{13}CO_2 = -43.15‰$, Heredia's compressed air: CO_2 mixing ratio: 419.1 ppmv, $\delta^{13}CO_2 = -10.09‰$, NOAA gas standard: CO_2 mixing ratio: 394.85 ppmv, $\delta^{13}CO_2 = -8.292‰$. Stable isotope compositions are expressed as $\delta^{13}C = (Rs/Rstd - 1)\cdot 1{,}000$, where R is the $^{13}C/^{12}C$ ratio in a sample (s) or standard (std) and reported in the delta-notation (‰) relative to VPDB scale (Craig, 1957).

The total alkalinity of each sample (reported as $CaCO_3$) was also calculated from the average CO_2 concentration measured during the isotope analysis. The CO_2 concentration measured by the CRDS was standardized against Na_2CO_3 solutions (Sigma Aldrich, >99.0%, 5–100 mg/L). These standard solutions were analyzed following the same procedure described above. Given the small volume of sample that is analyzed by this method, only those water samples with a total alkalinity >10 mg/L were quantified.

Major Ions Analysis

Major ions samples were filtered using 0.45 µm PTFE and/or PVDF filter into 5 mL HDPE pre-cleaned poly vials. Ion chromatography (Thermo Scientific ICS-5000+, CA, USA) was used to analyze lithium, ammonium, sodium, potassium, magnesium, calcium, bromide, chloride, fluoride, nitrite, nitrate, and sulfate [DL (mg/L): 0.011, 0.035, 0.64, 0.20, 0.19, 0.23, 0.50, 0.23, 0.18, 0.26, 0.22 and 0.20, respectively]; to ensure the quality of the analysis, procedural blanks and recovery standards were performed in each batch of samples.

DOC Optical Properties

The water column of these shallow wetland ponds may be completely anoxic but due to sampling constraints we were unable to confirm that water samples were kept anoxic prior to measurement of optical properties.

DOC excitation-emission matrices (EEMs) were measured (in absolute fluorescence intensity) using a FP-8300 spectrofluorometer (Jasco, Tokyo, Japan) equipped with double monochromators both at the excitation and the emission sides. Excitation wavelength were set up in the range of 200–450 nm with sequential increments of 5 nm with an integration time of 0.5 s and increments of 1 nm in the emission wavelength for the range of 280–550 nm. Spectral corrections for the EEMs were done with sample blanks that were carried out using deionized water type I in each batch of samples under the same conditions to eliminate water Raman peaks (Huguet et al., 2009).

Chromophoric dissolved organic matter (DOM) absorbance at 254.7 nm (a_{254}, m^{-1}) was measured (Method 5910B; APHA AWWA, 2005) using UV-spectrophotometry (Jasco VP-750, Tokyo, Japan). For all the samples, the absorbance of a blank sample prepared with deionized water type I was measured as a spectral correction using quartz cells of 1 cm path length. EEMs diagrams are presented following Chen et al. (2003) wavelength boundaries (in absolute fluorescence intensity). It is important to highlight that the lack of a proper inner filter effect correction limited the comparative potential of our exploratory results with other studies.

Statistical and Optical and Data Analysis

EEMs were analyzed as a qualitative tool for assessing the principal sources of DOC in the water. The well-known Coble peaks (Coble, 1996) (in absolute fluorescence intensity), the fluorescence index (FI) (Mcknight et al., 2001), the humification index (HIX) (Zsolnay et al., 1998) and the autochthonous biological activity index (BIX) (Huguet et al., 2009) were calculated using the staRdom package (Pucher et al., 2019). Fluorescence measurements (including ratios) are not corrected for the strong absorption of these samples, so our results must be considered carefully in comparison with other studies. Nonetheless, the patters we observed are informative for this preliminary investigation into DOM properties in these understudied systems.

Spearman (1904) correlation coefficients and principal component analysis (PCA) were applied to further identify common parameters that are influencing dissolved carbon optical properties. Data below quantification limits (QL) were transformed to tie ranks before performing multivariate analysis (Helsel, 2012). Principal component analysis was carried out through the FactoMineR package (Lê et al., 2008), the data were scaled for unit variance in order to avoid some variables become dominant because of their intrinsic differences in units and variability; while the number of final dimensions were not constrained. In order to evaluate differences between parameters within each sampling site or sample campaign, we compute a one factor permutation test (Blair et al., 1994; Good, 2000), where the results that have been obtained by parameter (e.g., BIX) for each factor (e.g., sampling site) were shuffled 5,000 times doing a resampling procedure to estimate different means.

In addition, an exploratory correlation matrix was conducted using the corr R package (Friendly, 2002; R., Core Team, 2019; Wei and Simko, 2021). All variables exhibiting collinearity were removed prior the analysis. Only significant Spearman correlations ($\alpha = 0.05$) were visualized (see **Supplementary Figure 1**).

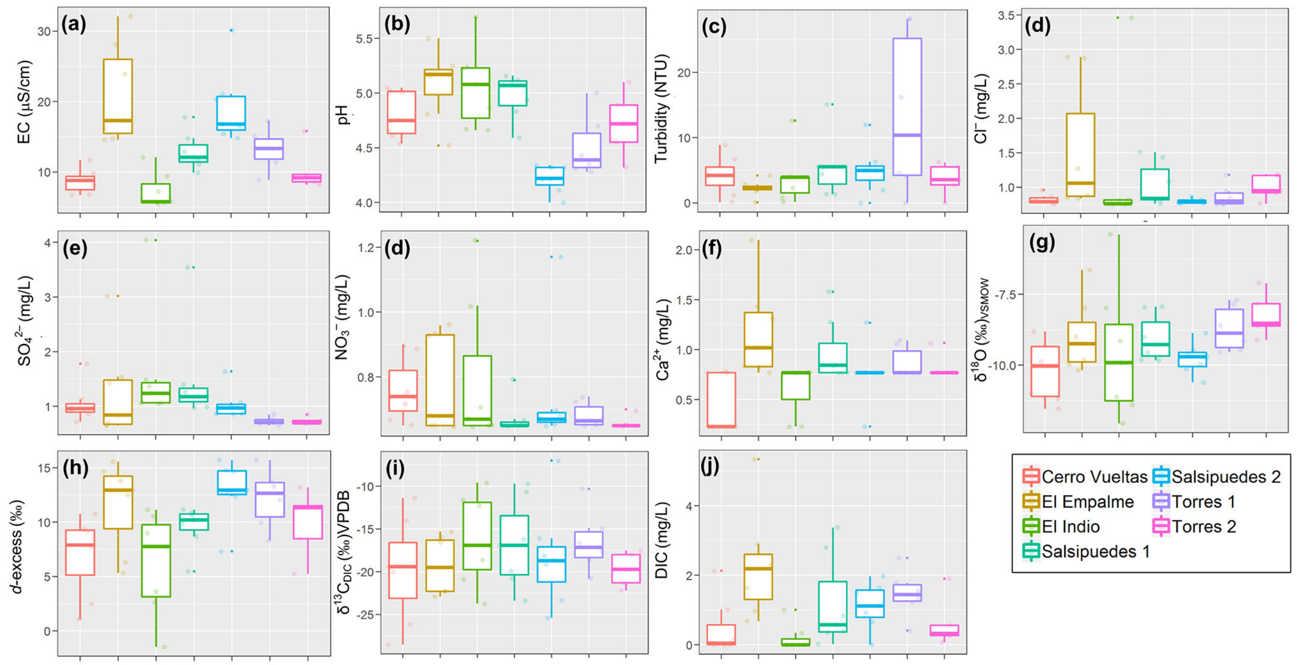


FIGURE 3 | Box plots (with individual data points) of physico-chemical (EC, pH, turbidity), water stable isotopes ($\delta^{18}O$ and *d*-excess), major ions (Cl^-, SO_4^{2-}, NO_3^-, Ca^{2+}), dissolved inorganic carbon (DIC), and carbon stable isotopes ($\delta^{13}C_{DIC}$). Sampling sites are color-coded and plots include 25th, 75th, median, and outliers. El Empalme and El Indio represent the lowest and highest sampling sites, respectively (see **Figure 1**).

RESULTS AND DISCUSSION

Peat Bogs Water Chemistry

Figure 3 shows a summary of the main water chemistry characteristics across the study sites. Peat bogs exhibited low EC in the range of precipitation (from 5 to 35 µS/cm) with acidic conditions between 4 and 6 pH units (**Figures 3a,b**). Major ion concentrations were relative low (Cl^- < 3.5 mg/L; SO_4^{-2} < 4 mg/L; NO_3^- < 1.2 mg/L; Ca^{+2} < 2 mg/L), suggesting slow rock weathering rates and minimal anthropogenic effects (Markham and Otárola, 2021) (**Figures 3d–f**). Turbidity averaged ~5 NTU across the hydrological year, with the exception of Torres 1 site (up to 30 NTU) (**Figure 3c**). Based on the water chemistry characteristics, the study sites are defined as ombrotrophic bogs (i.e., mainly fed by precipitation) (Lindsay, 2016).

DIC concentrations were below 2 mg/L and $\delta^{13}C_{DIC}$ compositions indicated a mixture between soil organic matter, soil CO_2, CO_2 in soil water, and in less degree DIC derived from bacterial CO_2 (mean $\delta^{13}C_{DIC} = -18.0 \pm 4.7‰$) (Pawellek and Veizer, 1994; Horgby et al., 2019) (**Figures 3i,j**). Water stable isotopes indicated a clear meteoric origin (Esquivel-Hernández et al., 2018, 2019, 2021) for this high elevation Páramo ecosystem, with $\delta^{18}O$ values ranging from −5.4 to −12.1‰, with a mean value of −9.2 ± 1.4‰. Similar to high elevation glacial lakes in this region (Esquivel-Hernández et al., 2018), the shallow peat bogs are also exposed to evaporation during the dry season (**Figures 3g**, **4**). Deuterium excess values indicated evaporation losses during the dry season (down to −1.5‰) and moisture recycling during the rain events of the wet season (up to +15.7‰) (Sánchez-Murillo et al., 2020) (**Figure 3h**). During the dry season (December to April), the strengthening of the easterly trade winds (Amador, 1998) results in low intensity horizontal rainfall events (Bruijnzeel et al., 2010), which in turn maintain the peat bogs near saturation; while water logging occurs mainly from mid-May to the end of November.

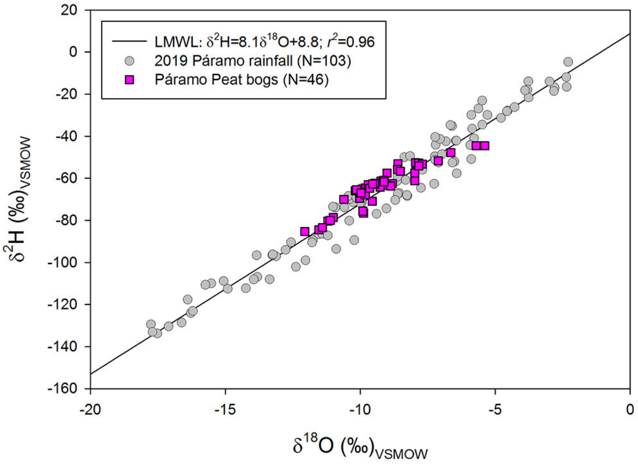


FIGURE 4 | Dual isotope diagram including peat bogs (pink squares) and rainfall (gray dots) during the 2019 hydrological year at the Páramo ecosystem. The local meteoric water line (LMWL) during 2019 (black line) is included as reference (Esquivel-Hernández et al., 2021).

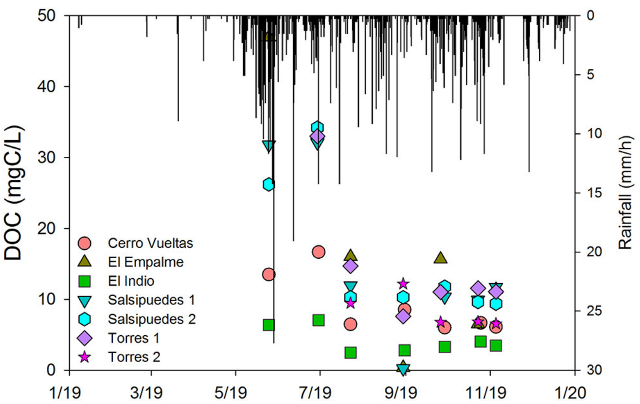


FIGURE 5 | DOC (mgC/L) temporal variation. Sampling sites are symbol and color-coded. El Empalme and El Indio represent the lowest and highest sampling sites, respectively (see **Figure 1**).

Dissolved Carbon Content and Optical Properties

Figure 5 shows a temporal variation of DOC across all sites vs. rainfall amounts. During the wet season (mid-May to November), rainfall inputs are critical to sustain saturated soils and maintain anoxic conditions, which in turn exert a large control on carbon storage and humification processes. At the end of the dry season (April to mid-May), DOC concentrations were relative high across almost all the sites (up to ∼50 mg/L). DOC concentrations were stabilized below ∼15 mg/L as the rainy season progressed. Cerro Vueltas and El Indio sites (highest elevation sites; **Figure 1**) exhibited a rather uniform DOC pattern from July to November, with DOC values below 10 mg/L.

Figure 6 summarizes the dissolved carbon optical properties throughout the hydrological year. In general, ∼91% of the dissolved carbon in the peat bogs corresponded to organic carbon. DOC concentrations ranged from 0.2 up to 47.0 mg/L (**Figure 6b**). Páramo DOC concentrations are relatively high when compared to other tropical water reservoirs. For example, Osburn et al. (2018) and Sánchez-Murillo et al. (2019) reported stream DOC concentrations ranging from 7.65 up to 11.52 mg/L, respectively, during storm events in swamp primary and secondary forests of the Caribbean slope of central Costa Rica. Williams et al. (2001) and Pesántez et al. (2018) reported mean DOC concentrations in Páramo wetlands and natural forests of the Ecuadorian Andes of 23.0 and 11.0 mg/L, respectively. Dalmagro et al. (2017) found significant DOC differences between the dry (∼2 mg/L) and wet (∼6 mg/L) seasons within the large Brazilian Pantanal wetland. Similarly, the authors reported greater DOC concentrations in forested areas than pasture regions. A global meta-analysis of DOC concentrations reported a mean value of 3.87 mg/L across South America river basins (Chaplot and Mutema, 2021).

El Indio site consistently reported low DOC concentrations and humic-like intensities, whereas protein-like intensity peaks were stronger than the rest of the peak bogs. Large DOC concentrations at El Empalme site (lower elevation and near anthropogenic influence, see **Figure 1**) were consistently linked to humic-like intensity peaks C, A, and M (i.e., recalcitrant material) (Coble, 1996; Hansen et al., 2016) (**Figures 6e–g**). Protein-like intensity peaks (B and T) (better refer as fresh-like material) were particularly lower in Salsipuedes 2 and Torres 1 sites (**Figures 6h,i**).

$SUVA_{254}$ is often related to hydrophobic organic acid fraction and constitutes a useful proxy of aromatic content and molecular weight of DOM (Weishaar et al., 2003). $SUVA_{254}$ values varied from 0.02 up to 8.72 L mgC^{-1} m^{-1}, with Salsipuedes 2 and Torres 1-2 sites exhibiting greater aromatic content than the rest of the peat bogs (**Figures 6a–c**). Despite the pristine nature of these high elevation peatbogs, it is important to remark that re-oxygenation of any reduced iron could lead to elevated absorbance (and thus $SUVA_{254}$ values) not linked to DOM absorbance (Poulin et al., 2014).

In general, the intensity of humic-like peaks was 6–7 times greater than fresh-like peaks across all sites. C:T ratios (an indication of the amount of degraded vs. fresh-like material) were greater at Salsipuedes 2 and Torres 1 (**Figure 6j**). Overall C:T ratios averaged 10.9, indicating a greater relative contribution of soil-derived DOM (Hansen et al., 2016). C:A ratios (an indication of the amount of humic-like vs. fulvic-like material) were <1 at three sites (Cerro Vueltas, El Indio, and Torres 2) indicating more prevalence of fulvic-like material (**Figure 6k**). C:M ratios at Salsipuedes 2 and Torres 1-2 sites exhibited greater values than the rest of the peat bogs, which in turn may suggest larger compositions of diagenetically altered material (microbial activity in sediments and soils) (**Figure 6l**).

The biological index (BIX) (an indication of autotrophic activity; Huguet et al., 2009) was below 1 across all sites (**Figure 6o**) indicating that DOM in the peat bogs was not recently produced. The latter clearly correspond to relative high values of humification (HIX; an indicator of humic substance content; Zsolnay et al., 1998) between 0.76 and 0.99. HIX values in natural waters range from 0.6 to 0.9 (Hansen et al., 2016) (**Figure 6n**). However, Dalmagro et al. (2017) have reported mean HIX values >10 whithin the large Brazilian Pantanal wetland. El Indio site exhibited the lower HIX values corresponding with low intensities in peaks C, A, and M (**Figure 6n**). The contribution of terrestrial to microbial sources to the DOM pool was evaluated with the fluorescence index (FI) (Mcknight et al., 2001). In natural water and tropical wetlands FI values ranged between 1.2 and 1.8 (Hansen et al., 2016; Dalmagro et al., 2017). In the peat bogs, FI values varied between 1.6 and 2.0, with a prevalence of microbial sources at Cerro Vueltas, El Indio, Salsipuedes 1, and Torres 2. FI values at Salsipuedes 2 and Torres 1 indicated a tendency toward more terrestrial DOM derived sources (**Figure 6m**).

Excitation/Emission Fluorescence Matrices (EEMs)

Based on the EEMs regions defined by wavelength boundaries (Chen et al., 2003 and references therein), El Empalme site exhibited strong intensities in a region dominated by humic acid-like substances (Region V), followed by a moderate intensity in

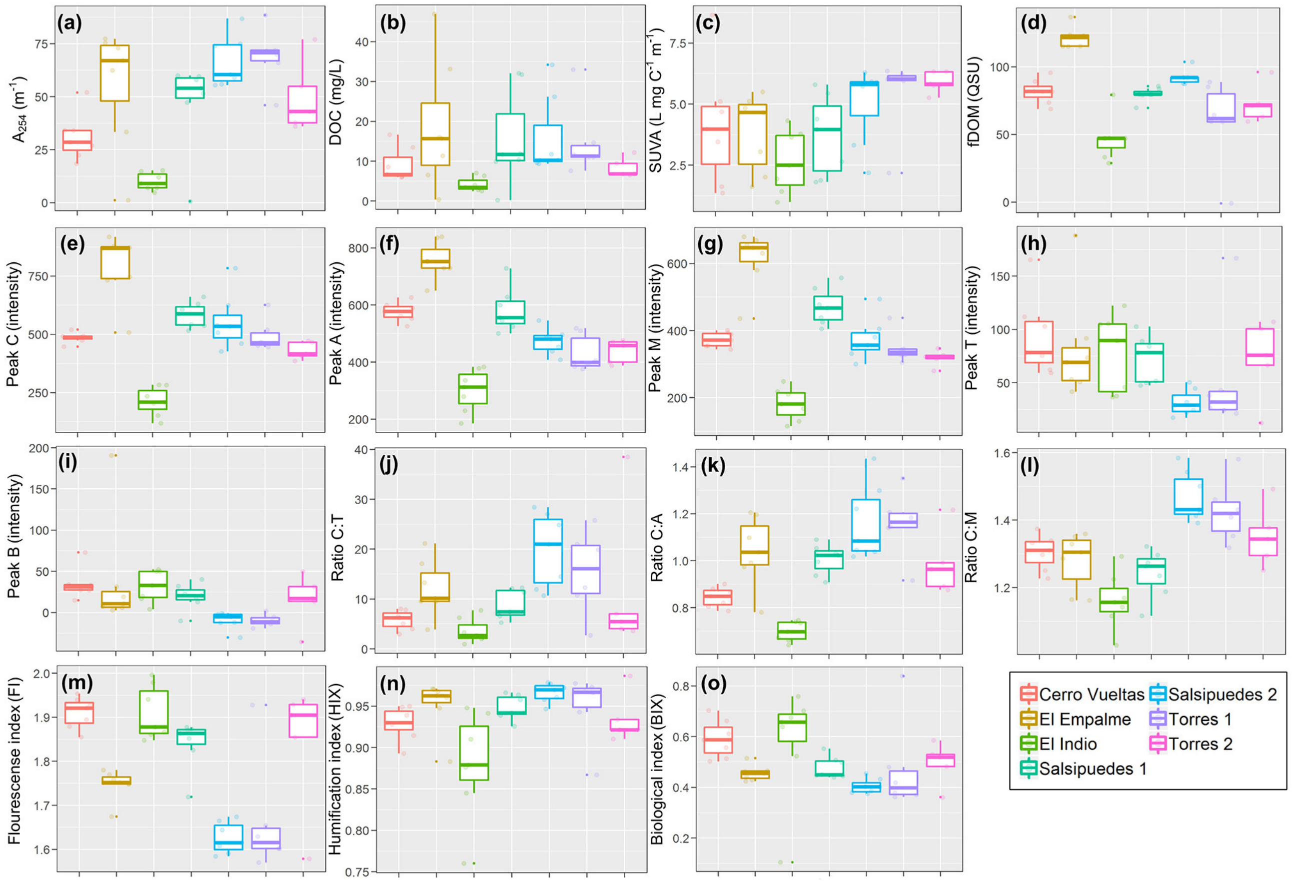


FIGURE 6 | Box plots (with individual data points) of dissolved and optical organic carbon properties, including: dissolved organic carbon (DOC), a_{254}, *f*DOM, SUVA index, Coble's peaks (C, A, M, T, and B; in absolute fluorescence intensity), and ratios (C:T; C:A; and C:M), fluorescence index (FI), humification index (HIX), and biological index (BIX). Sampling sites are color-coded and plots include 25th, 75th, median, and outliers. El Empalme and El Indio represent the lowest and highest sampling sites, respectively (see **Figure 1**).

the fulvic acid-like region (Region III), with minimal influence of protein or microbial by-product like contents (**Figure 7**). This is supported by strong C, A, and M peaks and high humification processes from terrestrial derived DOM (**Figures 6e–g**).

In contrast, El Indio and Torres 2 sites reported moderate intensities in Regions V-III, with also moderate signals in Regions I-II (i.e., aromatic proteins), corresponding to high FI and BIX values, and the lower HIX values. EEMs denoted a high prevalence of humic and fulvic acids in the peat bogs, with moderate to low intensities in soluble microbial by-product-like and aromatic proteins regions at Cerro Vueltas, El Indio, and Torres 2 (**Figure 7**).

Exploring Water Chemistry and Dissolved Carbon Relationships

PCA dimensions 1 and 2 together explained 68.8% of the total variance in the data set (**Figure 8**). Dimension 1, which explained 48.5% of the total variance, showed strong positive loadings for variables associated with humification processes (e.g., peaks C, A, and M; ratios, HIX, DOC, and a_{254}) and thus, fungal degraded by-products. On the other hand, dimension 2 which explained 20.4% of the total variance, showed positive loadings for parameters related with protein-like materials (peaks T and B), pH, and microbial derived DOM.

The PCA output also allowed the visualization of sampling sites within the framework of dimensions 1 and 2. El Empalme exhibited a high positive score on the dimension 1 axis, related with the presence of humic-like intensity peaks. Cerro Vueltas site showed a low positive score on the dimension 2, associated with microbial derived indexes (FI and BIX). Salsipuedes 2 and Torres 1 were related to humification ratios, whereas El Indio and Torres 2 were not related with a clear pattern between the PCA dimensions. Salsipuedes 1 was associated with moderate positive scores to humic-like and protein-like peaks.

CONCLUSIONS

Our study revealed that fluorescence intensity of humic-like peaks was 6–7 times greater than fresh-like peaks across all

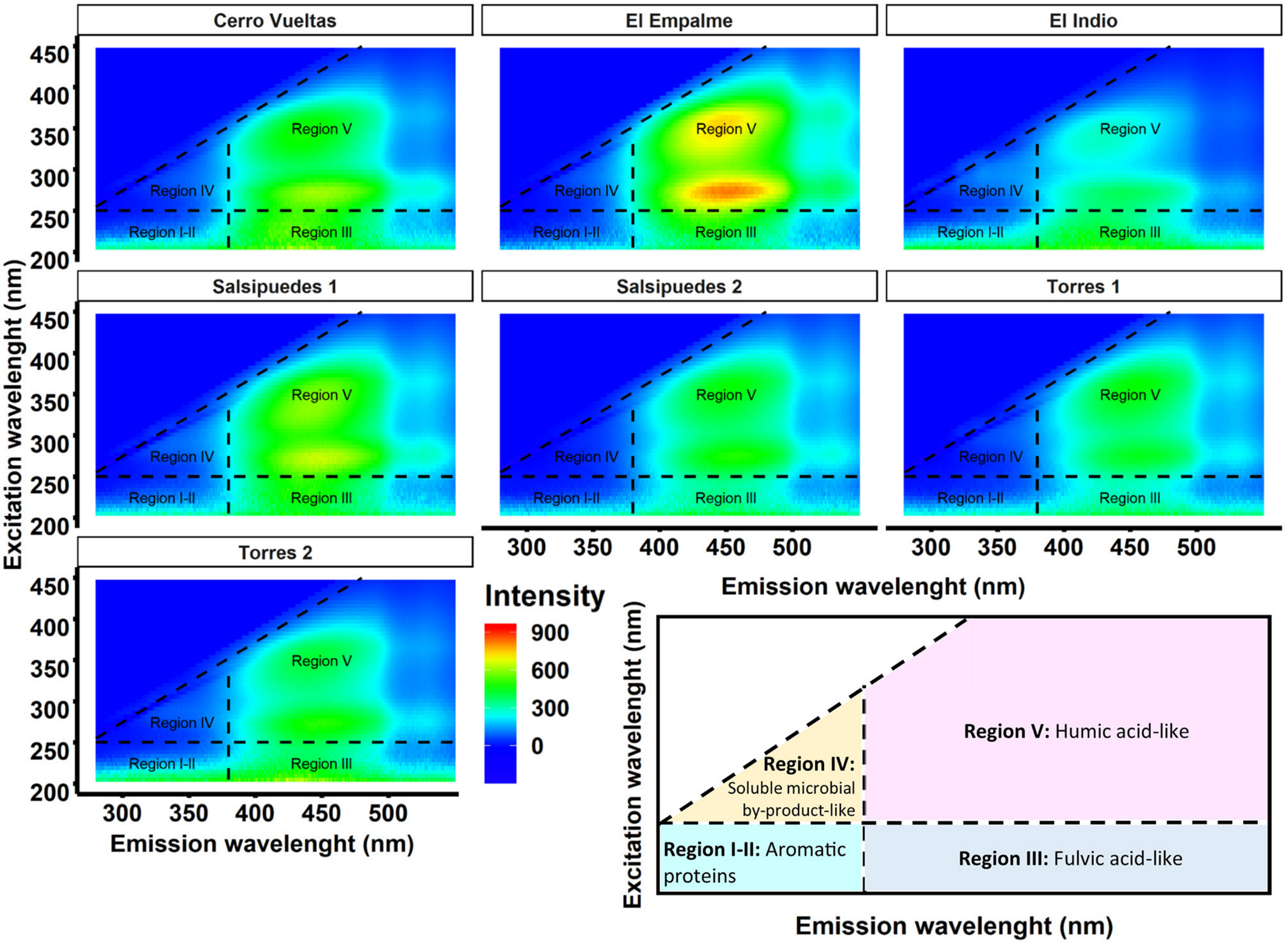


FIGURE 7 | Aggregated fluorescence excitation/emission matrices (EEMs) per sampling site. The inset shows five EEM regions defined by wavelength boundaries (Chen et al., 2003 and references therein) in absolute fluorescence intensity. El Empalme and El Indio represent the lowest and highest sampling sites, respectively (see **Figure 1**).

sites throughout a wet season. Second order indexes (BIX, and HIX) point to a greater relative contribution of recalcitrant soil-derived DOM with a high prevalence of humic and fulvic acids in the peat bogs. Evidence of fresh organic material from soluble microbial by-products and aromatic protein were detected in less degree, indicating also active microbial activity as a secondary decomposition pathway.

Our study comprises one exploratory sampling campaign during a complete hydrological year which limits the ability of exploring the impacts of inter-annual rainfall variability on DOC transport and export. However, from mid-May to the end of November, rainfall inputs are critical to sustain saturated soils, anoxic and acid conditions, which in turn exert a large control on carbon storage and humification processes. Our data revealed a clear seasonal pattern during the hydrological year. At the end of the dry season (April to mid-May), DOC concentrations were relatively high in almost all the sites. As the rainy season progressed, DOC concentrations decreased below 10 mg/L.

Since the Central America region has been experiencing recurrent droughts, it is expected that DOC concentrations may be significantly higher during post-drought events (enzymatic flush mechanism). Dry and oxic conditions reduces DOC decomposition due to the inhibitory effect of phenolic compounds on microbial activity (Fenner and Freeman, 2011). Our study also suggest that more robust sampling efforts using at higher temporal and spatial resolution using optical absorbance and fluorescence spectroscopic techniques (Dalmagro et al., 2017) are needed to underpin the potential impact of climate variability over DOM stocks, sources, and export, as well as microbial community changes between the dry and wet seasons (Stegen et al., 2016).

Further research on the relevance of photodegradation on temporal DOM dynamics of these highland peatbogs would also be of interest considering the high UV radiation to which the Páramo ecosystems are subjected to. Similarly, lowland drinking water operators should pay careful attention to DOM export (Osburn et al., 2018; Sánchez-Murillo et al., 2019), since high

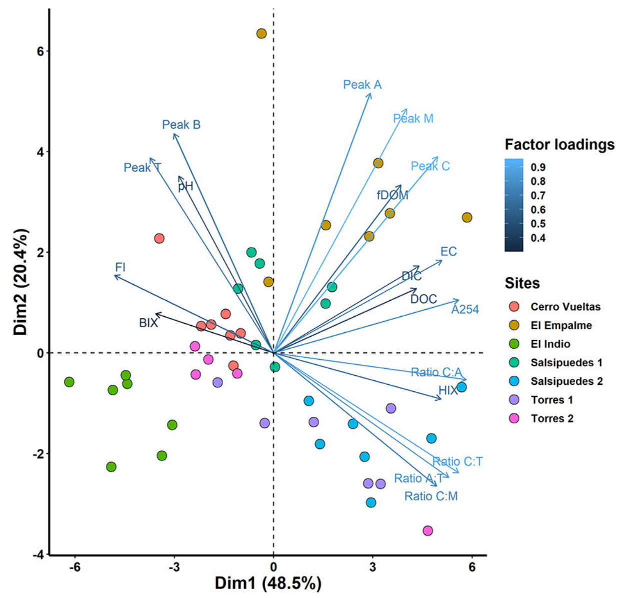


FIGURE 8 | Principal component analysis biplot explaining 68.9% of the total variance in two dimensions, including individual variable factor loadings (blue color scale) and distribution between the sampling sites (color-coded dots). El Empalme and El Indio represent the lowest and highest sampling sites, respectively (see **Figure 1**).

DOC concentrations particularly after prolonged droughts, may act as a substrate for microbial growth and may react during chlorination to form disinfection byproducts in the resultant drinking water (Ritson et al., 2017).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.4211/hs.d001620c2f104a0891dc18825ea2fe31, CUAHSI (Sánchez-Murillo et al., 2021).

AUTHOR CONTRIBUTIONS

RS-M and PG-A conceived the presented idea. RS-M, MP, and PG-A carried out all field sampling campaigns. RS-M, RS-G, GE-H, and RP-S conducted all chemical analyses. RS-M and RS-G collaborated with R statistical tests. RS-M initiated the first manuscript draft. All authors contributed to the final manuscript version.

FUNDING

This work was partially supported by International Atomic Energy Agency (IAEA) grants to RS-M: Isotope Techniques for the Evaluation of Water Sources for Domestic Water Supply in Urban Areas (CRP- F33024). Support from the Research Office of the Universidad Nacional of Costa Rica through grants SIA 0602-11, 0482-13, 0122-13, 0378-14, 0101-14, 0332-18, 0339-18, 0417-17, 0051-17, FECTE-2018, and 0487-20 was also fundamental.

ACKNOWLEDGMENTS

Analytical instrumental support from the IAEA Technical Cooperation Project (COS7005: Ensuring water security and sustainability of Costa Rica) is also acknowledged. Logistic assistance from the Wildlife and Health Laboratory (Universidad Estatal a Distancia, UNED, Costa Rica) is also acknowledged. Insights from three reviewers were incorporated in the final preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frwa.2021.742780/full#supplementary-material

REFERENCES

Amador, J. A. (1998). A climatic feature of the tropical Americas: the trade wind easterly jet. *Top. Meteor. Oceanogr*. 5, 91–102.

APHA and AWWA (2005). *Standard Methods for the Examination of Water and Wastewater, 20, Federation*. Washington, DC: American Public Health Association/American Water Works Association/Water Environment.

Arriagada, R. A., Sills, E. O., Ferraro, P. J., and Pattanayak, S. K. (2015). Do payments payoff? *Evidence from participation in Costa Rica's PES program. PLoS ONE* 10:e0136809. doi: 10.1371/journal.pone.0136809

Berbés-Blázquez, M., Bunch, M. J., Mulvihill, P. R., Peterson, G. D., and de Joode, B. V. W. (2017). Understanding how access shapes the transformation of ecosystem services to human well-being with an example from Costa Rica. *Ecosyst. Serv*. 28, 320–327. doi: 10.1016/j.ecoser.2017.09.010

Bispo, D. F. A., Silva, A. C., Christofaro, C., Silva, M. L. N., Barbosa, M. S., Silva, B. P. C., et al. (2016). Hydrology and carbon dynamics of tropical peatlands from Southeast Brazil. *Catena* 143, 18–25. doi: 10.1016/j.catena.2016.03.040

Blair, R. C., Higgins, J. J., Karniski, W., and Kromrey, J. D. (1994). A study of multivariate permutation tests which may replace Hotelling's T2 test in prescribed circumstances. *Multivar. Behav. Res*. 29, 141–163. doi: 10.1207/s15327906mbr2902_2

Bourgeau-Chavez, L. L., Endres, S. L., Graham, J. A., Hribljan, J. A., Chimner, R. A., Lillieskov, E. A., et al. (2018). "Mapping peatlands in boreal and tropical ecoregions," in *Comprehensive Remote Sensing*, ed S. Liang (Oxford: Elsevier), 24–44. doi: 10.1016/B978-0-12-409548-9.10544-5

Brak, B., Vroklage, M., Kappelle, M., and Cleef, A. M. (2005). "Comunidades vegetales de la turbera de altura "La Chonta" en Costa Rica," in *Páramos de Costa Rica*, eds M. Kappelle and S. P. Horn (Costa Rica: Editorial INBio), 607–630.

Bruijnzeel, L. A., Scatena, F. N., and Hamilton, L. S. (2010). *Tropical Montane Cloud Forests: Science for Conservation and Management*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511778384

Chaplot, V., and Mutema, M. (2021). Sources and main controls of dissolved organic and inorganic carbon in river basins: a worldwide meta-analysis. *J. Hydrol*. 603:126941. doi: 10.1016/j.jhydrol.2021.126941

Chaves, R. C., Figueredo, C. C., Boëchat, I. G., de Oliveira, J. T. M., and Gücker, B. (2020). Fluorescence indices of dissolved organic matter as early warning signals of fish farming impacts in a large tropical reservoir. *Ecol. Indic.* 115:106389. doi: 10.1016/j.ecolind.2020.106389

Chen, W., Westerhoff, P., Leenheer, J. A., and Booksh, K. (2003). Fluorescence excitation– emission matrix regional integration to quantify spectra for dissolved organic matter. *Environ. Sci. Technol.* 37, 5701–5710. doi: 10.1021/es034354c

Coble, P. G. (1996). Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Marine Chem.* 51, 325–346. doi: 10.1016/0304-4203(95)00062-3

Cook, S., Peacock, M., Evans, C. D., Page, S. E., Whelan, M. J., Gauci, V., et al. (2017). Quantifying tropical peatland dissolved organic carbon (DOC) using UV-visible spectroscopy. *Water Res.* 115, 229–235. doi: 10.1016/j.watres.2017.02.059

Cooper, D. J., Sueltenfuss, J., Oyague, E., Yager, K., Slayback, D., Cabero-Caballero, E. M., et al. (2019). Drivers of peatland water table dynamics in the central Andes, Bolivia and Peru. *Hydrol. Proc.* 33, 1913–1925, doi: 10.1002/hyp.13446

Cooper, D. J., Wolf, E. C., Colson, C., Vering, W., Granda, A., and Meyer, M. (2010). Alpine Peatlands of the Andes, Cajamarca, Peru. *Arctic Antarctic Alpine Res.* 42, 19–33, doi: 10.1657/1938-4246-42.1.19

Corrales Ulate, L. D. M. (2018). *Caracterización biofísica y determinación del intercambio de gases de efecto invernadero en las turberas de Talamanca, Costa Rica*. Available online at: http://201.207.189.89/handle/11554/8975

Couwenberg, J., Dommain, R., and Joosten, H. (2010). Greenhouse gas fluxes from tropical peatlands in Southeast Asia. *Glob. Change Biol.* 16, 1715–1732, doi: 10.1111/j.1365-2486.2009.02016.x

Craig, H. (1957). Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. *Geochim. Cosmochim. Acta* 12, 133–149. doi: 10.1016/0016-7037(57)90024-8

Dalmagro, H. J., Johnson, M. S., de Musis, C. R., Lathuillière, M. J., Graesser, J., Pinto-Júnior, O. B., et al. (2017). Spatial patterns of DOC concentration and DOM optical properties in a Brazilian tropical river-wetland system, J. *Geophys. Res. Biogeosci.* 122, 1883–1902. doi: 10.1002/2017JG003797

Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus* 16, 436–468. doi: 10.3402/tellusa.v16i4.8993

Drake, T. W., Van Oost, K., Barthel, M., Bauters, M., Hoyt, A. M., Podgorski, D. C., et al. (2019). Mobilization of aged and biolabile soil carbon by tropical deforestation. *Nature Geosci.* 12, 541–546. doi: 10.1038/s41561-019-0384-9

Esquivel-Hernández, G., Mosquera, G. M., Sánchez-Murillo, R., Quesada-Román, A., Birkel, C., Crespo, P., et al. (2019). Moisture transport and seasonal variations in the stable isotopic composition of rainfall in Central American and Andean Páramo during El Niño conditions (2015–2016). *Hydrol. Proc.* 33,1802–1817. doi: 10.1002/hyp.13438

Esquivel-Hernández, G., Sánchez-Murillo, R., Quesada-Román, A., Mosquera, G. M., Birkel, C., and Boll, J. (2018). Insight into the stable isotopic composition of glacial lakes in a tropical alpine ecosystem: Chirripó, Costa Rica. *Hydrol. Proc.* 32, 3588–3603. doi: 10.1002/hyp.13286

Esquivel-Hernández, G., Sánchez-Murillo, R., and Vargas-Salazar, E. (2021). Chirripó hydrological research site: advancing stable isotope hydrology in the Central American Páramo. *Hydrol. Proc.* 35:e14181. doi: 10.1002/hyp.14181

Fenner, N., and Freeman, C. (2011). Drought-induced carbon loss in peatlands. *Nature Geosci.* 4, 895–900. doi: 10.1038/ngeo1323

Friendly, M. (2002). Corrgrams: Exploratory displays for correlation matrices. *Am. Stat.* 56, 316–324. doi: 10.1198/000313002533

Gandois, L., Teisserenc, R., Cobb, A. R., Chieng, H. I., Lim, L. B. L., Kamariah, A. S., et al. (2014). Origin, composition, and transformation of dissolved organic matter in tropical peatlands. *Geochim. Cosmochim. Acta* 137, 35–47, doi: 10.1016/j.gca.2014.03.012

Gómez, L. D. (1986). *Vegetación y Clima de Costa Rica*. San José: Editorial de la Universidad Estatal a Distancia.

Good, P. (2000). *Permutation Test: A Practical Guide to Resampling Methods for Testing Hypotheses*, 2nd ed., eds P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, and S. Zeger. New York, NY: Springer. doi: 10.1007/978-1-4757-3235-1

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 δ^2H analysis of cumulative precipitation samples. *J. Hydrol.* 448, 195–200. doi: 10.1016/j.jhydrol.2012.04.041

Gumbricht, T., Roman-Cuesta, R. M., Verchot, L., Herold, M., Wittmann, F., Householder, E., et al. (2017). An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. *Glob. Change Biol.* 23, 3581–3599. doi: 10.1111/gcb.13689

Hansen, A. M., Kraus, T. E., Pellerin, B. A., Fleck, J. A., Downing, B. D., and Bergamaschi, B. A. (2016). Optical properties of dissolved organic matter (DOM): effects of biological and photolytic degradation. *Limnol. Oceanogr.* 61, 1015–1032. doi: 10.1002/lno.10270

Hapsari, K. A., Biagioni, S., Jennerjahn, T. C., Meyer Reimer, P., Saad, A., Achnopha, Y., et al. (2017). Environmental dynamics and carbon accumulation rate of a tropical peatland in Central Sumatra, Indonesia. *Quat. Sci. Rev.* 169, 173–187, doi: 10.1016/j.quascirev.2017.05.026

Helsel, D. (2012). *Statistics for Censored Environmental Data Using Minitab® and R*. New Jersey, NY: John Wiley and Sons. doi: 10.1002/9781118162729

Horgby, Å., Segatto, P. L., Bertuzzo, E., Lauerwald, R., Lehner, B., Ulseth, A. J., et al. (2019). Unexpected large evasion fluxes of carbon dioxide from turbulent streams draining the world's mountains. *Nat. Commun.* 10, 1–9. doi: 10.1038/s41467-019-12905-z

Hribljan, J. A., Suárez, E., Heckman, K. A., Lilleskov, E. A., and Chimner, R. A. (2016). Peatland carbon stocks and accumulation rates in the Ecuadorian paramo. *Wetlands Ecol. Manag.* 24, 113–127, doi: 10.1007/s11273-016-9482-2

Huguet, A., Vacher, L., Relexans, S., Saubusse, S., Froidefond, J. M., and Parlanti, E. (2009). Organic geochemistry properties of fluorescent dissolved organic matter in the Gironde Estuary. *Org. Geochem.* 40, 706–719. doi: 10.1016/j.orggeochem.2009.03.002

Itoh, M., Okimoto, Y., Hirano, T., and Kusin, K. (2017). Factors affecting oxidative peat decomposition due to land use in tropical peat swamp forests in Indonesia. *Sci. Total Environ.* 609, 906–915, doi: 10.1016/j.scitotenv.2017.07.132

Jiménez, J. A. (2016). "Bogs, marshes, and swamps of Costa Rica," in *Costa Rican Ecosystems*, ed M. Kappelle (Chicago: University of Chicago Press), 866–894.

Kiew, F., Hirata, R., Hirano, T., Xhuan Wong, G., Baran Aeries, E., Kemudang Musin, K., et al. (2018). CO_2 balance of a secondary tropical peat swamp forest in Sarawak, Malaysia. *Agr. Forest Meteorology* 248, 494–501. doi: 10.1016/j.agrformet.2017.10.022

Kimmel, K., and Mander, Ü. (2010). Ecosystem services of peatlands: implications for restoration. *Progr. Phys. Geogr.* 34, 491–514. doi: 10.1177/0309133310365595

Lachniet, M. S., and Seltzer, G. O. (2002). Late quaternary glaciation of Costa Rica. *Geol. Society Am. Bull.* 114, 547–558. doi: 10.1130/0016-7606(2002)114<0547:LQGOCR>2.0.CO;2

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: a package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01

Leng, L. Y., Haruna Ahmed, O., and Boyie Jalloh, M. (2019). Brief review on climate change and tropical peatlands. *Geosci. Front.* 10:18, doi: 10.1016/j.gsf.2017.12.018

Li, W., Dickinson, R. E., Fu, R., Niu, G.-Y., Yang, Z.-L., and Canadell, J. G. (2007). Future precipitation changes and their implications for tropical peatlands. Geophys. *Res. Lett.* 34:L01403. doi: 10.1029/2006GL028364

Lindsay, R. (2016). *Peatland (Mire Types): Based on Origin and Behavior of Water, Peat Genesis, Landscape Position, and Climate*. Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-6173-5_279-1

Lourençato, L. F., Caldeira, P. P., Bernardes, M. C., Buch, A. C., Teixeira, D. C., and Silva-Filho, E. V. (2017). Carbon accumulation rates recorded in the last 150 years in tropical high mountain peatlands of the Atlantic Rainforest, SE – Brazil. *Sci. Total Environ.* 579, 439–446, doi: 10.1016/j.scitotenv.2016.11.076

Markham, J., and Otárola, M. F. (2021). Bryophyte and lichen biomass and nitrogen fixation in a high elevation cloud forest in Cerro de La Muerte, Costa Rica. *Oecologia* 195, 489–497. doi: 10.1007/s00442-020-04840-4

Marx, A., Dusek, J., Jankovec, J., Sanda, M., Vogel, T., van Geldern, R., et al. (2017). A review of CO_2 and associated carbon dynamics in headwater streams: a global perspective. *Rev. Geophysics* 55, 560–585. doi: 10.1002/2016RG000547

Mcknight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T., and Andersen, D. T. (2001). Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* 46, 38–48. doi: 10.4319/lo.2001.46.1.0038

Moore, S., Evans, C. D., Page, S. E., Garnett, M. H., Jones, T. G., Freeman, C., et al. (2013). Deep instability of deforested tropical peatlands revealed by fluvial organic carbon fluxes. *Nature* 493, 660–664, doi: 10.1038/nature11818

Müller, D., Warneke, T., Rixen, T., Müller, M., Jamahari, S., Denis, N., et al. (2015). Lateral carbon fluxes and CO_2 outgassing from a tropical peat-draining river. *Biogeosciences* 12, 5967–5979. doi: 10.5194/bg-12-5967-2015

Osburn, C. L., Oviedo-Vargas, D., Barnett, E., Dierick, D., Oberbauer, S. F., and Genereux, D. P. (2018). Regional groundwater and storms are hydrologic controls on the quality and export of dissolved organic matter in two tropical rainforest streams, Costa Rica. *J. Geophys. Res.* 123, 850–866. doi: 10.1002/2017JG003960

Page, S. E., and Baird, A. J. (2016). Peatlands and global change: response and resilience. *Ann. Rev. Environ. Resourc.* 41, 35–57. doi: 10.1146/annurev-environ-110615-085520

Page, S. E., Siegert, F., Rieley, J. O., Boehm, H. D., Jaya, A., and Limin, S. (2002). The amount of carbon released from peat and forest fires in Indonesia during 1997. *Nature* 420, 61–65, doi: 10.1038/nature01131

Park, M., and Snyder, S. (2018). Sample handling and data processing for fluorescent excitation-emission matrix (EEM) of dissolved organic matter (DOM). *Chemosphere* 193, 530–537. doi: 10.1016/j.chemosphere.2017.11.069

Pawellek, F., and Veizer, J. (1994). Carbon cycle in the upper Danube and its tributaries: $\delta^{13}C_{DIC}$ constrains. *Israel J. Earth Sci.* 43, 187–194.

Pesántez, J., Mosquera, G. M., Crespo, P., Breuer, L., and Windhorst, D. (2018). Effect of land cover and hydro-meteorological controls on soil water DOC concentrations in a high-elevation tropical environment. *Hydrol. Proc.* 32, 2624–2635. doi: 10.1002/hyp.13224

Poulin, B. A., Ryan, J. N., and Aiken, G. R. (2014). Effects of iron on optical properties of dissolved organic matter. *Environ. Sci. Technol.* 48, 10098–10106. doi: 10.1021/es502670r

Pucher, M., Wünsch, U., Weigelhofer, G., Murphy, K., Hein, T., and Graeber, D. (2019). staRdom : versatile software for analyzing spectroscopic data of dissolved organic matter in R. *Water* 11, 1–19. doi: 10.3390/w11112366

R., Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/ (accessed June 14, 2021).

Ribeiro, K., Pacheco, F. S., Ferreira, J. W., de Sousa-Neto, E. R., Hastie, A., Krieger Filho, G. C., et al. (2021). Tropical peatlands and their contribution to the global carbon cycle and climate change. *Glob. Change Biol.* 27, 489–505. doi: 10.1111/gcb.15408

Ritson, J. P., Brazier, R. E., Graham, N. J. D., Freeman, C., Templeton, M. R., and Clark, J. M. (2017). The effect of drought on dissolved organic carbon (DOC) release from peatland soil and vegetation sources. *Biogeosciences* 14, 2891–2902. doi: 10.5194/bg-14-2891-2017

Sánchez-Murillo, R., Esquivel-Hernández, G., Birkel, C., Correa, A., Welsh, K., Durán-Quesada, A. M., et al. (2020). Tracing water sources and fluxes in a dynamic tropical environment: from observations to modeling. *Front. Earth Sci* 8:571477. doi: 10.3389/feart.2020.571477

Sánchez-Murillo, R., Gastezzi, P., Sánchez-Gutiérrez, R., Esquivel-Hernández, G., and Pérez-Salazar, R. (2021). Optical properties of dissolved carbon in a high elevation tropical peatland ecosystem, *HydroShare*. doi: 10.4211/hs.d001620c2f104a0891dc18825ea2fe31

Sánchez-Murillo, R., Romero-Esquivel, L. G., Jiménez-Antillón, J., Salas-Navarro, J., Corrales-Salazar, L., Álvarez-Carvajal, J., et al. (2019). DOC transport and export in a dynamic tropical catchment. *J. Geophys. Res.* 124, 1665–1679. doi: 10.15359/cicen.1.35

Silva, A. C., Barbosa, M. S., Barral, U. M., Silva, B. P. C., Fernandes, J. S. C., Viana, A. J. S., et al. (2019). Organic matter composition and paleoclimatic changes in tropical mountain peatlands currently under grasslands and forest clusters. *Catena* 180, 69–82. doi: 10.1016/j.catena.2019.04.017

Spearman, C. E. (1904). The proof and measurement of association between two things. *Am. J. Psychol.* 15, 72–101. doi: 10.2307/1412159

Stan, K., and Sanchez-Azofeifa, A. (2019). Deforestation and secondary growth in Costa Rica along the path of development. *Reg. Environ. Change* 19, 587–597. doi: 10.1007/s10113-018-1432-5

Stegen, J. C., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., Arntzen, E. V., et al. (2016). Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nat. Commun.* 7:11237. doi: 10.1038/ncomms11237

Villa, J. A., Mejía, G. M., Velásquez, D., Botero, A., Acosta, S. A., Marulanda, J. M., et al. (2019). Carbon sequestration and methane emissions along a microtopographic gradient in a tropical Andean peatland. *Sci. Total Environ.* 654, 651–661, doi: 10.1016/j.scitotenv.2018.11.109

Waldron, S., Vihermaa, L., Evers, S. , Garnett, M. H., Newton, J., and Henderson, A. C. G. (2019). C mobilisation in disturbed tropical peat swamps: old DOC can fuel the fluvial efflux of old carbon dioxide, but site recovery can occur. *Sci. Rep.* 9:11429. doi: 10.1038/s41598-019-46534-9

Wallbott, L., G., and Siciliano, and Lederer, M. (2019). Beyond PES and REDD+: Costa Rica on the way to climate-smart landscape management? *Ecol. Soc.* 24:24. doi: 10.5751/ES-10476-240124

Wei, T., and Simko, V. (2021). *R package 'corrplot': Visualization of a Correlation Matrix. (Version 0.90)*. Available online at: https://github.com/taiyun/corrplot (accessed June 14, 2021).

Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., and Mopper, K. (2003). Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* 37, 4702–4708. doi: 10.1021/es030360x

Williams, M. W., Hood, E. W., Ostberg, G., Francou, B., and Galarraga, R. (2001). Synoptic survey of surface water isotopes and nutrient concentrations, Páramo high-elevation region, Antisana Ecological Reserve, Ecuador. *Arctic Antarctic Alpine Res.* 33, 397–403. doi: 10.1080/15230430.2001.12003447

Yu, Z. (2011). Holocene carbon flux histories of the world's peatlands: Global carbon-cycle implications. *Holocene* 21, 761–774, doi: 10.1177/0959683610386982

Zsolnay, A., Baigar, E., and Jimenez, M. (1998). Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. *Chemosphere* 38, 45–50. doi: 10.1016/S0045-6535(98)00166-0

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.

frontiers
in Water

PERSPECTIVE
published: 26 January 2022
doi: 10.3389/frwa.2021.786936

Rivers as Plastic Reservoirs

Tim van Emmerik, Yvette Mellink, Rahel Hauk, Kryss Waldschläger and Louise Schreyers*

Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, Netherlands

Land-based plastic waste, carried to the sea through rivers, is considered a main source of marine plastic pollution. However, most plastics that leak into the environment never make it into the ocean. Only a small fraction of plastics that are found in the terrestrial and aquatic compartments of river systems are emitted, and the vast majority can be retained for years, decades, and potentially centuries. In this perspective paper we introduce the concept of river systems as plastic reservoirs. Under normal conditions, hydrometeorological variables (such as wind, runoff and river discharge) mobilize, transport and deposit plastics within different river compartments (e.g., riverbanks, floodplains, lakes, estuaries). The emptying of these plastic reservoirs primarily occurs under extreme hydrological conditions (e.g., storms, floods). In this paper we specifically focus on the retention mechanisms within different river compartments, and their effect on the fate of the plastics that are accumulated on various timescales. We aim to introduce the concept of rivers as (long-term) sinks for plastic pollution, and provide suggestions for future research directions.

Keywords: macroplastic, hydrology, marine debris, plastic soup, pollution, water quality

Edited by:
Boris M. Van Breukelen,
Delft University of
Technology, Netherlands

Reviewed by:
Luca Solari,
University of Florence, Italy
Ellen Wohl,
Colorado State University,
United States
Mauricio Arias,
University of South Florida,
United States

****Correspondence:***
Tim van Emmerik
tim.vanemmerik@wur.nl

Specialty section:
This article was submitted to
Environmental Water Quality,
a section of the journal
Frontiers in Water

Received: *30 September 2021*
Accepted: *22 December 2021*
Published: *26 January 2022*

Citation:
van Emmerik T, Mellink Y, Hauk R,
Waldschläger K and Schreyers L
(2022) Rivers as Plastic Reservoirs.
Front. Water 3:786936.
doi: 10.3389/frwa.2021.786936

INTRODUCTION

In contrast to what is commonly assumed, most land-based plastic waste does not flow into the ocean. The available observations of plastic pollution within and across river catchments in fact support the hypothesis that most plastics may be retained in and around rivers (e.g., van Emmerik et al., 2019b; Tramoy et al., 2020a; Newbould et al., 2021; Ryan and Perold, 2021). In this paper we present our perspective on what happens to the plastics that do not flow into the sea, and what future work is needed to better understand and quantify the role of rivers as plastic reservoirs. Note that in this paper, we mainly focus on macroplastics (>0.5 cm), although we do include relevant science on other size ranges.

Plastic pollution is an emerging environmental risk due to its negative impact on ecosystem health and human livelihood. Macroplastics in particular can directly (lethally) injure animal and vegetation species, damage vessels and hydraulic infrastructures, and lead to increased urban flood risk due to clogging (van Emmerik and Schwarz, 2020). Plastics enter the environment through various pathways (**Figure 1A**), including improper disposal, direct dumping, leaking from waste infrastructure or industry, sewage discharge, hydrometeorological variables such as wind and surface runoff, and disasters including floods, storms or landslides (Lechthaler et al., 2020; van Emmerik and Schwarz, 2020). As such, river systems play a major role in transporting land-based plastic waste into the world's oceans.

Plastic transport processes in river systems are complex. Various hydrometeorological variables have been correlated to the mobilization, transport, dispersion, and accumulation of plastics in rivers. Examples include wind (Bruge et al., 2018; Roebroek et al., 2021a), river flow velocity and discharge (Haberstroh et al., 2021), and precipitation-induced runoff (Treilles et al., 2021). Hydrology is not only important for mobilization and transport processes, but also determines when and where plastics accumulate in river systems (Liro et al., 2020). The combination of river dynamics and characteristics are key to whether plastics are deposited on banks and floodplains, get entrapped in riparian vegetation, settle into the sediment, or get transported to downstream areas (van Emmerik and Schwarz, 2020). Once plastics enter the estuary, the combination of riverine and tidal dynamics determine the fate of (specific) plastics (**Figure 1A**). A growing amount of evidence suggests that most plastics that enter the environment are never exported into the ocean (van Emmerik et al., 2019b, 2020a; Tramoy et al., 2020a,b; Ryan and Perold, 2021). Globally, only a very small fraction of the mismanaged plastic waste is emitted into the sea (Meijer et al., 2021). The most important science and policy questions therefore remain where the plastic accumulates within river systems, and at what spatial and temporal scales they are retained.

In this perspective paper we elaborate on the hypothesis that rivers act as reservoirs for plastic pollution. We present an overview of the various river compartments that act as accumulation zones for plastic pollution, as well as the retention and (re)mobilization mechanisms. We also elaborate on the fate of plastics retained in these zones, including long-term storage, degradation and remobilization processes. Finally, we present a perspective for future research directions, including specific scientific questions that need to be addressed to effectively prevent, mitigate and reduce plastic pollution in the environment.

MOST RIVER PLASTICS DO NOT FLOW INTO THE OCEAN

On global (Jambeck et al., 2015; Borrelle et al., 2020; Meijer et al., 2021), regional (Bai et al., 2018; van Calcar and van Emmerik, 2019; González-Fernández et al., 2021), and local levels (van Emmerik et al., 2019a; Schöneich-Argent et al., 2020; Vriend et al., 2020), several assessment strategies have been developed to estimate the plastic mass flows from land into rivers, and finally into the ocean. In recent years, the focus on estimating plastic emissions has shifted more toward understanding what happens with plastic within river systems before it makes it into the ocean. Meijer et al. (2021) quantified that globally, less than 2% of the total plastic waste leaked into the environment ends up in the marine environment. Depending on basin characteristics and the locations of plastic leakage, this ratio varies considerably between rivers. For relatively short rivers with high population densities close to the river and the coast, such as found in the Philippines and Indonesia, plastics are more likely to be emitted into the ocean. For long rivers with larger upstream population, infrastructure such as dams, and larger floodplains (for example the Rhône or the Rhine) the likelihood of emission is even less. On a local level, annual emissions of rivers in Jakarta, Indonesia, were compared to the total amount of plastic leaked into the environment, showing that only 3% of the mismanaged plastic waste is transported into the ocean (van Emmerik et al., 2019b). Substantial work on the Seine, France, demonstrated that plastics can have long retention times, already surpassing decades (Tramoy et al., 2020a). Other studies demonstrated the short travel distances of plastic waste once they enter a river system (Weideman et al., 2020; Newbould et al., 2021), suggesting they accumulate somewhere along the river. All available evidence considered, river systems can be a major plastic reservoir, rather than only a source of marine plastic pollution.

WHERE ARE RIVER PLASTICS RETAINED?

When rivers do function as reservoirs for plastic pollution, it is key to better understand and quantify where plastics are retained, what factors determine retention, and for how long they are retained. In this section, we provide examples from observational studies that found plastics entrapped in various compartments of river systems, which we categorized into six main retention mechanisms (**Figures 1B–G**).

Retained on Riverbanks and Floodplains

Plastic pollution found on riverbanks and floodplains originates either from terrestrial pathways (direct littering or dumping) or from riverine transport. Similar to some sediments, macroplastics are more easily deposited on the riverbanks along areas with stagnant water, low flow velocities, low channel slopes and high riparian vegetation densities (Bruge et al., 2018). During floods, riverbank plastic can be transported and deposited in floodplains as well, where it will most likely be buried by natural sediments (Lechthaler et al., 2020; Roebroek et al., 2021b). However, the retention mechanisms and time scales of plastic on riverbanks and in floodplains remain largely unresolved, especially as sedimentation rates and river characteristics play an important, yet unknown role (Lechthaler et al., 2021). Roebroek et al. (2021a) found retention times of shorter than six months, whereas Tramoy et al. (2020b), who investigated the use-by dates on medical packaging, showed that 37% of the litter was discarded before 1984 (Tramoy et al., 2020b). Both the spatiotemporal distribution and the retention times of macroplastics on riverbanks and in floodplains must be further explored in the future, for which we recommend upscaling and expanding methods. The latter may include using expiration dates or the language on packaging, sampling plastics at deeper layers in the riparian zone, and quantifying retention within specific features such as vegetation.

Retained in Plants

Vegetation along the riverbanks, in estuaries and floating at the river surface can function as both a carrier and a reservoir of macroplastics. Variations in river water level can result in the deposition of macroplastics in grass, bushes and trees (Williams

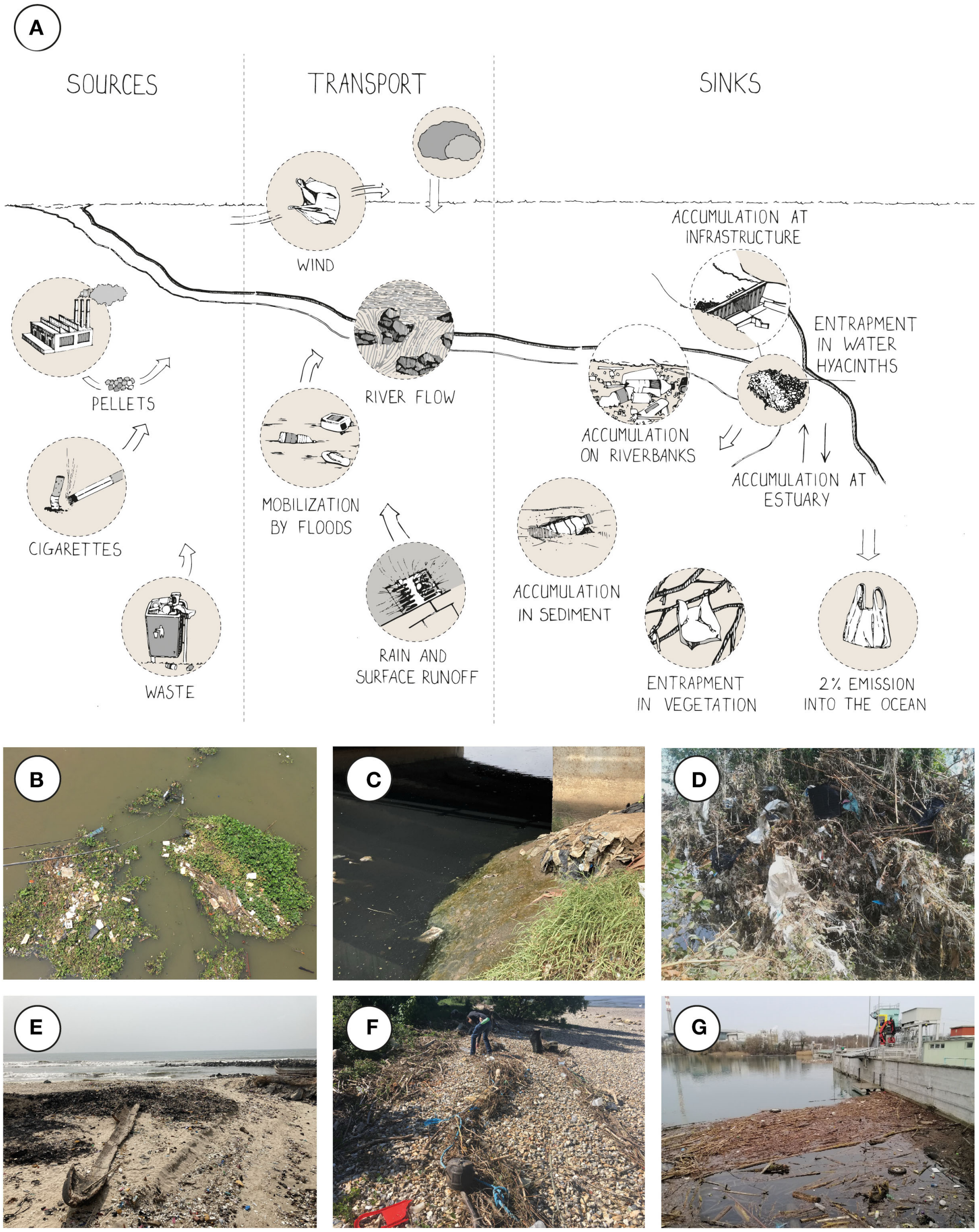


FIGURE 1 | (A) Conceptual overview of sources, transport mechanisms and sinks (retention zones) of river plastic pollution (Graphic design by Cher van den Eng). Sources include mismanaged plastic waste, littering from individuals (e.g., cigarette butts), and leakage from industry (e.g., pellets). Transport processes include wind, surface runoff, river flow dynamics, and floods. Retention zones include infrastructure, stationary and mobile vegetation (the latter illustrated by the water hyacinth), riverbanks, and the estuary. Finally, only <2% is estimated to be emitted into the ocean globally. The photos show examples of plastic retention **(B)** in water hyacinth (Saigon river, Vietnam), **(C)** within riverbank sediments, (Odaw river, Ghana), **(D)** in riparian vegetation (Meuse river, the Netherlands), **(E)** in the estuary mouth (Odaw river, Ghana), **(F)** on riverbanks (Seine river, France), and **(G)** at infrastructure (Rhône river, Switzerland). All photos are taken by the authors.

and Simmons, 1996). The retention time of plastics trapped in riparian vegetation greatly depends on (seasonally varying) flow dynamics, and the characteristics of the vegetation, such as the height, plant density and spatial configuration, architecture, and geometry. These factors determine to what extent the water interacts with the vegetation, and in turn influences the likelihood of entrapment. A recent study observed that arboreal and shrubby vegetation entrap more macroplastic compared to herbaceous, reed and bush types (Cesarini and Scalici, 2022). The presence of branches, thicker foliage, as well as higher height and density compared to the other vegetation types might play an important role in macroplastic trapping and retention. Moreover, floating aquatic vegetation has been found to trap and carry considerable amounts of macroplastics downstream. In the Saigon river, drifting patches of water hyacinths were found to carry as much as 78% of the floating plastic (Schreyers et al., 2021). However, the effects of bidirectional flows due to the influence of tidal regimes, and the stranding of large hyacinth patches on the riverbanks further complicates this conception of floating vegetation as a transport carrier of plastic into the sea.

Retained in Riverbed Sediments

Dense macroplastic items (>1,000 kg/m^3) can sink to the river bed in the absence of strong turbulence or currents. These negatively buoyant plastic types include items made of polyvinyl chloride, polyethylene terephthalate, polypropylene and polystyrene. During the transport process, initially buoyant macroplastic items, such as bottles, can fill with water or have biofilm grown on their surface, and thus lose their buoyancy and sink (Gabbott et al., 2020; Lechthaler et al., 2020; Al-Zawaidah et al., 2021). These items can be deposited in areas of the river with low flow velocities (Liro et al., 2020) and high natural sedimentation rates. The macroplastics can either be buried by natural sediments, which increases their remobilization threshold, be transported as bed-load transport, or get resuspended into the water column due to stronger flow velocities or turbulence (van Emmerik and Schwarz, 2020). Possible methods to measure deposited macroplastics include the analysis of dredged sediment (Constant et al., 2021), or analyzing samples from sediment traps (Enders et al., 2019; Saarni et al., 2021). There are only a few measurement methods to quantify macroplastics in riverbed sediments or as bed-load, yet, those are essential to identify macroplastic retention zones and understand their spatiotemporal distribution.

Retained at Infrastructure

Only a small fraction of the world's rivers remain free flowing (Grill et al., 2019). Many river systems have been anthropogenically altered to guarantee shipping, water supply for consumption and irrigation, or generate power. Examples include dams, water inlets, groins, ports, sluices and canals. This type of infrastructure can act as accumulation zones for macroplastics. Larger plastic concentrations have been found upstream of dams, both in surface waters and sediment (Zhang et al., 2015; Castro-Jiménez et al., 2019; Watkins et al., 2019). Macroplastics become trapped around hydroelectric power plants, and are deposited on riverbanks and in sediments around groins (Skalska et al., 2020; Tramoy et al., 2020a). In urban water systems, plastics are trapped by rack structures, pumping stations or around bridges (Tasseron et al., 2020). Consequently, the water system becomes (partially) obstructed, which leads to higher upstream water levels (Honingh et al., 2020). Targeted infrastructure, such as floating litter booms or litter traps, is placed to intentionally accumulate plastics. Besides capturing plastic for removal, these traps also provide a rich source of data. Several assessments to estimate plastic mass transport, identify plastic mass flows, or characterize plastic waste, analyzed waste captured at such infrastructure in the Netherlands (Vriend et al., 2020), France (Gasperi et al., 2014) and Malaysia (Malik et al., 2020). Further work at the river system scale is needed to shed additional light on the role of infrastructure on plastic accumulation, and the plastic mass balance on catchment level.

Retained in Lakes

Exorheic lakes, i.e., lakes that ultimately drain into the ocean, are found in river systems all over the world. Macroplastics have been found afloat in lake surface waters (Faure et al., 2015), buried in bottom lake sediments (Egessa et al., 2020), and deposited on lake shorelines/beaches (Faure et al., 2015; Egessa et al., 2020). These plastics either come from local activities (e.g., littering, fishing gear, direct wastewater drainage from a nearby urban area or direct surface runoff) or have been conveyed by rivers that discharge into the lake. Once a river flows into a lake, the plastics are carried by the lake surface currents and can concentrate in small temporary gyres (Faure et al., 2015). Wind-induced surface currents, especially during storms, also transport and deposit substantial amounts of plastic on the shorelines of lakes (Zbyszewski et al., 2014). The full retention capacity of lakes remains unresolved, and as many river systems worldwide feature lakes, quantifying how much upstream riverine plastics make it through a lake to the downstream river, is crucial for accurate estimates on riverine plastic transport and emissions.

Retained in Estuaries

Estuaries are influenced by tidal dynamics and freshwater discharge, leading to diurnally changing water level, salt concentrations, flow velocity, and flow direction (Savenije, 2012). The scarce observational evidence available suggests that plastic transport and accumulation is impacted by these dynamics at different timescales. Within tidal cycles, plastic flux close to the river mouth has been found to be about the same during both ebb and flood tide, suggesting that the actual net transport from rivers into the sea is very limited (van Emmerik et al., 2020c). As long as the net discharge is low, the plastics accumulating in the estuary have a growing likelihood of beaching on the riverbanks (van Emmerik et al., 2020a), getting entrapped in riparian vegetation (Martin et al., 2020), deposited in the sediment (Acha et al., 2003), or to degrade and fragment into smaller particles (Lebreton et al., 2019). Increased freshwater discharge, especially during flood conditions, may cause a peak of plastic export into the near coastal zone. Other work has suggested that plastics may be retained on estuary shores for years or even longer (Tramoy et al., 2020b). The role of estuaries on plastic transfer between rivers and the ocean, and the time scales of retention, remain largely unresolved. Future work should therefore explore the factors that drive plastic accumulation, (re)mobilization, and the time scales of retention in estuaries.

WHAT HAPPENS TO ACCUMULATED PLASTICS?

In the previous sections we discussed the various plastic transport mechanisms, followed by examples of where and why plastics are entrapped, demonstrating the function of rivers as plastic reservoirs. In this section we further discuss what happens to the retained plastics, and how the reservoirs may be emptied.

Breaking Down Into Smaller Particles

Plastics in river systems are commonly fragments of soft and hard plastics or foam (Castro-Jiménez et al., 2019;

Tramoy et al., 2019; van Emmerik et al., 2020a). While some plastic items remain intact in the environment for decades (Tramoy et al., 2020b), others break down into smaller and smaller particles, due to physical fragmentation or chemical degradation (Delorme et al., 2021). Physical fragmentation is caused by mechanical influences such as waves, and abrasion through sediment (Barnes et al., 2009) but also chewing and digestive fragmentation through organisms (Dawson et al., 2018; Mateos-Cárdenas et al., 2020). Chemical degradation processes include biodegradation, photodegradation, thermo-oxidative degradation, thermal degradation, and hydrolysis. Exposure to UV-light and air accelerate the degradation process (Andrady, 2011; Waldschläger et al., 2020). During fragmentation and degradation different processes interact, depending on environmental conditions and properties of the plastic items (Andrady, 2011; Song et al., 2017).

Once broken down into microplastic particles, they could either be deposited in the sediment, be transported in the water column, or ingested by biota (Leslie et al., 2017). The interaction of fragmentation and degradation, and transport of plastics through different river compartments remains largely unknown, and a quantification method for degradation rates has not been found yet (Ford et al., 2022). We do expect smaller particles to be remobilized more easily, travel further downstream, and eventually reach the ocean. Microplastics have been found to be flushed away after floods or storm events (Hurley et al., 2018; Treilles et al., 2022), and do not experience the same interaction with many of the entrapment factors as mentioned in section Where are River Plastics Retained? The plastic reservoir within rivers may therefore be emptied on the form of fragmented secondary microplastic particles.

Retention Leads to Longer Exposure Times

We have seen that plastics accumulate within various compartments in river systems, leading to the exposure of biota to plastics, with all its potential detrimental consequences. The longer plastic accumulates in natural ecosystems, the higher the number of species that will encounter plastics and therefore the higher the chance it leaves a negative impact. Recent evidence shows that macroplastics can be retained in rivers for several decades (Tramoy et al., 2020a), increasing the likelihood of ingestion by biota. In the Amazon, as much as 80% of the fish species had ingested plastic particles (Andrade et al., 2019), which can be life-threatening, especially when (toxic) waterborne contaminants are adsorbed to the plastics (Bellasi et al., 2020). Given the pivotal role of freshwater biota in the food chain, future research must continue to explore the degree and duration of the exposure of biota to plastics in order to clarify the threats they face in different river compartments.

Emptying of the Plastic Reservoir by Extreme Events

During high wind speeds, high rainfall intensities, or high river discharges, accumulated plastics are likely to be remobilized and removed from retention zones. Ample studies have demonstrated an increased riverine plastic transport during flood events (Moore et al., 2011; Jang et al., 2014; Mihai, 2018; van Emmerik et al., 2019c; Roebroek et al., 2021b). Flood events inundate the riverbanks, floodplains and tidal flats and thereby (re)mobilize the plastics that have been accumulating in these reservoirs during normal conditions. Additional plastics also enter the river system during such events. Estuaries that during normal conditions tend to have a zero net transport of plastics into the oceans, export plastics during and just after a flood event

TABLE 1 | Overview of the retention and release dynamics, and their timescales, discussed in this paper.

	Retention dynamics		Release dynamics		Key references
Compartments	**Timescale**	**Main mechanisms**	**Timescale**	**Main mechanisms**	
Riverbanks	Continuous accumulation	Stagnant water, low flow velocities, low channel slopes	Annual to decadal	Floods	Tramoy et al. (2020b)
Floodplains	Annual to decadal	Extreme events	Annual to decadal	Floods	Lechthaler et al. (2021) and Liro et al. (2020)
Floating vegetation	Continuous accumulation	Growing cycle of plants	Seasonal cycle	Disintegration of aquatic plants (flushing effect)	Schreyers et al. (2021)
Riparian vegetation	Continuous accumulation	Height, density, configuration, distance to river channel	Annual to decadal	Floods	Cesarini and Scalici (2022)
Riverbed sediments	Continuous accumulation	Buoyancy of items, low flow velocities and turbulences	Monthly to annual to decadal	Hydrologic regime with high flow velocities and turbulences	Hurley et al. (2018)
Infrastructure	Continuous accumulation	Entrapment upstream of infrastructure	Annual to decadal	Discharging retained water	Castro-Jiménez et al. (2019), Watkins et al. (2019), and Zhang et al. (2015)
Lakes	Continuous accumulation	Reduced flow velocity than in the inflowing river	Decadal to centennial	Floods and increased water levels	Egessa et al. (2020) and Faure et al. (2015)
Estuaries	Continuous accumulation	Linked to the tidal cycle	Annual to decadal	Increased freshwater discharge and extreme-events	Tramoy et al. (2020b)

Note that the timescales indicated are based on the observational evidence available to date, but remain uncertain due to limited research globally.

(Liro et al., 2020). Even plastics buried within riverbed sediments can get removed from this reservoir when the active sediment layer of the riverbed is destabilized by extremely high water flow velocities (Ockelford et al., 2020). However, extreme events do not solely (re)mobilize plastics, they can lead to plastic deposition as well. When the water flow velocities in the river channel are much higher than above the submerged floodplains, plastics can become transferred from the faster flow (channel) to the slower flow velocity area (floodplain) and settle there (Ciszewski and Grygar, 2016). Another phenomenon, often referred to as the "Christmas tree effect," whereby plastic debris carried by the flood becomes entrapped in riparian trees and is left behind when the water level drops again, demonstrates the reservoir filling capacity of a flood event (Williams and Simmons, 1996). Future research must explore the relative importance between (re)mobilization and (re)deposition mechanisms during extreme events.

DISCUSSION AND OUTLOOK

In this paper we demonstrate the function of rivers as plastic reservoirs, through the retention and release mechanisms in various river compartments (**Table 1**). One of the crucial open questions that remains are the spatial and temporal scales of the accumulation and retention mechanisms across different river compartments. In estuaries, for example, the retention of floating plastics may be in the order of days (due to the tidal cycle), whereas riverbanks are shown to store plastics for several decades already. These findings suggest that even if no more plastics would leak into the environment as of today, the plastics retained within river systems will continue to have negative impacts, break down in smaller particles, and function as a potential source of marine plastic pollution for the decades and centuries to come. We therefore call on more research on investigating the factors that determine the processes, spatial extent and time scales of retention and release mechanisms of plastics on riverbanks, lakes, estuaries, vegetation, sediment, and infrastructure.

Detailed investigations of fundamental retention and release mechanisms are complicated due to a lack of observations across all river compartments, and adequate monitoring methods (Broere et al., 2021; van Emmerik et al., 2022). On top of that, most work focuses on tracking specific items (e.g., bottles) through both modeling and observations (Duncan et al., 2020; Newbould et al., 2021), or quantifying plastics in single river cross-sections. Even though such targeted and local studies provide useful snapshots, we advocate for large-scale and holistic studies in which the transport and retention in all river compartments is examined. Furthermore, we reiterate the need for conceptual and observation based approaches to accurately quantify the amount and characteristics of plastics within each compartment, and the mass flow between them (van Emmerik et al., 2022). Assessments that focus on the entire river system can be facilitated through for example large-scale tracker experiments, exploring the use of remote sensing, and promotion of citizen science apps.

Tracker experiments, both using GPS trackers and marked items, have yielded new insights in the transport and accumulation processes, demonstrating that the travel distance is limited in many cases. River focused experiments have used relatively low number of trackers (less than a hundred), and large-scale experiments using thousands of trackers have to date only been used at sea (e.g., Duncan et al., 2020; Schöneich-Argent and Freund, 2020; Tramoy et al., 2020b; Newbould et al., 2021; Ryan and Perold, 2021). Close-range and satellite remote sensing also offer new avenues for large-scale data collection. Marine plastic debris can already be observed from space using Sentinel-2 imagery (Biermann et al., 2020). Rivers are generally smaller and more dynamic systems, requiring more research to fully explore the potential for river plastic monitoring from space. Lab experiments to establish hyperspectral reflectance databases will accelerate the development of new sensors, and disentangle and downscale the aggregated signals observed from space (Tasseron et al., 2021). Finally, several citizen science mobile applications facilitate rapid upscaling of data collection of plastic pollution on land (Litterati; Ballatore et al., 2021) and in river systems (CrowdWater; van Emmerik et al., 2020b). Projects that target specific locations have been shown to stimulate the use of such apps and rapidly increase observations done by citizen scientists (Tasseron et al., 2020).

With this paper we aim to demonstrate the current state of the science on river plastic pollution, and elaborate on the hypothesis that rivers function as plastic reservoirs. River plastic pollution and hydrology are unequivocally connected, emphasizing the need for more transdisciplinary studies to improve the understanding of this emergent pollutant.

AUTHOR CONTRIBUTIONS

TvE, YM, RH, KW, and LS: conceptualization, formal analysis, investigation, writing—original draft, and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

The work of TvE was supported by the Veni Research Program, the River Plastic Monitoring Project with project number 18211, which was (partly) financed by the Dutch Research Council (NWO). The work of LS was supported by NWO Open Mind grant 18127. The work of KW was supported by Investment Plan for strengthening the Technical Sciences at Wageningen University.

ACKNOWLEDGMENTS

We thank Cher van den Eng for providing the artwork for **Figure 1**. We thank Paul Vriend and Lourens Meijer for providing valuable input on an earlier version of this manuscript. We are grateful to the three reviewers and the editor whose suggestions helped improving the paper.

REFERENCES

Acha, E. M., Mianzan, H. W., Iribarne, O., Gagliardini, D. A., Lasta, C., and Daleo, P. (2003). The role of the Rio de la Plata bottom salinity front in accumulating debris. *Mar. Pollut. Bull*. 46, 197–202. doi: 10.1016/S0025-326X(02)00356-9

Al-Zawaidah, H., Ravazzolo, D., and Friedrich, H. (2021). Macroplastics in rivers: present knowledge, issues and challenges. *Environ. Sci*. 23:535, doi: 10.1039/D0EM00517G

Andrade, M. C., Winemiller, K. O., Barbosa, P. S., Fortunati, A., Chelazzi, D., Cincinelli, A., et al. (2019). First account of plastic pollution impacting freshwater fishes in the Amazon: ingestion of plastic debris by piranhas and other serrasalmids with diverse feeding habits. *Environ. Pollut*. 244, 766–773. doi: 10.1016/j.envpol.2018.10.088

Andrady, A. L. (2011). Microplastics in the marine environment. *Mar. Pollut. Bull*. 62, 1596–1605. doi: 10.1016/j.marpolbul.2011.05.030

Bai, M., Zhu, L., An, L., Peng, G., and Li, D. (2018). Estimation and prediction of plastic waste annual input into the sea from China. *Acta Oceanologica Sinica* 37, 26–39. doi: 10.1007/s13131-018-1279-0

Ballatore, A., Verhagen, T. J., Li, Z., and Cucurachi, S. (2021). This city is not a bin: crowdmapping the distribution of urban litter. *J. Ind. Ecol*. 1–16. Available online at: https://doi.org/10.1111/jiec.13164

Barnes, D. K., Galgani, F., Thompson, R. C., and Barlaz, M. (2009). Accumulation and fragmentation of plastic debris in global environments. *Philos. Trans. R. Soc. Lond., B, Biol. Sci*. 364, 1985–1998. doi: 10.1098/rstb.2008.0205

Bellasi, A., Binda, G., Pozzi, A., Galafassi, S., Volta, P., and Bettinetti, R. (2020). Microplastic contamination in freshwater environments: a review, focusing on interactions with sediments and benthic organisms. *Environments* 7:30. doi: 10.3390/environments7040030

Biermann, L., Clewley, D., Martinez-Vicente, V., and Topouzelis, K. (2020). Finding plastic patches in coastal waters using optical satellite data. *Sci. Rep*. 10, 1–10. doi: 10.1038/s41598-020-62298-z

Borrelle, S. B., Ringma, J., Law, K. L., Monnahan, C. C., Lebreton, L., McGivern, A., et al. (2020). Predicted growth in plastic waste exceeds efforts to mitigate plastic pollution. *Science* 369, 1515–1518. doi: 10.1126/science.aba3656

Broere, S., van Emmerik, T., González-Fernández, D., Luxemburg, W., de Schipper, M., Cózar, A., et al. (2021). Towards underwater macroplastic monitoring using echo sounding. *Front. Earth Sci*. 9:628704. doi: 10.3389/feart.2021.628704

Bruge, A., Barreau, C., Carlot, J., Collin, H., Moreno, C., and Maison, P. (2018). Monitoring litter inputs from the Adour River (Southwest France) to the marine environment. *J. Mar. Sci. Eng*. 6:24. doi: 10.3390/jmse6010024

Castro-Jiménez, J., González-Fernández, D., Fornier, M., Schmidt, N., and Sempéré, R. (2019). Macro-litter in surface waters from the Rhone River: plastic pollution and loading to the NW Mediterranean Sea. *Mar. Pollut. Bull*. 146, 60–66. doi: 10.1016/j.marpolbul.2019.05.067

Cesarini, G., and Scalici, M. (2022). Riparian vegetation as a trap for plastic litter. *Environ. Pollut*. 292:118410.

Ciszewski, D., and Grygar, T. M. (2016). A review of flood-related storage and remobilization of heavy metal pollutants in River Systems. *Water Air Soil Pollut*. 227:239doi: 10.1007/s11270-016-2934-8

Constant, M., Alary, C.l., De Waele, I., Dumoulin, D., Breton, N., and Billon, G. (2021). To what extent can micro- and macroplastics be trapped in sedimentary particles? A case study investigating dredged sediments. *Environ. Sci. Technol*. 55, 5898–5905 doi: 10.1021/acs.est.0c08386

Dawson, A. L., Kawaguchi, S., King, C. K., Townsend, K. A., King, R., Huston, W. M., et al. (2018). Turning microplastics into nanoplastics through digestive fragmentation by Antarctic krill. *Nat. Commun*. 9, 1–8. doi: 10.1038/s41467-018-03465-9

Delorme, A. E., Koumba, G. B., Roussel, E., Delor-Jestin, F., Peiry, J. L., Voldoire, O., et al. (2021). The life of a plastic butter tub in riverine environments. *Environ. Pollut*. 287:117656. doi: 10.1016/j.envpol.2021.117656

Duncan, E. M., Davies, A., Brooks, A., Chowdhury, G. W., Godley, B. J., Jambeck, J., et al. (2020). Message in a bottle: open source technology to track the movement of plastic pollution. *PLoS ONE* 15:e0242459. doi: 10.1371/journal.pone.0242459

Egessa, R., Nankabirwa, A., Basooma, R., and Nabwire, R. (2020). Occurrence, distribution and size relationships of plastic debris along shores and sediment of northern Lake Victoria. *Environ. Pollut*. 257:113442. doi: 10.1016/j.envpol.2019.113442

Enders, K., Käppler, A., Biniasch, O., Feldens, P., Stollberg, N., Lange, X., et al. (2019). Tracing microplastics in aquatic environments based on sediment analogies. *Sci. Rep*. 9:15207doi: 10.1038/s41598-019-50508-2

Faure, F., Demars, C., Wieser, O., Kunz, M., and De Alencastro, L. F. (2015). Plastic pollution in Swiss surface waters: nature and concentrations, interaction with pollutants. *Environ. Chem*. 12, 582–591. doi: 10.1071/EN14218

Ford, H. V., Jones, N. H., Davies, A. J., Godley, B. J., Jambeck, J. R., Napper, I. E., et al. (2022). The fundamental links between climate change and marine plastic pollution. *Sci. Total Environ*. 806:1150392. doi: 10.1016/j.scitotenv.2021.150392

Gabbott, S., Russell, C., Yohan, Y., and Zalasiewicz, J. (2020). The geography and geology of plastics: their environmental distribution and fate. *Plast. Waste Recycl*. 2020, 33–63 doi: 10.1016/B978-0-12-817880-5.00003-7

Gasperi, J., Dris, R., Bonin, T., Rocher, V., and Tassin, B. (2014). Assessment of floating plastic debris in surface water along the Seine River. *Environ. Pollut*. 195, 163–166. doi: 10.1016/j.envpol.2014.09.001

González-Fernández, D., Cózar, A., Hanke, G., Viejo, J., Morales-Caselles, C., Bakiu, R., et al. (2021). Floating macrolitter leaked from Europe into the ocean. *Nat. Sustain*. 4, 474–483. doi: 10.1038/s41893-021-00722-6

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world's free-flowing rivers. *Nature* 569, 215–221. doi: 10.1038/s41586-019-1111-9

Haberstroh, C. J., Arias, M. E., Yin, Z., and Wang, M. C. (2021). Effects of urban hydrology on plastic transport in a subtropical river. *ACS EST Water* 8, 1714–1727. doi: 10.1021/acsestwater.1c00072

Honingh, D., van Emmerik, T., Uijttewaal, W., Kardhana, H., Hoes, O., and van de Giesen, N. (2020). Urban river water level increase through plastic waste accumulation at a rack structure. *Front. Earth Sci*. 8:28. doi: 10.3389/feart.2020.00028

Hurley, R., Woodward, J., and Rothwell, J. J. (2018). Microplastic contamination of river beds significantly reduced by catchment-wide flooding. *Nat. Geosci*. 11, 251–257. doi: 10.1038/s41561-018-0080-1

Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., et al. (2015). Plastic waste inputs from land into the ocean. *Science* 347, 768–771. doi: 10.1126/science.1260352

Jang, Y. C., Lee, J., Hong, S., Mok, J. Y., Kim, K. S., Lee, Y. J., et al. (2014). Estimation of the annual flow and stock of marine debris in South Korea for management purposes. *Mar. Pollut. Bull*. 86, 505–511. doi: 10.1016/j.marpolbul.2014.06.021

Lebreton, L., Egger, M., and Slat, B. (2019). A global mass budget for positively buoyant macroplastic debris in the ocean. *Sci. Rep*. 9, 1–10. doi: 10.1038/s41598-019-49413-5

Lechthaler, S., Esser, V., Schüttrumpf, H., and Stauch, G. (2021). Why analysing microplastics in floodplains matters: application in a sedimentary context. *Environ. Sci*. 23, 117–131 doi: 10.1039/D0EM00431F

Lechthaler, S., Waldschläger, K., Stauch, G., and Schüttrumpf, H. (2020). The way of macroplastic through the environment. *Environments* 7:73doi: 10.3390/environments7100073

Leslie, H. A., Brandsma, S. H., Van Velzen, M. J. M., and Vethaak, A. D. (2017). Microplastics en route: field measurements in the Dutch river delta and Amsterdam canals, wastewater treatment plants, North Sea sediments and biota. *Environ. Int*. 101, 133–142. doi: 10.1016/j.envint.2017.01.018

Liro, M., van Emmerik, T., Wyzga, B., Liro, J., and Miku,ś, P. (2020). Macroplastic storage and remobilization in rivers. *Water* 12:2055. doi: 10.3390/w12072055

Malik, N. K. A., Abd Manaf, L., Jamil, N. R., Rosli, M. H., Ash'aari, Z. H., and Adhar, A. S. M. (2020). Variation of floatable litter load and its compositions captured at floating debris boom (FDB) structure. *J. Mater. Cycles Waste Manag*. 22, 1744–1767. doi: 10.1007/s10163-020-01065-8

Martin, C., Baalkhuyur, F., Valluzzi, L., Saderne, V., Cusack, M., Almahasheer, H., et al. (2020). Exponential increase of plastic burial in mangrove sediments as a major plastic sink. *Sci. Adv*. 6:eaaz5593. doi: 10.1126/sciadv.aaz5593

Mateos-Cárdenas, A., O'Halloran, J., van Pelt, F. N., and Jansen, M. A. (2020). Rapid fragmentation of microplastics by the freshwater amphipod Gammarus duebeni (Lillj.). *Sci. Rep*. 10, 1–12. doi: 10.1038/s41598-020-69635-2

Meijer, L. J., van Emmerik, T., van der Ent, R., Schmidt, C., and Lebreton, L. (2021). More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean. *Sci. Adv*. 7:eaaz5803. doi: 10.1126/sciadv.aaz5803

Mihai, F.-C. (2018). Rural plastic emissions into the largest mountain lake of the Eastern Carpathians. *R. Soc. Open Sci*. 5:172396. doi: 10.1098/rsos.172396

Moore, C. J., Lattin, G. L., and Zellers, A. F. (2011). Quantity and type of plastic debris flowing from two urban rivers to coastal waters and beaches of Southern California. *J. Integr. Coast. Zone Manag*. 11, 65–73. doi: 10.5894/rgci194

Newbould, R. A., Powell, D. M., and Whelan, M. J. (2021). Macroplastic debris transfer in rivers: a travel distance approach. *Front. Water* 3:724596. doi: 10.3389/frwa.2021.724596

Ockelford, A., Cundy, A., and Ebdon, J. E. (2020). Storm response of fluvial sedimentary microplastics. *Sci. Rep*. 10:1865doi: 10.1038/s41598-020-58765-2

Roebroek, C. T., Harrigan, S., Van Emmerik, T. H., Baugh, C., Eilander, D., Prudhomme, C., et al. (2021b). Plastic in global rivers: are floods making it worse? *Environ. Res. Lett*. 16:025003. doi: 10.1088/1748-9326/abd5df

Roebroek, C. T., Hut, R., Vriend, P., De Winter, W., Boonstra, M., and Van Emmerik, T. H. (2021a). Disentangling variability in riverbank macrolitter observations. *Environ. Sci. Technol*. 55, 4932–4942. doi: 10.1021/acs.est.0c08094

Ryan, P. G., and Perold, V. (2021). Limited dispersal of riverine litter onto nearby beaches during rainfall events. *Estuarine Coast. Shelf Sci*. 251:107186. doi: 10.1016/j.ecss.2021.107186

Saarni, S., Hartikainen, S., Meronen, S., Uurasjärvi, E., Kalliokoski, M., and Koistinen, A. (2021): Sediment trapping - an attempt to monitor temporal variation of microplastic flux rates in aquatic systems. *Environ. Pollut*. 274:116568. doi: 10.1016/j.envpol.2021.116568

Savenije, H. H. G. (2012). *Salinity and Tides in Alluvial Estuaries, 2nd Edn*. Delft: Delft University of Technology.

Schöneich-Argent, R. I., Dau, K., and Freund, H. (2020). Wasting the North Sea?–a field-based assessment of anthropogenic macrolitter loads and emission rates of three German tributaries. *Environ. Pollut*. 263:114367. doi: 10.1016/j.envpol.2020.114367

Schöneich-Argent, R. I., and Freund, H. (2020). Trashing our own "backyard" –investigating dispersal and accumulation of floating litter from coastal, riverine, and offshore sources in the German Bight using a citizen science-based wooden drifter recapture approach. *Mar. Environ. Res*. 162:105115. doi: 10.1016/j.marenvres.2020.105115

Schreyers, L., van Emmerik, T., Luan Nguyen, T., Castrop, E., Phung, N. A., Kieu-Le, T. C., et al. (2021). Plastic plants: the role of water hyacinths in plastic transport in tropical rivers. *Front. Environ. Sci*. 9:177. doi: 10.3389/fenvs.2021.686334

Skalska, K., Ockelford, A., Ebdon, J. E., and Cundy, A. B. (2020). Riverine microplastics: Behaviour, spatio-temporal variability, and recommendations for standardised sampling and monitoring. *J. Water Process Eng*. 38:101600. doi: 10.1016/j.jwpe.2020.101600

Song, Y. K., Hong, S. H., Jang, M., Han, G. M., Jung, S. W., and Shim, W. J. (2017). Combined effects of UV exposure duration and mechanical abrasion on microplastic fragmentation by polymer type. *Environ. Sci. Technol*. 51, 4368–4376. doi: 10.1021/acs.est.6b06155

Tasseron, P., van Emmerik, T., Peller, J., Schreyers, L., and Biermann, L. (2021). Advancing floating macroplastic detection from space using experimental hyperspectral imagery. *Remote Sens*. 13:2335. doi: 10.3390/rs13122335

Tasseron, P., Zinsmeister, H., Rambonnet, L., Hiemstra, A. F., Siepman, D., and van Emmerik, T. (2020). Plastic hotspot mapping in urban water systems. *Geosciences* 10:342. doi: 10.3390/geosciences10090342

Tramoy, R., Colasse, L., Gasperi, J., and Tassin, B. (2019). Plastic debris dataset on the Seine river banks: plastic pellets, unidentified plastic fragments and plastic sticks are the Top 3 items in a historical accumulation of plastics. *Data Brief* 23:103697. doi: 10.1016/j.dib.2019.01.045

Tramoy, R., Gasperi, J., Colasse, L., Silvestre, M., Dubois, P., Noûs, C., et al. (2020a). Transfer dynamics of macroplastics in estuaries–new insights from the Seine estuary: part 2. *Short-term dynamics based on GPS-trackers. Mar. Pollut. Bull*. 160:111566. doi: 10.1016/j.marpolbul.2020.111566

Tramoy, R., Gasperi, J., Colasse, L., and Tassin, B. (2020b). Transfer dynamic of macroplastics in estuaries—New insights from the Seine estuary: Part 1. Long term dynamic based on date-prints on stranded debris. *Mar. Pollut. Bull*. 152:110894. doi: 10.1016/j.marpolbul.2020.110894

Treilles, R., Gasperi, J., Gallard, A., Mohamed, S. A. A. D., Rachid, D. R. I. S., Partibane, C., et al. (2021). Microplastics and microfibers in urban runoff from a suburban catchment of Greater Paris. *Environ. Pollut*. 287:117352. doi: 10.1016/j.envpol.2021.117352

Treilles, R., Gasperi, J., Tramoy, R., Dris, R., Gallard, A., Partibane, C., et al. (2022). Microplastic and microfiber fluxes in the Seine River: flood events versus dry periods. *Sci. Total Environ*. 805:150123. doi: 10.1016/j.scitotenv.2021.150123

van Calcar, C. J., and van Emmerik, T. H. M. (2019). Abundance of plastic debris across European and Asian rivers. *Environ. Res. Lett*. 14:124051. doi: 10.1088/1748-9326/ab5468

van Emmerik, T., Loozen, M., Van Oeveren, K., Buschman, F., and Prinsen, G. (2019b). Riverine plastic emission from Jakarta into the ocean. *Environ. Res. Lett*. 14:084033. doi: 10.1088/1748-9326/ab30e8

van Emmerik, T., Roebroek, C., De Winter, W., Vriend, P., Boonstra, M., and Hougee, M. (2020a). Riverbank macrolitter in the Dutch Rhine–Meuse delta. *Environ. Res. Lett*. 15:104087. doi: 10.1088/1748-9326/abb2c6

van Emmerik, T., and Schwarz, A. (2020). Plastic debris in rivers. *WIREs Water* 7:e1398doi: 10.1002/wat2.1398

van Emmerik, T., Seibert, J., Strobl, B., Etter, S., Den Oudendammer, T., Rutten, M., et al. (2020b). Crowd-based observations of riverine macroplastic pollution. *Front. Earth Sci*. 8:298. doi: 10.3389/feart.2020.00298

van Emmerik, T., Strady, E., Kieu-Le, T. C., Nguyen, L., and Gratiot, N. (2019a). Seasonality of riverine macroplastic transport. *Sci. Rep*. 9, 1–9. doi: 10.1038/s41598-019-50096-1

van Emmerik, T., Tramoy, R., van Calcar, C., Alligant, S., Treilles, R., Tassin, B., et al. (2019c). Seine plastic debris transport tenfolded during increased river discharge. *Front. Mar. Sci*. 6:642. doi: 10.3389/fmars.2019.00642

van Emmerik, T., Van Klaveren, J., Meijer, L. J., Krooshof, J. W., Palmos, D. A. A., and Tanchuling, M. A. (2020c). Manila river mouths act as temporary sinks for macroplastic pollution. *Front. Mar. Sci*. 7:770. doi: 10.3389/fmars.2020.545812

van Emmerik, T., Vriend P., and Copius Peereboom, E. (2022). Roadmap for long-term macroplastic monitoring in rivers. *Front. Environ. Sci*. 9:802245. doi: 10.3389/fenvs.2021.802245

Vriend, P., Van Calcar, C., Kooi, M., Landman, H., Pikaar, R., and Van Emmerik, T. (2020). Rapid assessment of floating macroplastic transport in the Rhine. *Front. Mar. Sci*. 7:10. doi: 10.3389/fmars.2020.00010

Waldschläger, K., Born, M., Cowger, W., Gray, A., and Schüttrumpf, H. (2020). Settling and rising velocities of environmentally weathered micro-and macroplastic particles. *Environ. Res*. 191:110192. doi: 10.1016/j.envres.2020.110192

Watkins, L., McGrattan, S., Sullivan, P. J., and Walter, M. T. (2019). The effect of dams on river transport of microplastic pollution. *Sci. Total Environ*. 664:834–840. doi: 10.1016/j.scitotenv.2019.02.028

Weideman, E. A., Perold, V., and Ryan, P. G. (2020). Limited long-distance transport of plastic pollution by the Orange-Vaal River system, South Africa. *Sci. Total Environ*. 727:138653. doi: 10.1016/j.scitotenv.2020.138653

Williams, A. T., and Simmons, S. L. (1996). The degradation of plastic litter in rivers: implications for beaches. *J. Coastal Conserv*. 2, 63–72. doi: 10.1007/BF02743038

Zbyszewski, M., Corcoran, P. L., and Hockin, A. (2014). Comparison of the distribution and degradation of plastic debris along shorelines of the Great Lakes, North America. *J. Great Lakes Res*. 40, 288–299. doi: 10.1016/j.jglr.2014.02.012

Zhang, K., Gong, W., Lv, J., Xiong, X., and Wu, C. (2015). Accumulation of floating microplastics behind the Three Gorges Dam. *Environ. Pollut*. 204, 117–123. doi: 10.1016/j.envpol.2015.04.023

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.

frontiers in Water

ORIGINAL RESEARCH
published: 15 February 2022
doi: 10.3389/frwa.2021.750802

Assessing Microbial Water Quality, Users' Perceptions and System Functionality Following a Combined Water Safety Intervention in Rural Nepal

Carola Bänziger [1]†, Ariane Schertenleib [1], Bal Mukunda Kunwar [2], Madan R. Bhatta [2] and Sara J. Marks [1]*

[1] *EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland,* [2] *Helvetas Swiss Intercooperation, Kathmandu, Nepal*

Risk-based water safety interventions are one approach to improve drinking water quality and consequently reduce the number of people consuming faecally contaminated water. Despite broad acceptance of water safety planning approaches globally, there is a lack of evidence of their effectiveness for community-managed piped water supplies in rural areas of developing countries. Our research, in the form of a cluster-based controlled pre-post intervention analysis, investigated the impact of a combined water safety intervention on outcomes of microbial water quality, users' perceptions and piped system functionality in rural Nepal. The study enrolled 21 treatment systems and 12 control systems across five districts of the Karnali and Sudurpaschim provinces. Treatment group interventions included field laboratories for microbial analysis, regular monitoring of water quality including sanitary inspections, targeted treatment and infrastructure improvements, household hygiene and water filter promotion, and community training. In certain systems, regular system-level chlorination was implemented. Before and after the interventions, the microbial water quality was measured at multiple points within the water system. This information was complemented by household interviews and sanitary inspections. The main result to emerge from this study is that chlorination is the only identified intervention that led to a significant reduction in *E. coli* concentration at the point of consumption. Secondly, the effectiveness of other interventions was presumably reduced due to higher contamination at endline in general, brought about by the monsoon. All the interventions had a positive impact on users' perceptions about their water system, as measured by expectations for future functionality, satisfaction with the services received, and awareness of the potential health risks of drinking contaminated water. For future applications we would recommend the more broadly applied use of chlorination methods at system level as a key component of the package of risk-based water safety interventions.

Keywords: drinking water, water safety planning, microbial quality, piped systems, rural communities, Nepal

Edited by:
Abbas Roozbahani,
University of Tehran, Iran

Reviewed by:
Katarzyna Pietrucha-Urbanik,
Rzeszów University of Technology, Poland
Lisa Paruch,
Norwegian Institute of Bioeconomy Research (NIBIO), Norway

***Correspondence:**
Carola Bänziger
carola.baenziger@fhnw.ch

†Present address:
Carola Bänziger,
FHNW, University of Applied Sciences Northwestern Switzerland, Muttenz, Switzerland

Specialty section:
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

Received: *31 July 2021*
Accepted: *20 December 2021*
Published: *15 February 2022*

Citation:
Bänziger C, Schertenleib A, Kunwar BM, Bhatta MR and Marks SJ (2022) Assessing Microbial Water Quality, Users' Perceptions and System Functionality Following a Combined Water Safety Intervention in Rural Nepal. Front. Water 3:750802. doi: 10.3389/frwa.2021.750802

INTRODUCTION

Access to safe drinking water is a basic human right still denied to many. In 2017, 93% of the global population had access to an improved drinking water source, while 579 million people still used unimproved sources such as unprotected springs, open wells and surface water (WHO/UNICEF Joint Monitoring Programme, 2019). However, an improved water source does not always ensure adequate drinking water quality (Bain et al., 2014a; Shaheed et al., 2014). Although improved sources such as piped supplies or protected point sources have the potential to deliver safe water by nature of their design and construction (WHO/UNICEF Joint Monitoring Programme, 2019), this alone does not guarantee safe water.

In order to reliably assess the safety of drinking water, data on its microbial quality is necessary (Baum et al., 2014). However, many countries lack representative data on drinking water quality (WHO/UNICEF Joint Monitoring Programme, 2019). It has been estimated that around 10% of improved water sources are heavily faecally contaminated, with >100 *Escherichia coli* (*E. coli*) colony forming units (CFU)/100 mL (Bain et al., 2014b). This corresponds to 1.8 billion people globally using a faecally contaminated drinking water source (Bain et al., 2014b). These studies suggest that having access to an improved source is not enough to ensure safe drinking water and meet the global target 6.1 of the Sustainable Development Goals (SDGs) to "achieve universal and equitable access to safe and affordable drinking water for all" by 2030. To meet the criteria, households must use an improved water source that is accessible on premises, available when needed, and free from fecal and priority chemical contamination (WHO/UNICEF Joint Monitoring Programme, 2019).

This study focuses on rural communities in Karnali and Sudurpaschim provinces, in Western Nepal, where the majority of the population relies on a faecally contaminated drinking water source (WHO/UNICEF Joint Monitoring Programme, 2019). Nepal had 29.3 million inhabitants in 2017, of which 81% were living in rural areas (WHO/UNICEF Joint Monitoring Programme, 2019). The Human Development Index ranks Nepal 147th out of 189 countries and territories (UNDP, 2019). Although the majority of people in Nepal have water on premises (65%) and available when needed (82%), only 29% of the population had access to a water source free from *E. coli* contamination at the point of collection in 2014 (WHO/UNICEF Joint Monitoring Programme, 2019). Water quality deteriorates further at point of consumption, leaving only 18% of the population with safe water (WHO/UNICEF Joint Monitoring Programme, 2019).

10% of the country's population are children under the age of 5 years (Government of Nepal, 2017). Diarrhea is one of the most common illnesses among children in Nepal and continues to be a major cause of childhood morbidity and mortality (Government of Nepal, 2018). The diarrhea prevalence among children under the age of 5 was 8% in 2016 (Government of Nepal, 2017), with the highest incidence in the Karnali province (Government of Nepal, 2018). It is difficult to obtain precise data on childhood mortality due to diarrhoeal diseases in Nepal. However, it has been estimated that around 25% of all child deaths are associated with acute diarrhea (Bista, 2001).

To confront the health risks due to compromised water quality, risk assessment and risk management approaches have been applied to rural drinking water systems in a variety of ways and with varying degrees of success. The establishment of a water safety plan (WSP) is one mean to reduce risks of contamination (Mahmud et al., 2007) and comprises all steps of a water supply system from the catchment to the consumer (WHO, 2017a). Experience on the implementation of WSPs is mostly documented for water distribution systems managed by water utilities in high income contexts. The implementation of WSPs for small piped systems and in the context of rural areas in low- and middle-income countries (LMICs) is less described (Mahmud et al., 2007). The few existing studies showed contrasting results. While String et al. (2020) found more *E. coli* contamination in the WSP communities than in the non-WSP communities, Mahmud et al. (2007) observed improvements in microbial water quality related to the WSPs; however, such improvements were not uniform. Other studies mostly focused on the set-up of WSPs and how they can be adapted to rural community-managed systems (Barrington et al., 2013), or on how their impact can be measured (Kumpel et al., 2018), rather than on systematically assessing the actual impact. These studies highlighted the need for more robust and standardized methods and indicators to assess WSP outcomes and impacts (Kumpel et al., 2018).

The development literature also identified prerequisites and challenges for the implementation of WSPs in LMICs. For implementation of WSPs to be successful, community education, behavior change, and the distribution of simplified documentation are essential (Barrington et al., 2013). Capacity building, according to Ferrero et al. (2019), is a crucial factor for the successful implementation of a WSP (Ferrero et al., 2019). Others have found that financial constraints often pose challenges to the implementation of WSPs (Kumpel et al., 2018), despite WSPs being relatively inexpensive, adaptable, and requiring minimal water quality testing (Barrington et al., 2013).

The present study is motivated by the aforementioned evidence gaps regarding the requisite enabling conditions and potential benefits of risk-based water safety strategies for water supplies in remote rural areas. In the frame of this research, a combination of water safety interventions, informed by risk assessment and risk management practices with Helvetas-Nepal's Integrated Water Resource Management (IWRM) Programme, were adapted to the local context and implemented in 21 water supply systems in rural Nepal. We use a cluster-based controlled pre-post intervention design to assess the impact of this package of water safety interventions on multiple outcomes: microbial water quality, users' perception of their water services and the functionality of the water system.

STUDY SITE AND METHODS

The Integrated Water Resource Management Programme

This study was conducted across five districts (Achham, Kalikot, Jajarkot, Dailekh, and Surkhet) in Sudurpaschim and Karnali

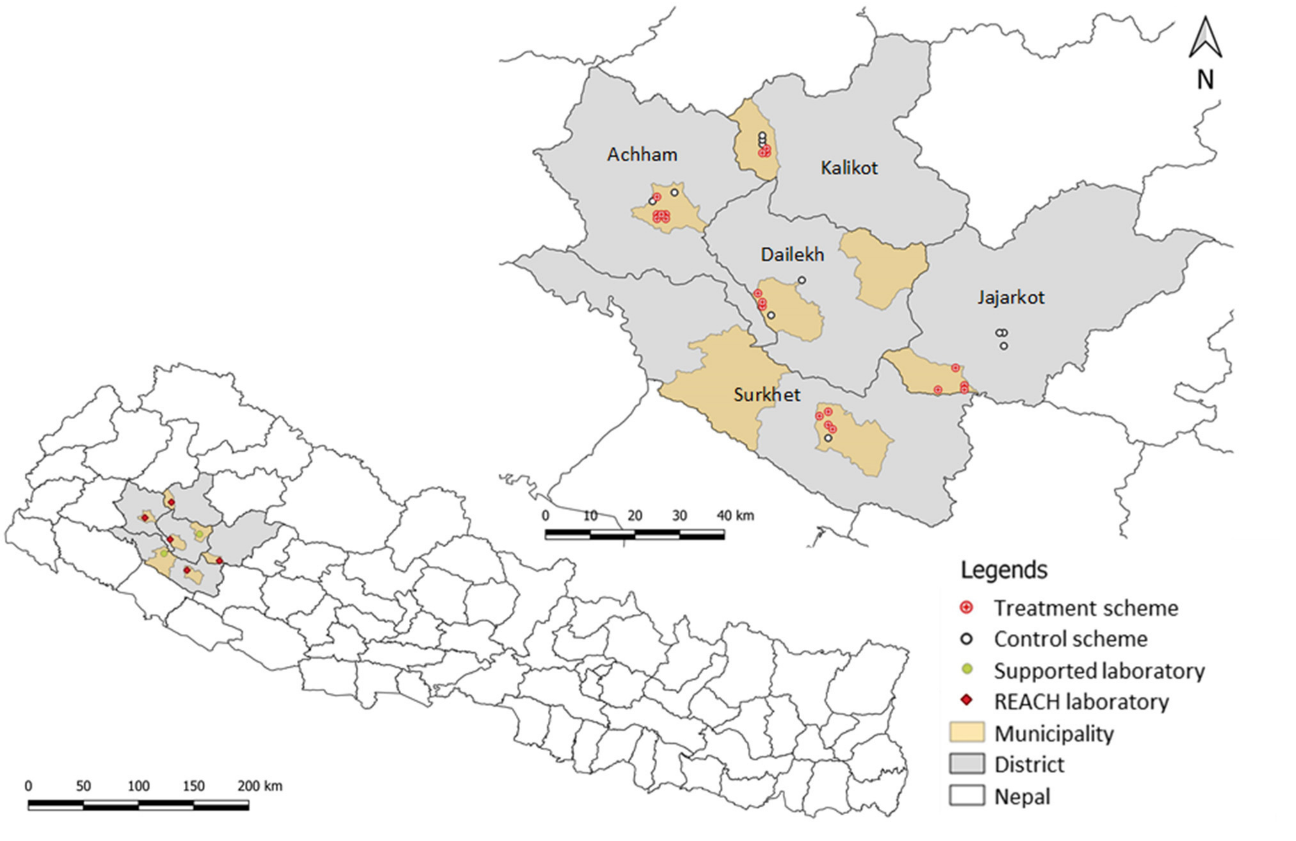


FIGURE 1 | Map of Nepal with the five districts involved in the Integrated Water Resource Management (IWRM) Programme, showing the locations of the treatment and control drinking water systems (DWS) and the laboratories implemented for this study (REACH laboratories).

provinces, in the hilly region of Nepal (see **Figure 1**). Between the years 2011 and 2017 Helvetas Swiss Intercooperation-Nepal (hereafter referred to as Helvetas) constructed drinking water systems (DWS) and implemented total sanitation interventions across these districts in the frame of their Integrated Water Resource Management (IWRM) Programme. In all districts, the construction of the DWS was followed by a series of interventions targeted at improving the water, sanitation and hygiene (WASH) conditions locally. Specifically, this consisted of the establishment of a water and sanitation users' committee, the promotion of improved household hygiene practices, the use of ceramic candle filters for household water treatment and safe storage, and the training of a female community health volunteer on a total sanitation approach and a village maintenance worker to operate the DWS (Marks and Shrestha, 2020). Some systems provide continuous (24 h) service, while others provide intermittent service with variable opening times and service durations throughout the year. The systems all have similar branched design layouts, consisting of a spring source connected to a reservoir tank by a distribution line, with water then flowing to private or public taps as described by Tosi Robinson et al. (2018) (**Figure 2**).

Community Selection and Intervention Package

Within the five districts, 33 communities were selected based on the existence of a functioning DWS and the agreement of the community to participate. Within these, 21 were defined as treatment systems and 12 as control systems. System selection occurred in parallel to the planning and selection of the locations of the field laboratories. A criterion for the choice of treatment systems was the access to the nearest laboratory within 2 h ours for the furthest sampling point. Control systems were selected in such a way that they would not be affected by the interventions from treatment systems, nor influence the treatment systems themselves.

In the treatment systems, different interventions took place between pre- and post-treatment data collection. The interventions included the setting up of field laboratories for microbial analysis, regular monitoring of water quality including sanitary inspections and standard questionnaires, centralized data management, targeted infrastructure improvements such as source protection, installation of roughing sand filters in front of the reservoir tanks, chlorination of reservoir tanks, pipeline improvement and maintenance, and local watershed restoration. Further, training of community water safety task forces and

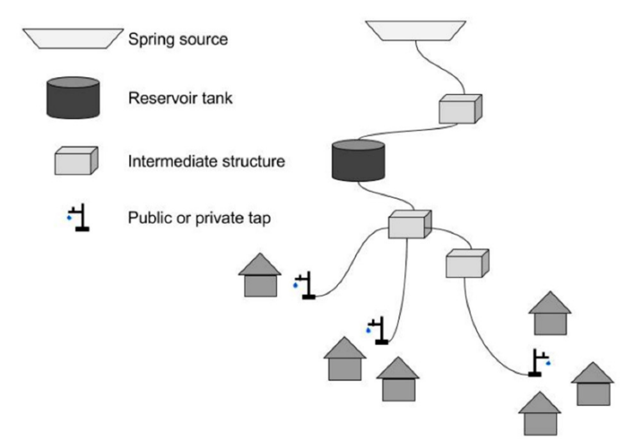


FIGURE 2 | Sketch of a typical gravity-fed piped water system or sub-system in the study area. A water source can feed one or several reservoir tanks that distribute water to different taps. The intermediate structures can be distribution, collection, purge valve, air valve or interruption chambers, or break pressure tanks (Image source: Tosi Robinson et al., 2018).

laboratory technicians as well as promotion of household hygiene and ceramic candle filters for point-of-use water treatments were part of the interventions. The regular monitoring involved monthly water quality analysis of *E. coli* and total coliform concentrations, visual turbidity, pH and free residual chlorine at one reservoir tank, one tap and one corresponding household per system. Quarterly sanitary inspections were collected at the protected spring, reservoir tank, tap and household, and brief household surveys were conducted during each inspection round. A subset of six systems, minimum one per district, were selected to receive regular system-level chlorination, either delivered manually in reservoir tanks or automatically using passive chlorinator devices. In the control systems, none of the interventions described above was implemented.

Data Collection

The study included data collection at both pre-treatment (April–June 2018) and post-treatment time points (August–October 2019), hereafter referred to as baseline and endline. The study design was akin to that of Tosi Robinson et al., 2018. The field teams were composed of staff members of Eawag, Helvetas, and local non-governmental organizations (NGOs). The teams conducted household and key informant interviews, collected a water sample (100 mL) in each of the interviewed households at the point of consumption, and performed sanitary inspections. The sanitary inspections were carried out in each system of the reservoir tank(s), the pipelines, taps, and if present, structures like roughing sand filter, break pressure tank, distribution chamber and intake. Water samples were collected during the sanitary inspections at the inlet of all reservoir tanks, six randomly selected taps, and from other structures (for example, inlet of roughing sand filter and distribution chamber) if present and accessible.

Household Surveys

From each of the 33 water systems, a random selection of 15 households was enrolled resulting in a total of 493 surveys (in one system it was not possible to reach 15). All the questionnaires were translated to and conducted in Nepali using tablets (Samsung Galaxy Tab A, Seoul, Korea) with ODK software (open data kit, https://www.opendatakit.org/). Only households using the IWRM Programme drinking water system were interviewed. Eligible households were selected randomly from the village's household list and enrolled following informed consent about the Programme's purpose and anonymity of the questionnaire. At the study baseline, if the household declined to participate in the study or if no adult was available at the time of the visit, another household was selected randomly as a replacement. During the endline period, enumerators requested to interview the same person as during baseline. If this person was not available, another person from the same household was interviewed. If nobody within the same household was present (e.g., due to migration of the entire household), a neighboring household was interviewed. The survey questions probed the households' drinking water supply characteristics, sanitation and hygiene practices, socio-economic statuses, users' perception, and sense of ownership for their water system.

Drinking Water Quality

Sample Collection

The method for water sampling and processing corresponds to the one described by Tosi Robinson et al. (2018). Briefly, in each household that participated in the survey, a 100 mL drinking water sample was collected as if study participants were providing a cup of water to drink. Water samples at the reservoir tanks were collected directly from the inlet, which is the closest point to the water source accessible for sampling; therefore, the sample collected is representative of the water entering but not of the water stored at the reservoir tank. For systems with a roughing sand filter installed upstream of the reservoir, the sample was collected at the inlet of the filter. At the taps, water was run for 30 s before sampling to wash out any deposited residue and ensure a representative sample from the piped system. All the water samples from a single system were collected on the same day in sterile 100 mL Whirl-Pak Thio-bags (Nasco, Fort Atkinson, USA) containing sodium thiosulfate to inactivate any residual chlorine.

Membrane Filtration

The samples were processed within 2 h after collection using the membrane filtration method, as described by Tosi Robinson et al. (2018). Briefly, each water sample was filtered through a sterilized filter funnel (DelAgua, UK) with a 0.45 µm Millipore cellulose membrane filter (Merck, Germany). The membrane was placed on Nissui EC Compact Dry plates (Nissui Pharmaceuticals, Japan) pre-moistened with sterile water (Nissui, 2009). The plates were incubated for 24 h at 35 +/− 2°C. The incubators used during this study were constructed as described in Schertenleib et al. (2019).

Colony Enumeration

After incubation of the EC Compact Dry plates, blue colonies were counted as *E. coli* colony forming units per 100 mL (CFU/100 mL) and the sum of purple and blue colonies as total coliform according to the manufacturer's instructions. Counts higher than 300 colonies per plate were reported as too numerous to count (TNTC). To estimate variation in the assay, one duplicate was performed for every seven household water samples, resulting in two household duplicates per system. In addition, a duplicate was collected from one randomly chosen scheme-level site (source, roughing sand filter, reservoir tank or tap) at each system. Positive and negative controls were processed daily.

pH and Free Residual Chlorine

Water testing included free residual chlorine (FRC) concentration and pH measurements using the Lovibond® Three-Chamber-Tester Chlorine LR-pH (Lovibond® Tintometer Group, England), following the manufacturer's instruction) for using DPD and Phenol Red. The visual test indicates pH values between 6.8 and 8.2 and chlorine concentration between 0.1 and 3.0 mg/L. Every tenth sample was processed in duplicate.

Data Analysis

Water quality, sanitary inspection and survey data were compiled and cleaned using Microsoft Excel 10 (Microsoft, Redmond, WA, USA). Coding and statistical tests of intervention effects were performed using IBM SPSS Statistics 25 (IBM, New York, NY, USA). For the bivariate analysis, microbial data were log-transformed after non-detect samples were set to half of the lower limit of detection (0.5 CFU/100 mL) and TNTC values were set to 300 CFU/100 mL. The central tendency of microbial concentrations and other variables are reported in terms of mean, standard deviation, and median values. Non-parametric tests (Wilcoxon signed-rank test, Mann-Whitney test and Spearman's correlation) were used for bivariate comparisons because most variables were not normally distributed based on the Shapiro-Wilk test. The Wilcoxon signed-rank test was used for comparison between baseline and endline and the Mann-Whitney test for comparison between the different systems at one time point.

In order to identify factors associated with fecal contamination of the drinking water, an ordinal logistic regression model was estimated with *E. coli* risk categories as dependent variable. Based on the results of the bivariate comparisons described above, we ran the ordinal logistic regression model with a reduced selection of the independent variables and the *E. coli* concentration at endline in ordinal risk categories, with risk tiers from lowest to highest as follows: <1 CFU/100 mL, 1–10 CFU/100 mL, 11–100 CFU/100 mL, and >100 CFU/100 mL. The likelihood ratio Pearson chi-square test and the deviance test were both used to conduct the Goodness-of-Fit analysis.

Ethics Statement

All participants gave their oral and written informed consent before the interviews. The research was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Eawag ethics committee (protocol 16_09_022018). The study received approval in Nepal as part of Helvetas-Nepal's IWRM research program.

RESULTS

Description of Study Population

The 33 water systems within this study served 29–250 households [Mean (M) = 67.4, Standard Deviation (SD) = 44.0], of which we interviewed 15 each. The respondents were 67% female and on average 38.4 years old (SD = 14.2). Mean household size was 6.4 people (SD = 2.5), of which 2.7 (SD = 1.6) were children under 18 years old, and 0.82 (SD = 0.87) under 5 years old. Less than half of the study participants (44%) completed primary or higher-level education, and 16% were illiterate. In almost all households (97%) there was at least one person working in the agricultural sector. The mean monthly regular expenditures per household were 10'861 NPR (SD = 7981) [equivalent to 96 USD (SD = 74) in October 2019], ranging from 1'000 to 80'000 NPR (9–704 USD). The houses were predominantly made of stone and mud (95%) with an earth floor (95%), and a stone slate roof (56%). The majority of households had access to electricity, either through their own solar panel (83%) or through connection to an electrical grid (15%), while 6% did not have access to electricity. Water supply (34%) was the biggest concern mentioned by people living in the village at baseline, followed by transportation and roads (23%). More details about the households' characteristics are described in **Supplementary Table S1** and **Supplementary Table S2** in **Supplementary Material**.

Water Supply and Sanitation Access

Most respondents had piped water connections and access to improved sanitation. For all study systems, Helvetas had previously implemented drinking water systems and total sanitation in the frame of their IWRM Programme between the years 2011 and 2017. In all 33 study communities the water system had a branched design, with 10 having private taps and 23 having public taps shared by multiple households. The total number of taps per system ranged from 5 to 250 (M = 30.8, SD = 43.1). The water originated from 1 to 5 different sources per system. There was at least 1 and up to 6 reservoir tanks per system (see **Supplementary Table S3** for more details). Households most commonly reported their main drinking water source to be piped village connections (71%) and piped household connections (26%). Other sources included protected sources, unmanaged piped waters, open sources, rainwater or river water (see **Supplementary Table S2** for additional information on main drinking water sources used). The average round trip time for gathering water was 10.7 min (SD = 8.5), with extremes ranging from 1 to 65 min, including time required for queuing and filling the containers. In addition to the main system, some respondents (21%) reported also using other sources for drinking water. Among these households, their main reasons for source switching were intermittent supply (52%), seasonal unavailability (33%), and closer proximity to another source (12%).

Most of the households had their own ventilated improved pit latrine (80%) and only one interviewed household indicated practicing open defecation at endline. About half of all interviewed households reported keeping their animals overnight in the same house where they lived (see **Supplementary Table S4** for more information).

Baseline to Endline Comparison

Water Quality

Point of Consumption

The visually assessed turbidity at the point of consumption at endline was clear in 98.2% of the samples and somewhat turbid in 1.8% of samples. In 7.2% of the samples debris was present. The pH was on average 7.3 (SD = 0.2), ranging from 6.8 to 8.2 with a median of 7.4. With 96.3% and 99.2% of the water samples at baseline and endline having a pH lower than 8, respectively, the vast majority was within the preferable pH for effective disinfection with chlorine (WHO, 2017b).

Among all treatment systems, there was no significant difference between baseline and post-treatment *E. coli* concentrations at the point of consumption (**Table 1**). Control systems exhibited an increasing contamination trend from baseline to endline, although the difference was also not statistically significant (**Figure 3**). When comparing between all treatment systems and the control systems at endline, we saw significantly less *E. coli* in treatment systems ($Z = -3.982$, $p < 0.001$). At baseline, there was no significant difference in *E. coli* in treatment systems as compared to control systems ($Z = -1.061$, $p = 0.289$).

The percentage of households that met the WHO drinking water quality guideline of <1 *E. coli* CFU/100 mL was about the same at baseline in the control and treatment systems at around 10% (see green category in **Figure 4** and **Supplementary Table S5** in the **Supplemental Material**). By endline in the treatment systems, the share of households with no detectable *E. coli* in their stored water containers increased significantly ($Z = -3.130$; $p = 0.002$) to 19.8%. By contrast, there was a decrease in the share of control systems' samples without detectable *E. coli* to 7.8%, but the difference from baseline was not statistically significant ($Z = -0.784$; $p = 0.433$). With an increase from 5.3% of the households with no detected *E. coli* at baseline to 78.9% at endline, the households with measurable chlorine showed the largest improvement in microbial water quality ($Z = -3.500$; $p < 0.001$) (**Figure 4**).

Point of Collection

Generally speaking, by study endline the mean *E. coli* concentrations were lower at the taps than at the point of consumption (**Table 2**). Taps with measurable chlorine concentration showed a significant decrease in *E. coli* concentration from baseline to endline ($Z = -2.185$, $p = 0.029$). However, for other comparisons, such as the changes in fecal contamination levels from baseline to endline among both control systems and in the treatment systems, as well as the differences in *E. coli* concentration at the tap between the control and treatment systems, were not statistically significant.

In the treatment systems there were significantly more taps that met the WHO standards of <1 *E. coli* CFU/100 mL at endline (15.9%) than at baseline (9.8%) ($Z = -2.000$, $p = 0.046$). Among the control systems, 6.9% had no detectable *E. coli* at endline and 13.9% at baseline. This decrease in households meeting the *E. coli* target was not statistically significant ($Z = -1.155$, $p = 0.248$). At all taps where free residual chlorine was measured at endline, samples had no detectable *E. coli* and hence 100% met the WHO guideline for microbial water safety.

Quality Control

We processed duplicates for one out of every seven samples to estimate variation in the assay, resulting in two duplicates per 15 household water samples. The log10 difference between the duplicates of the household water samples showed low variation in water quality as indicated by a mean of 0.05 (SD = 0.3; Median = 0.0) difference, corresponding to 5.4 CFU/100 mL (SD = 42.9; Median = 0.0). Negative controls were processed daily through the same filtration funnel as was a preceding positive

TABLE 1 | *E. coli* concentration at the point of consumption at baseline and endline in control and treatment systems, including comparison from baseline to endline with the Wilcoxon signed-rank test.

	Baseline			Endline			Wilcoxon signed-rank test	
	N	Median [CFU/100 mL]	Mean (SD) [log10(CFU/100 mL]	*N*	Median [CFU/100 mL]	Mean (SD) [log10(CFU/100 mL]	Z	*p*
Control systems	179	22	1.30 (0.93)	180	38	1.48 (0.90)	−1.567	0.117
Treatment systems (all)	313	24	1.22 (0.89)	308	16	1.13 (0.98)	−0.049	0.961
Treatment systems (intended chlorination)	90	10	1.05 (0.92)	90	3	0.67 (0.98)	−1.710	0.087
Treatment systems (measured chlorine)	19	25	1.33 (0.82)	19	0	-0.05 (0.70)	−2.811	0.005

Treatment systems are reported as: All systems, those with intended chlorination, and those with detectable chlorine at endline.

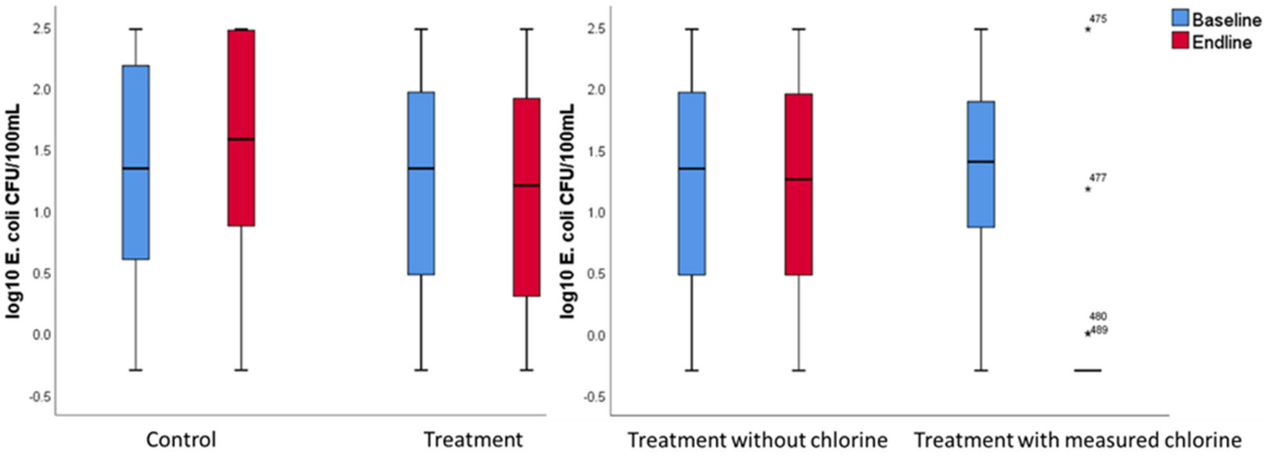


FIGURE 3 | Log10 *E. coli* concentration in CFU/100 mL at the point of consumption at baseline and endline in the control and treatment systems (left) and in treatment systems with and without measured chlorine at endline (right). Outlier values are represented by asterisks (*), with sample numbers shown.

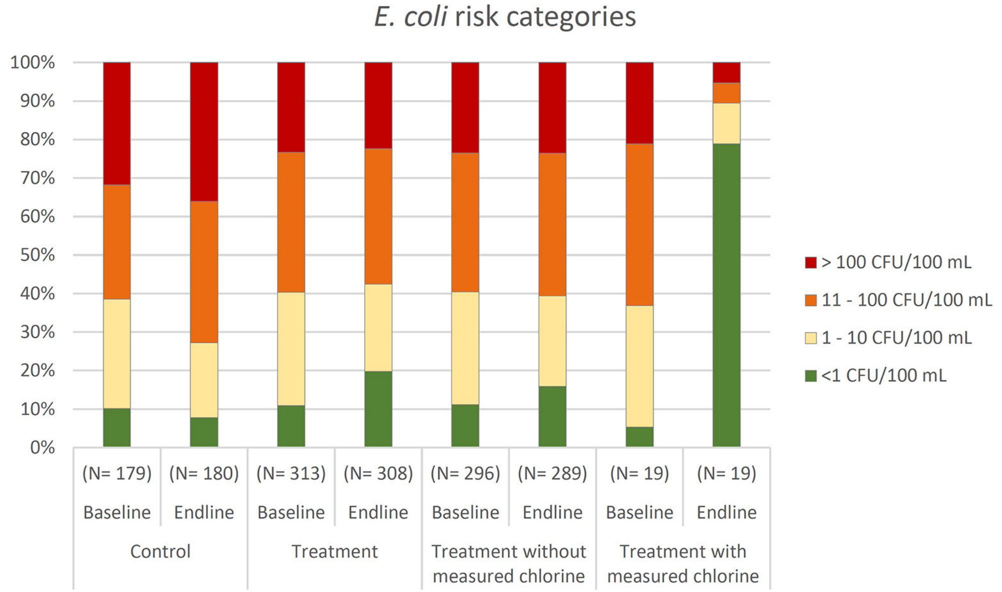


FIGURE 4 | Percentage of households within the different *E. coli* risk categories of low risk (<1 CFU/100 mL), intermediate risk (1–10 CFU/100 mL), high risk (11–100 CFU/100 mL), and very high risk (>100 CFU/100 mL) at baseline and endline in the control and treatment systems, including treatment systems with and without detectable chlorine at endline.

control. All negative control samples processed had no detectable *E. coli*.

System Functionality

Sanitary Inspections

At the time of the research team's visits the water systems were generally functioning well. According to the sanitary inspections, on average 79.2% of the taps per system were working at the endline visit (SD = 22.0), 7.5% were leaking (SD = 17.0), 8.2% broken (SD = 11.9) and 5.2% not functioning (SD = 8.1). The share of taps functioning, leaking or being broken did not differ significantly between control and treatment systems at endline. However, the control systems had a significantly higher share of taps not functioning than the treatment systems ($p = 0.036$).

Households' Perceptions of Functionality

In addition to the sanitary inspection, enumerators asked households for their opinion on the current functionality of their main drinking water source. Most respondents rated the water system as functioning well and only a few as not functioning (**Table 3**). In the treatment systems, there was a significant improvement in the functionality rating from baseline to endline, while no such difference was observed among control systems. At

TABLE 2 | *E. coli* concentrations at each sample location for the intervention and control systems, with bivariate comparisons of baseline and endline measurements.

		Treatment				Control				Chlorination			
		N	**Median CFU/100 mL**	**Mean (SD) log10(CFU/100 mL)**	**Wilcoxon signed-rank test**	***N***	**Median CFU/100 mL**	**Mean (SD) log10(CFU/100 mL)**	**Wilcoxon signed-rank test**	***N***	**Median CFU/100 mL**	**Mean (SD) log10(CFU/100 mL)**	**Wilcoxon signed-rank test**
Household	Baseline	313	24	1.22 (0.89)	$Z = -0.049$	179	22	1.30 (0.93)	$Z = -1.567$	19	25	1.33 (0.82)	$Z = -2.811$
	Endline	308	16	1.13 (0.98)	$p = 0.961$	180	38	1.48 (0.895)	$p = 0.117$	19	0	−0.045 (0.701)	$p = 0.005$
Tap	Baseline	125	15	1.19 (0.84)	$Z = -1.119$	72	23.5	1.12 (0.87)	$Z = -0.520$	12	1	−0.30 (0.0)	$Z = -2.201$
	Endline	129	19	1.08 (0.84)	$p = 0.263$	71	22	1.31 (0.85)	$p = 0.603$	12	0	0.34 (0.73)	$p = 0.028$
Reservoir tank	Baseline	50	12.5	1.16 (0.88)	$Z = -0.189$	36	8.5	0.89 (0.89)	$Z = -0.956$				
	Endline	45	16	1.17 (0.82)	$p = 0.850$	36	19.5	1.24 (0.75)	$p = 0.339$				

The Chlorination group refers to samples with detectable chlorine.

endline the rating of the system's functionality in the treatment group was significantly better than in the control group ($Z = -3.777$, $p < 0.001$), whereas at baseline there was no significant difference ($Z = -0.815$, $p = 0.415$).

In the control systems, significantly more households reported interruptions of water services for more than a week at endline compared to baseline, whereas no such change was observed among treatment systems (**Table 3**). The percentage of households that reported interruptions at endline did not differ significantly between treatment and control systems ($Z = -1.264$, $p = 0.206$). Yet, there were more reported interruptions in the treatment systems at baseline than in the control systems ($Z = -2.089$, $p = 0.037$). The most commonly mentioned reasons for interruptions at endline were broken or burst pipes (49.3%), lack of water (27.8%), intake problems (19.0%) and landslides (17.7%). Households' confidence that repairs to the system could be performed within 1 week decreased from baseline to endline among both treatment and control systems (**Table 3**). At endline, the confidence that problems would be fixed within 1 week was significantly higher in the treatment than in the control group ($Z = -2.122$, $p = 0.034$). Whereas, at baseline, they did not differ significantly ($Z = -1.345$, $p = 0.179$).

At endline we asked the interviewees how confident they were that their water system would be functional 1 year later using a 5-dot scale, with the largest dot being the highest confidence level (see **Supplementary Figure S1** for the 5-dot scale). Households served by treatment systems were significantly more confident (M = 3.8, SD = 1.0) that the system would be functional in 1 year than people in the control systems (M = 3.4, SD = 1.1) ($Z = -3.897$, $p < 0.001$). There was no significant difference in the confidence that the system will be functional 1 year later between treatment and control schemes at baseline ($Z = -1.222$, $p = 0.222$).

Water Availability

The endline survey also probed the duration of water availability at the tap. The distribution of the hours of water availability per day reported at endline is bimodal with around one third of respondents (38%) having water services available for 12 h a day or less, which we defined as intermittent supply, and two thirds (62%) for more than 12 h, which was defined as continuous water supply. The cut-off at 12 h was linked to the JMP definition of water availability when needed (WHO/UNICEF Joint Monitoring Programme, 2019). Most of the households with more than 12 h water supply had water for a full 24 h (61.2%). Among the control systems, more households (69.8%) had continuous water supply than in the treatment systems (57.3%) ($Z = -2.748$, $p = 0.006$). In the control systems, water services were available for an average of 18 h daily (SD = 9.2), which were significantly more hours than in the treatment systems (M = 15.8 h, SD = 9.6) ($Z = -2.145$, $p = 0.032$). The hours of water availability showed a small and marginally significant negative correlation with the *E. coli* concentration at the point of consumption ($r_s = -0.087$, $p = 0.055$). Hence, having less hours of water availability at the tap was associated with higher *E. coli* concentration.

TABLE 3 | System functionality as measured by the percentage of households that reported the corresponding answer in treatment and control systems at baseline and endline, including comparison from baseline to endline with the Wilcoxon signed-rank test.

	Treatment			**Control**		
	Baseline	**Endline**	**Wilcoxon signed-rank test**	**Baseline**	**Endline**	**Wilcoxon signed-rank test**
Water system functioning						
Yes, well	79.2%	92.9%	$Z = -4.517$ $p < 0.001$	82.2%	81.6%	$Z = 0.000$ $p = 1.000$
Yes, but not so well	18.5%	6.5%		16.1%	17.3%	
No	2.2%	0.6%		1.7%	1.1%	
Interruption: Households with interruptions that lasted longer than 1 week within the last 6 months	18.2%	14.6%	$Z = -1.222$ $p = 0.222$	11.1%	19.0%	$Z = -2.271$ $p = 0.023$
Problem fixed within 1 week						
Very confident	68.7%	43.3%	$Z = -4.802$ $p < 0.001$	62.7%	33.7%	$Z = -4.700$ $p < 0.001$
Somewhat confident	22.9%	43.6%		28.8%	50.9%	
Not at all confident	8.4%	13.1%		8.5%	15.4%	

The sample size was in treatment system at baseline N = 313, at endline N = 308, and in control systems at baseline N = 180, at endline N = 179.

TABLE 4 | Percentage of households that reported practicing household water treatment in the treatment and control systems at baseline and endline, including comparison from baseline to endline with the Wilcoxon signed-rank test and at baseline and endline with the Mann-Whitney test.

Practicing household water treatment	**Baseline**	**Endline**	**Wilcoxon signed-rank test**
Treatment	52.1%	65.9%	$Z = -4.364$ $p < 0.001$
Control	48.9%	35.6%	$Z = -3.159$ $p = 0.002$
Mann-Whitney test	$Z = -0.681$ $p = 0.496$	$Z = -6.440$ $p < 0.001$	

The sample size was in treatment system at baseline N = 313, at endline N = 308, and in control systems at baseline N = 180, at endline N = 179.

Users' Experience

Satisfaction, Taste, and Perceived Safety

According to the interviews at endline, people served by the treatment systems were significantly more satisfied with their main drinking water source than households served by the control systems ($Z = -2.255$, $p = 0.024$). Further, in the treatment systems, significantly more respondents indicated that the source is available when needed than in control systems ($Z = -3.568$, $p < 0.001$). At endline the tests revealed no significant difference in the perception of taste between control and treatment ($Z = -1.470$, $p = 0.142$), nor whether or not chlorine was detected in the household storage containers ($Z = -1.559$, $p = 0.119$). Better functioning sources were associated with the availability of the source when needed, and both were associated with higher users' satisfaction (see text in **Supplemental Material** and **Supplementary Table S6** for further details on the relationship between users' satisfaction and system functionality metrics).

There was no significant difference in households' perception of water safety at baseline among treatment and control systems ($Z = -1.529$, $p = 0.126$). When looking at the evolution of the perception of safety from baseline to endline, the change was not significant in the control ($Z = -1.262$, $p = 0.207$) nor in the treatment systems ($Z = -0.653$, $p = 0.513$). However, contrary to expectations, at endline households in control systems perceived their water as significantly safer than households in treatment systems ($Z = -2.361$, $p = 0.018$). The perception of water safety was significantly correlated with people treating their water at household level ($r_s = 0.184$, $p < 0.001$). The riskier people perceived the water to be, the higher the likelihood that they would treat it.

Household Water Treatment and Hand Hygiene

While the percentage of households treating their water increased significantly in the treatment systems from baseline to endline, it decreased in the control systems (**Table 4**). At endline, significantly more people reported treating their water at household level within the treatment systems compared to the control systems. Whereas, at baseline, there was no significant difference. 93% of the households treating their water used a filter and 9% boiled their water. 70% of the people treating their water indicated to do this always, 18% most of the times, 9% rarely and 3% never.

Microbial water quality was correlated with indicators for safe storage. The presence of a lid on the storage container correlated negatively with the *E. coli* concentration at the point of consumption ($r_s = -0.103$, $p = 0.023$). Further, the cleanliness of the storage container correlated negatively with the *E. coli* concentration ($r_s = -0.177$, $p = 0.035$). Meaning that the presence of a lid and the cleanliness of the storage container were both associated with lower *E. coli* concentrations (more details about water storage in **Supplementary Table S7**).

We observed at endline that hand hygiene practices were also improved by the interventions. Besides more households having access to hand washing facilities in the treatment systems at endline (82.7% compared to 70.0% in control systems), the state of the handwashing facilities was better (87.0% of the handwashing facilities being in good and clean condition) than in the control systems (68.8%) ($Z = -4.243$, $p < 0.001$). Additionally, more households served by the treatment systems (80.3%) had soap available as compared to the control systems, where less than half of the households (48.0%) had soap at endline ($Z = -6.425$, $p < 0.001$). Furthermore, the number of times respondents in the treatment group washed their hands daily (Median = 7) was significantly greater than respondents in the control group (Median = 6) ($Z = -2.097$, $p = 0.036$).

Identifying the Factors Explaining Fecal Contamination of Drinking Water in the Household: Regression Analysis

An ordinal logistic regression model was estimated in order to identify factors associated with fecal contamination of stored drinking water. The dependent variable was the *E. coli* concentration at the point of consumption at endline, split into ordered risk categories, taking the value of 0 for samples that were <1 *E. coli* CFU/100 mL (non-detect), 1 for 1–10 *E. coli* CFU/100 mL, 2 for 11–100 *E. coli* CFU/100 mL and 3 for >100 *E. coli* CFU/100 mL. Bivariate comparisons (Spearman's correlation for continuous data and Mann-Whitney test in case of binary data) between the dependent variable and its hypothesized explanatory variables were made to identify significant associations and inform the construction of a parsimonious model (bivariate test results shown in **Supplementary Table S8**). The reduced model is presented in **Table 5** and the full model shown in **Supplementary Table S9**. The assumption of no multicollinearity was satisfied. To check for the assumption for an ordinal logistic regression of proportional odds, we conducted the test of parallel lines. With the test being non-significant [$\chi^2(38) = 44.065$, $p = 0.230$], the assumption of proportional odds was satisfied. The likelihood ratio chi-square test revealed a significant improvement in fit of the final model relative to the intercept only model [$\chi^2(19) = 81.719$, $p < 0.001$]. In the Goodness-of-Fit analysis, the Pearson chi-square test [$\chi^2(1412) = 1448.68$, $p = 0.243$] and the deviance test [$\chi^2(1412) = 1200.52$, $p = 1.000$] were both non-significant, which suggested a good model fit.

All else constant, belonging to the treatment scheme, measured chlorine concentration and the presence of a lid on the storage container were significant predictors and positively affected the microbial safety of drinking water at the point of consumption. The odds ratio (OR) indicated that treatment systems were associated with decreased odds in *E. coli* risk category. Meaning that all else constant, for households receiving the treatment intervention package, the odds of having unsafe water are multiplied by 0.538 as compared to those receiving no intervention ($p = 0.002$). For those with detectable chlorine, the odds of having unsafe water are greatly decreased, i.e., multiplied by 0.030 as compared to those without chlorine ($p = 0.003$). Safe storage, as measured by the presence of a lid on the storage container, was associated with decreased odds in *E. coli* risk category, although the association was only marginally significant in the model ($p = 0.055$). All else constant, for households having safe water storage, the odds of having unsafe water are multiplied by 0.641 as compared to those without safe storage.

TABLE 5 | Ordinal logistic regression model of the factors explaining risk of fecal contamination of stored drinking water (0 for <1 *E. coli* CFU/100 mL, 1 for 1–10 *E. coli* CFU/100 mL, 2 for 11–100 *E. coli* CFU/100 mL, 3 for >100 *E. coli* CFU/100 mL) at endline, with the odds ratio (OR), Wald statistic, and the level of significance reported.

Variable	OR	Wald	Sig.
Treatment scheme (yes = 1, no = 0)	0.538	9.930	**0.002**
Measured chlorine concentration (mg/L)	0.030	8.782	**0.003**
WASH information received (yes = 1, no = 0)	1.351	2.172	0.141
Practicing household water treatment (yes = 1, no = 0)	1.317	1.439	0.230
Presence of lid on storage container (yes = 1, no = 0)	0.641	3.696	0.055
Confidence of functionality in 1 year (5 = very confident – 1 = not confident at all)	0.755	1.285	0.257
Confidence of problem fixed in 1 week (3 = very confident – 1 = not confident at all)	1.080	0.174	0.676
Handwashing facilities with soap available (yes = 1, no = 0)	0.866	0.431	0.512
Hours of water availability (hours)	0.991	0.872	0.350
People in household (number)	1.044	1.289	0.256
Presence of children under 5 years (yes = 1, no = 0)	1.171	0.733	0.392
Monthly expenditure (USD)	1.209	0.218	0.641
Respondent completed primary or higher education (yes = 1, no = 0)	0.999	1.138	0.286
Toilet clean (yes = 1, no = 0)	0.971	0.020	0.888
Animal in same house overnight (yes = 1, no = 0)	1.332	1.326	0.250
District: Achham	1.234	0.563	0.453
District: Jajarkot	0.526	3.915	**0.048**
District: Kalikot	0.672	1.602	0.206
District: Surkhet	0.414	7.126	**0.008**
Quasi-R^2 (Nagelkerke) = 0.17			
$N = 478$ (missing: 15)			

p-values <0.05 are shown in bold.

Many of the project-associated independent variables were non-significant predictors of *E. coli* contamination risk in this model, such as the reception of WASH information ($p = 0.141$), the use of household water treatment ($p = 0.230$), the confidence that the water system will be functional in one year ($p = 0.257$), the confidence that in case of a problem it will be fixed within 1 week ($p = 0.676$), the presence of handwashing facilities with soap ($p = 0.512$), and the hours of water availability ($p = 0.350$).

Contrary to expectations, the majority of the confounders (i.e., contextual factors and non-project-related factors) were not significantly associated with *E. coli* contamination risk. Solely, being in Surkhet (OR = 0.414, $p = 0.008$) or in Jajarkot (OR = 0.526, $p = 0.048$) were associated with decreased odds in

E. coli risk category, everything else held constant. Meaning that the odds of having unsafe water are multiplied by 0.414 for households situated in Surkhet and multiplied by 0.526 for households situated in Jajarkot as compared to those among the other districts.

DISCUSSION

This research provides new evidence of the effectiveness of a combined risk-based water safety intervention for piped supplies in rural Nepal. It builds on and expands findings from prior studies by incorporating a broader geographic region to capture increased variability in households' and water systems' characteristics as compared with a previous study at the same field site (Tosi Robinson et al., 2018). The novel aspect of this work is the expanded set of outcome measures, in particular on users' perceptions and system functionality. Furthermore, the present study included a longer study period to improve understanding of intervention longevity and quarterly tracking of performance indicators using basic sanitary inspections to assess temporal changes. We followed control communities alongside the intervention communities to account for potential variation in study outcomes due to external factors.

Microbial Quality

Our study reveals that chlorination is the only component of the combined intervention package that leads to a major improvement in the microbial water quality of collection taps and household stored water. The *E. coli* concentration in stored water containers decreased significantly from baseline to endline solely among households in treatment communities where free residual chlorine was detected at the point of consumption. The effectiveness of other interventions to provide safe water was relatively lower, possibly due to seasonal effects of the monsoon causing higher contamination in general. For example, among all treatment systems, the share of household storage containers meeting WHO guidelines for microbial safety (<1 CFU/100 mL) increased modestly from 11% at baseline to 20% at endline. In the control systems, by contrast, we observed an increasing trend in the *E. coli* contamination of drinking water supplies over time. It's notable that during the rainy season fecal contamination can be mobilized and therefore cause degradation of the microbial water quality of surface and groundwater (Johnson et al., 2010). Therefore, it is plausible that the broader intervention package, while not dramatically improving water quality from baseline levels if chlorination was absent, still prevented increased contamination due to the seasonal monsoon. However, water safety interventions should ensure year-round availability of safe water.

Collection taps with measurable chlorine concentration had no detectable *E. coli*. Despite chlorination successfully improving microbial water quality at the tap, a general trend among all systems of water quality deteriorating from the tap to the point of consumption was observed. This deterioration is in line with previous studies (Kumpel and Nelson, 2013; Shields et al., 2015), which conclude that unsafe storage and handling practices can lead to recontamination of the drinking water (Shields et al., 2015).

Finally, we found a negative correlation between the hours of water availability and the *E. coli* concentration. Hence, more hours of water availability were associated with lower *E. coli* concentration. Previous studies have shown that intermittent water supply systems have higher *E. coli* concentrations, with greater frequency, at the tap compared to continuous water supply (Kumpel and Nelson, 2013). Although water storage can be reduced through continuous water supply, households served by continuous systems in our study site still did not avoid all storage, which is in line with previous studies (Kumpel and Nelson, 2013). Consequently, a continuous water supply system does not guarantee prevention of recontamination due to unsafe storage. This further supports the use of chlorination as it reduces recontamination.

Chlorination

Chlorination is only effective when applied in appropriate doses (WHO, 2017b). The number of systems that manual chlorination was successfully implemented within was lower than planned by the research team. This underlines that proper implementation of chlorination in remote rural settings includes various technical and capacity challenges. Besides the substantial time burden of manual chlorination, correct dosing in order to reach the appropriate chlorine concentration at the point of consumption, with minimal disinfection by-products, requires technical skills and regular verification monitoring. The lack of proper equipment, insufficient chlorine supply, and limited operator training might have been additional hindering factors for the uptake of regular manual chlorination at system scale among the systems included in this study. To amend this capacity shortage, we recommend closer or more continuous supervision of village maintenance workers, including regular training opportunities (Ferrero et al., 2019) and strengthening of local laboratories for operational monitoring (Diener et al., 2017; Peletz et al., 2018).

This study has not confirmed previous research on people refusing chlorination as a water treatment method due to the taste (Crider et al., 2018). Taste and odor have been shown to be key factors influencing the acceptance of drinking water (Francis et al., 2015). In contrast to the results of similar studies carried out in this area (Crider et al., 2018), the perception of taste of the drinking water did not differ significantly between the systems with and without chlorination. However, according to the WHO drinking water quality guidelines, people are able to taste or smell chlorine at levels as low as 0.3 mg/L (WHO, 2017b) and in other studies people detected the taste at levels of 0.7 mg/L (Crider et al., 2018). In our measurements, we did not have many samples above these thresholds, potentially explaining the general acceptance of chlorinated water among households receiving this intervention.

Users' Perception

Our results indicate that the study interventions had a considerable and counterintuitive influence on users' perceptions of the safety of their drinking water. Water in the control systems was perceived as safer than in the treatment systems, despite measured fecal contamination levels indicating the opposite. This might be due to the fact that among the treatment systems people received more information on water treatment and hygiene, which potentially increased the awareness of risks linked to water contamination. In addition, regular monitoring among the treatment systems may have had an impact on the perception of water safety. According to Trent et al. (2018), providing a household with specific information about drinking water quality can cause behavior change and improve microbial drinking water quality (Trent et al., 2018).

Our study also reveals a positive correlation between the perception of water safety and the use of household water treatment technologies. As mentioned above, greater awareness the water contamination could have been a reason for people perceiving the water as more potentially harmful. This might have led them to treat their water in their household. The aforementioned is in line with previous studies, which showed that households' perception of their own water quality being low, and therefore presenting a health risk, is the most important precursor for the adoption of household water treatment (Daniel et al., 2018). However, household water treatment does not necessarily improve water quality, due to inadequate handling practice or the treatment products being of bad quality (Meierhofer et al., 2018). Therefore, interventions may have caused an increase in the perception of water safety associated with the use of household water treatment, which did not necessarily translate into a positive impact on water quality.

The interventions were correlated with higher user satisfaction with the water system and the water source being more available when needed. However, this stands in contrast to the reported hours of water availability, which was higher in the control systems than in the treatment systems. But, as the hours of water availability were not assessed at baseline, we do not know how many hours of water availability they had prior to the interventions. It is possible that the interventions may have led operators to be more prudent about opening hours, resulting in lower but more targeted availability. At the same time, there were less reported interruptions in the treatment systems. This is another possible explanation for the higher satisfaction and availability when needed among the treatment systems.

System Functionality

This study reveals that the interventions presumably increased households' rated current functionality of the system and their confidence that the system will be functional in 1 year. Apart from the sanitary inspections indicating more dysfunctional taps in the control systems, the systems with interventions showed a better rated functionality of the water system. Furthermore, we observed an increase in people reporting that the source was functioning well from baseline to endline among the treatment systems, compared with no change in the control systems. Higher water availability due to the monsoon could explain favorable effects on the rated functionality and satisfaction during this season. At the same time, increased surface runoff due to the monsoon can mobilize organic and inorganic matter and therefore cause turbidity (Brasington and Richards, 2000). The turbidity potentially affected the rated system functionality and satisfaction negatively. However, in the treatment systems, roughing sand filters were installed as part of the interventions, which can effectively reduce turbidity throughout the distribution system (Hashimoto et al., 2019). Hence, they could have caused the water to be less turbid during monsoon compared to other years and therefore be a reason for higher satisfaction and better rated functionality in the treatment systems.

According to the results, the combined intervention potentially prevented interruptions. The reported interruptions that lasted longer than 1 week in the control systems were higher at endline than at baseline. However, the proportion of people reporting such interruptions did not change significantly from baseline to endline within the treatment systems. Hence, the interventions either prevented interruptions from occurring or any interruptions could be fixed faster among the treatment systems. At the same time, among all systems we observed a decrease from baseline to endline in users' confidence that a problem could be fixed within 1 week. It is difficult to interpret a clear reason for this finding. Potentially the monsoon caused problems (e.g., landslides or impassable roads) that were insurmountable within a short time period. In summary, results indicate that the water safety interventions reduced the number of interruptions, potentially due to infrastructure improvements and better trained village maintenance workers.

Limitations

It is plausible that a number of aspects of the study design may limit our interpretation of the results obtained. First, the baseline and endline data collection were not conducted during the same season. The baseline data collection took place during the dry season and the endline after monsoon. Hence, seasonal effects likely influenced study outcomes across both treatment and control groups. Rainfall prior to endline data collection increased water availability, which could have caused increased levels of contamination due to mobilized fecal contaminants. Thus, seasonality could be responsible for differences not only in water quality but also in users' perception and system functionality (Kumpel et al., 2017).

We note also that correct application of system-chlorination was scarce at our study site. At endline, there were only two systems with measurable chlorine concentration at the point of consumption. Future research at this and similar field sites should aim to include more systems with correctly applied chlorination at the system level.

E. coli is sensitive and accordingly has some limitations as an indicator bacterium. They might not survive as long as certain pathogens, such as cryptosporidium, especially after exposure to chlorine (WHO, 2017b). The results reported here may therefore overstate the actual risk reductions of drinking chlorinated water. Furthermore, the data for many *E. coli* samples were censored since their concentrations were at the upper and lower detection limits. The inclusion of discrete values in continuously distributed microbial data may bias the results of statistical analyses (Chik et al., 2018).

CONCLUSION

The evidence from this study implies that chlorination led to a dramatic decrease of fecal contamination in drinking water at the point of consumption, while no other intervention was as effective. However, the number of systems across the study site where chlorination was correctly applied was limited. As compared to chlorination, the effectiveness of other interventions was relatively low, which may have been due to higher contamination brought about by the monsoon. Future research should prioritize interventions that effectively counteract the seasonal effects of the monsoon in Nepal to achieve consistent safety and availability of safe water. Additionally, the interventions increased households' awareness of water contamination issues, which in turn appeared to motivate the uptake of household water treatment methods. In addition, the interventions achieved higher general satisfaction and better rated functionality among the households enrolled in this study. As our research highlights the effectiveness of chlorination in ensuring acceptable drinking water quality, for future studies we would encourage more broadly applying chlorination methods at system level.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material** and water quality data are available at https://opendata.eawag.ch/. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The research activities involving human participants were reviewed and approved by EAWAG Ethical Review Committee. All participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CB, AS, BK, MB, and SM: conceived and designed the research and contributing to the writing of the article. CB, AS, BK, and MB: conducted the field research. CB and SM: analyzed the results. CB: wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Swiss Agency for Development Cooperation (Contract number 81053992) and this document is an output from the REACH programme funded by UK Aid from the UK Foreign, Commonwealth and Development Office (FCDO) for the benefit of developing countries (Programme Code 201880). However, the views expressed, and information contained in it are not necessarily those of or endorsed by those agencies, which can accept no responsibility for such views or information or for any reliance placed on them.

ACKNOWLEDGMENTS

We particularly acknowledge Rubika Shrestha, Ram Bahadur Shrestha, Mohan Bhatta, Binod Pandey, Kushmakar Pant, Ramesh Chaudhary, Prakash Ayer, Neerima Thapa, Netra Thapa, Ramesh Yogi, Meghraj Pandey, Ujir Singh Nepali, Ganesh Bista, Mahendra Adhikari, Lalita Bhandari, Lok Bohora, Satya Shahi, Chandra Dangi, Mukesh Sharma, Guillaume Clair-Caliot, and Benjamin Ambuehl for significant support during the implementation and assessment of the study interventions. Vasco Lepori, Thomas Lauber, and Urs Bänziger supported the writing process and reviewed the first draft of the manuscript. Timothy Julian provided valuable feedback on data collection, analysis and interpretation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frwa.2021.750802/full#supplementary-material

REFERENCES

Bain, R., Cronk, R., Hossain, R., Bonjour, S., Onda, K., Wright, J., et al. (2014a). Global assessment of exposure to faecal contamination through drinking water based on a systematic review. *Trop. Med. Int. Health* 19, 917–927. doi: 10.1111/tmi.12334

Bain, R., Cronk, R., Wright, J., Yang, H., Slaymaker, T., and Bartram, J. (2014b). Fecal contamination of drinking-water in low- and middle-income countries: a systematic review and meta-analysis. *PLoS Med.* 11:e1001644. doi: 10.1371/journal.pmed.1001644

Barrington, D., Fuller, K., and McMillan, A. (2013). Water safety planning: adapting the existing approach to community-managed systems in rural Nepal. *J. Water Sanit. Hyg. Dev.* 3, 392–401. doi: 10.2166/washdev.2013.120

Baum, R., Kayser, G., Stauber, C., and Sobsey, M. (2014). Assessing the microbial quality of improved drinking water sources: results from the Dominican Republic. *Am. J. Trop. Med. Hyg.* 90, 121–123. doi: 10.4269/ajtmh.13-0380

Bista, M. (2001). Infectious diseases in Nepal: a collection of selected publications on communicable diseases including vector-borne diseases 1992-2000. *J. Inst. Med. Nepal* 12, 247–257.

Brasington, J., and Richards, K. (2000). Turbidity and suspended sediment dynamics in small catchments in the Nepal Middle Hills. *Hydrol. Process.* 14, 2559–2574. doi: 10.1002/1099-1085(20001015)14:14<2559::AID-HYP114>3.0.CO;2-E

Chik, A. H. S., Schmidt, P. J., and Emelko, M. B. (2018). Learning something from nothing: the critical importance of rethinking microbial non-detects. *Front. Microbiol.* 9:2304. doi: 10.3389/fmicb.2018.02304

Crider, Y., Sultana, S., Unicomb, L., Davis, J., Luby, S. P., and Pickering, A. J. (2018). Can you taste it? taste detection and acceptability thresholds for chlorine residual in drinking water in Dhaka, Bangladesh.

Sci. Total Environ. 613–614, 840–846. doi: 10.1016/j.scitotenv.2017.09.135

Daniel, D., Marks, S. J., Pande, S., and Rietveld, L. (2018). Socio-environmental drivers of sustainable adoption of household water treatment in developing countries. *Npj Clean Water* 1, 1–6. doi: 10.1038/s41545-018-0012-z

Diener, A., Schertenleib, A., Daniel, D., Kenea, M., Pratama, I., Bhatta, M., et al. (2017). "Adaptable drinking-water laboratory unit for decentralised testing in remote and alpine regions" in *40th WEDC International Conference* (Loughborough), 1–6.

Ferrero, G., Setty, K., Rickert, B., George, S., Rinehold, A., DeFrance, J., et al. (2019). Capacity building and training approaches for water safety plans: a comprehensive literature review. *Int. J. Hyg. Environ. Health* 222, 615–627. doi: 10.1016/j.ijheh.2019.01.011

Francis, M. R., Nagarajan, G., Sarkar, R., Mohan, V. R., Kang, G., and Balraj,. V. (2015). Perception of drinking water safety and factors influencing acceptance and sustainability of a water quality intervention in rural southern India. *BMC Public Health* 15:731. doi: 10.1186/s12889-015-1974-0

Government of Nepal (2017). *Nepal–Demographic and Health Survey 2016*. Kathmandu: Ministry of Health and Population.

Government of Nepal (2018). *Annual Report - Department of Health Services 2074/75 (2017/18), Vol. 75*. Kathmandu: Ministry of Health and Population.

Hashimoto, Y., Takashima, H., and Jayamohan, S. (2019). Application of roughing filter to pre-treat 1,000 NTU raw water for slow sand filter. *Water Pract. Technol.* 14, 355–364. doi: 10.2166/wpt.2019.021

Johnson, K. M., Kumar, M. R., and Ponmurugan, P. (2010). Degradation of the quality of water during monsoon and the related outbreak of waterborne diseases. *Ecol. Environ. Conserv.* 16, 277–280.

Kumpel, E., Cock-Esteb, A., Duret, M., de Waal, D., and Khush, R. (2017). Seasonal variation in drinking and domestic water sources and quality in Port Harcourt, Nigeria. *Am. J. Trop. Med. Hyg.* 96, 437–445. doi: 10.4269/ajtmh.16-0175

Kumpel, E., Delaire, C., Peletz, R., Kisiangani, J., Rinehold, A., De France, J., et al. (2018). Measuring the impacts of water safety plans in the Asia-Pacific Region. *Int. J. Environ. Res. Public Health* 15:1223. doi: 10.3390/ijerph15061223

Kumpel, E., and Nelson, K. L. (2013). Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Res.* 47, 5176–5188. doi: 10.1016/j.watres.2013.05.058

Mahmud, S. G., Shamsuddin, S. A. J., Ahmed, M. F., Davison, A., Deere, D., and Howard, G. (2007). Development and implementation of water safety plans for small water supplies in Bangladesh: benefits and lessons learned. *J. Water Health* 5, 585–597. doi: 10.2166/wh.2007.045

Marks, S. J., and Shrestha, R. (2020). "Improving drinking water quality in rural communities in Mid-Western Nepal," in *Women in Water Quality, Women in Engineering and Science*, ed D. O'Bannon (Cham: Springer), 47–59. Available online at: https://link.springer.com/chapter/10.1007%2F978-3-030-17819-2_3

Meierhofer, R., Bänziger, C., Deppeler, S., Kunwar, B. M., and Bhatta, M. (2018). From water source to tap of ceramic filters—factors that influence water quality between collection and consumption in rural households in Nepal. *Int. J. Environ. Res. Public Health* 15:2439. doi: 10.3390/ijerph15112439

Nissui (2009). *Compact Dry "Nissui" EC for Coliform and E. coli, User's Manual. 1–4*. Tokyo: Nissui Pharmaceuticals Co. Ltd.

Peletz, R., Kisiangani, J., Bonham, M., Ronoh, P., Delaire, C., Kumpel, E., et al. (2018). Why do water quality monitoring programs succeed or fail? a qualitative comparative analysis of regulated testing systems in sub-Saharan Africa. *Int. J. Hyg. Environ. Health* 221, 907–920. doi: 10.1016/j.ijheh.2018.05.010

Schertenleib, A., Sigrist, J., Friedrich, M. N. D., Ebi, C., Hammes, F., and Marks, S. J. (2019). Construction of a low-cost mobile incubator for field and laboratory use. *J. Vis. Exp.* 145:e58443. doi: 10.3791/58443

Shaheed, A., Orgill, J., Montgomery, M. A., Jeuland, M. A., and Brown, J. (2014). Why "improved" water sources are not always safe. *Bull. World Health Organ.* 92, 283–289. doi: 10.2471/BLT.13.119594

Shields, K. F., Bain, R. E., Cronk, R., Wright, J. A., and Bartram, J. (2015). Association of supply type with fecal contamination of source water and household stored drinking water in developing countries: a bivariate meta-analysis. *Environ. Health Perspect.* 123, 1222–1231. doi: 10.1289/ehp.1409002

String, G. M., Singleton, R. I., Mirindi, P. N., and Lantagne, D. S. (2020). Operational research on rural, community-managed Water Safety Plans: case study results from implementations in India, DRC, Fiji, and Vanuatu. *Water Res.* 170:115288. doi: 10.1016/j.watres.2019.115288

Tosi Robinson, D., Schertenleib, S., Kunwar, B., Shrestha, R., Bhatta, M., and Marks, S. J. (2018). Assessing the impact of a risk-based intervention on piped water Quality in rural communities: The case of Mid-Western Nepal. *Int. J. Environ. Res. Public Health* 15:1616. doi: 10.3390/ijerph15081616

Trent, M., Dreibelbis, R., Bir, A., Tripathi, S. N., Labhasetwar, P., Nagarnaik, P., et al. (2018). Access to household water quality information leads to safer water: a cluster randomized controlled trial in India. *Environ. Sci. Technol.* 52, 5319–5329. doi: 10.1021/acs.est.8b00035

UNDP (2019). *Human Development Report 2019 Inequalities in Human Development in the 21 st Century Nepal*. New York, NY: UNDP, 1–10.

WHO (2017a). *Global Status Report on Water Safety Plans: A Review of Proactive Risk Assessment and Risk Management Practices to Ensure the Safety of Drinking-water*. Geneva: World Health Organization.

WHO (2017b). *Guidelines for Drinking-Water Quality: Fourth Edition Incorporating the First Addendum*. Geneva: World Health Organization.

WHO/UNICEF Joint Monitoring Programme (2019). *Progress on Household Drinking Water, Sanitation and Hygiene 2000-2017*. Geneva: World Health Organization and UNICEF.

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.

frontiers | Frontiers in Water

TYPE Original Research
PUBLISHED 15 July 2022
DOI 10.3389/frwa.2022.786040

OPEN ACCESS

EDITED BY
Samrat Chatterjee,
Pacific Northwest National Laboratory (DOE), United States

REVIEWED BY
Hiba Baroud,
Vanderbilt University, United States
Siddharth Saksena,
Virginia Tech, United States

*CORRESPONDENCE
S. Samadi
samadi@clemson.edu

SPECIALTY SECTION
This article was submitted to Water and Built Environment, a section of the journal Frontiers in Water

RECEIVED 29 September 2021
ACCEPTED 27 June 2022
PUBLISHED 15 July 2022

CITATION
Samadi S (2022) The convergence of AI, IoT, and big data for advancing flood analytics research. *Front. Water* 4:786040. doi: 10.3389/frwa.2022.786040

The convergence of AI, IoT, and big data for advancing flood analytics research

S. Samadi*

Department of Agricultural Sciences, Clemson University, Clemson, SC, United States

Floods are among the most destructive natural hazards that affect millions of people across the world leading to severe loss of life and damage to properties, critical infrastructure, and the environment. The combination of artificial intelligence (AI), big data, and the Internet of Things (IoTs), has the potential to more accurately predict these extreme events and accelerate the convergence of advanced techniques for flood analytics research. This convergence—so called the Artificial Intelligence of Things (AIoT)—is transformational for both technologies and science-based decision making since AI adds value to IoT through interpretable machine learning (ML) while IoT leverages the power of AI *via* connectivity and data intelligence. The aim of this research is to discuss the workflow of a Flood Analytics Information System (FAIS; version 4.00) as an example of AIoT prototype to advance and drive the next generation of flood informatics systems. FAIS integrates crowd intelligence, ML, and natural language processing (NLP) to provide flood warning with the aim of improving flood situational awareness and risk assessments. Various image processing algorithms, i.e., Convolutional Neural Networks (CNNs), were also integrated with the FAIS prototype for image label detection, and floodwater level and inundation areas calculation. The prototype successfully identifies a dynamic set of at-risk locations/communities using the USGS river gauge height readings and geotagged tweets intersected with watershed boundary. The list of prioritized locations can be updated, as the river monitoring system and condition change over time (typically every 15 min). The prototype also performs flood frequency analysis (FFA) by fitting multiple probability distributions to the annual flood peak rates and calculates the uncertainty associated with the model. FAIS was operationally tested (beta-tested) during multiple hurricane driven floods in the US and was recently released as a national-scale flood data analytics pipeline.

KEYWORDS

artificial intelligence of things, flood analytics information system, at-risk locations, flood image processing, convolutional neural networks

Introduction

The combination of Internet of Things (IoTs), the big data it creates, and the ability to use it *via* Artificial Intelligence (AI) is at a tipping point where significant changes, transformation, and innovation are poised to take place (e.g., Samadi and Pally, 2021). This convergence-so called Artificial Intelligence of Things (AIoT)—has made possible the development of many intelligent and context-awareness systems which are able to efficiently monitor and model environmental systems and provide timely and more accurate forecasts and decisions. The application of AIoT empowers a systematic study of domain-inspired architectures to answer fundamental questions and enable data-driven discovery and problem solving (Donratanapat et al., 2020). Indeed, AIoT creates intelligent connected systems to collect and transmit data from multiple sources–supporting the "learning" process involved in training AI to assist with decision making. This leads to interoperable networks and systems that are becoming increasingly more capable of solving real-time forecasting problems across scales.

AIoT also improves the contextually aware decision-making process for resolving complex flood operational decisions enabling the use of machine intelligence to be implemented more efficiently in real-time. In flood studies, AIoT can provide insights from both big data and AI to further enhance real-time forecasts and improve the capabilities of early warning systems. However, AIoT brings with it the burden of developing custom interfaces and handling huge amounts of structured and unstructured data. In flood analytics research, modeling surface runoff requires linking attributes across river networks, to hierarchically nested sub-basins and river reaches at multiple scales. Further, the data characteristics of flood modeling problems have also grown from static to dynamic and spatio-temporal, and centralized to distributed, and grow in both scope and size. The effective integration of big data for flood forecasting decision-making also requires the advances of data mining and analytics systems.

To tackle these challenges, many data providers such as the US Geological Survey (USGS) and the National Weather Service (NWS) focused on building application programming interfaces (APIs) to ease data management strategies and deliver data services *via* a network connection. Providing information and services through Web APIs supports interoperability and openness, and eases access to the data. The ability of APIs to stream the data and extract valuable attributes for real-time decision making added an entirely new dimension to flood forecasting research. When embedded within an AIoT system, API ensures interoperability between all connected components to optimize processes and extract valuable insights from big data provided by the connected IoT devices.

In the past few years, the use of APIs for extracting flood images and data produced *via* social media and river cameras in the context of flood crises has opened a new avenue and opportunities for developing new AIoT systems. The potential of AIoT for providing connectivity between IoTs and big data along with high-speed data transfer capabilities can be used to implement real time image processing and machine learning (ML) systems for flood studies. Both image processing and ML algorithms are powerful methods of image data exploration that can be embedded with river cameras for monitoring flood conditions (Donratanapat et al., 2020). Recently, image processing and ML algorithms have been used to label time lapse camera imagery, crowdsourcing and tabular data, and user generated texts and photos to extract road flooding condition and inundation extend (De Albuquerque et al., 2015; Starkey et al., 2017; Feng and Sester, 2018). In addition, ML algorithms such as natural language processing (NLP) have been used to extract and analyze social media geodata and provide actionable intelligence for real-time flood assessment. In many flood situations, big data from other sources such as USGS, NWS, National Hurricane Center (NHC), in situ sensors, existing authoritative geographic data, etc. are also available which can profitably be leveraged upon in order to make the real-time flood situational assessment more efficient and successful.

Despite the importance of using AIoT in flood related research, few studies have focused on its application in flood monitoring and modeling approaches. For example, Demir et al. (2018) investigated a web-based platform- the Iowa Flood Information System (IFIS)- as the next-generation decision support systems for flood studies. IFIS provides real-time information on streams and weather conditions that incorporates advanced rainfall-runoff models for flood prediction and mapping. Fries and Kerkez (2018) used water level sensors across the state of Iowa and outputs from the National Water Model (NWM) to dynamically map large-scale models to site-scale forecasts for flood warnings. In a consequence, Barker and Macleod (2019) developed a prototype as a real-time social geodata pipeline for flood data collection and visualization across Scotland. Ning et el. (2019) implemented a prototype system to screen flooding photos driven from social media for Hurricanes Harvey (August 2017) and Florence (September 2018) using Convolutional Neural Networks (CNNs) algorithms. Their analysis revealed that CNNs on average can detect 46% to 95% of flood objects in an image.

This study aims to discuss the workflow of a Flood Analytics Information System (FAIS version 4.00; Donratanapat et al., 2020) as an example of AIoT system. FAIS can be utilized for real-time flood data analytics and situational awareness as well as post flood impact assessment. FAIS workflow includes collecting and visualizing data from various sources, analyzing the potential at risk areas to flooding using Twitter geotagged data, providing image-based data analytics models for floodwater depth and inundation area calculation, as well as flood frequency analysis (FFA) for the USGS gauging stations. FAIS proved to be a robust and user-friendly tool for both real-time and post-event analysis of flooding at regional scale that could help stakeholders for rapid assessment of real-time flood situation and damages.

This paper is intended for researchers and developers engaged in the areas of AIoT and software development in flood related research and is organized as follows. Section Methodologies discusses different components of an AIoT system followed by a discussion on multiple algorithms used in the FAIS workflow design. Section Application presents FAIS implementation results including data collection, at-risk area calculation, data analytics and FFA. Finally, the conclusions and future work of this study is presented in Section Discussion and Future Works.

Methodologies

In this section, a general approach about AIoT system is presented along with the algorithms used to design the FAIS workflow.

Components of the AIoT system

Three main components of an AIoT system including IoT data analytics system, APIs, and AI/ML systems are explained. As shown in Figure 1, having the capability to monitor and collect real-time data through IoT devices can provide a solution for flood forecasting process and decision making. With an AIoT, AI is embedded into infrastructure components, such as programs and data analytics, all interconnected with IoT networks. APIs are then used to extend interoperability between components at the device level, software level, and platform level. In this way, AIoT acts as software-as-a-service (SaaS) to collect and analyze the data and provide timely assessment. SaaS enables data and model provision as a service and provides controlled access to data through APIs. Each component of AIoT is discussed briefly below.

IoT analytics systems

Devices and technology connected over the IoT can monitor and measure big data in real time. These data can offer valuable insights to simulate flooding events across large drainage system. Around 80% of big data are image-based data (Donratanapat et al., 2020) that can be collected through IoT sensors and commercial devices such as the USGS river cameras, Department of Transportation (DOT) traffic monitoring devices, and weather tracking systems. The data can then be transmitted, saved, and retrieved at any time. The huge amount of data that IoT sensors and devices generate must be processed before the information can be used in flood related studies. However, because the data often come in different formats, there are several steps that users must take before processing or applying any type of analytics to the real-time data. These steps are:

1. Standardize or transform the data to a uniform format, ensuring that format is compatible with flood forecasting models.
2. Store or create a backup of the newly transformed format.
3. Filter any repetitive, outdated, or unwanted data to help improve forecast accuracy.
4. Integrate additional structured (or unstructured) data from other sources to help enrich flood data sets.

Flood related dataset can greatly benefit from several types of data analytics procedures including (i) Time series analytics which are based on time-based data, and data are analyzed to reveal any anomalies, patterns, or trends, (ii) Streaming analytics, referred to as event stream processing, facilitates the analysis of massive "in-motion" datasets necessary for real-time streaming and forecasting, and (iii) Spatial analytics which can be used to analyze location-based dataset to reveal various geographic patterns, determining any type of spatial relationship between various parameters in flood forecasting. The best example of spatial analytics is using geospatial patterns and data to track and map flood inundation areas in real-time.

Application programming interface

The inherent programmability and the extended use of open APIs enable innovation in areas related to monitoring, modeling, and overall operational management of flood forecasting research. APIs are used by modelers/developers for implementing various features in flood modeling tools. Developers simply use an API call within their software to implement complex features and interact one piece of code with other portions. The goal is not only to collect real-time flood dataset, but also to monitor and forecast these extreme events. To perform any flood modeling and analytics study, a data retrieval approach is required to collect data from various sources before modeling set up. The interfaces of these application programs convey data in the form of Web Feeds, such as Really Simple Syndication (RSS), Web Services and Screen Scraping. Many APIs can be accessed over the Web using the HTTP protocol based on RESTFUL or Streaming services. Web API can be used in a web server or a web browser. The speedy collection and transmission offered by various APIs provided a seamless data integration within any AIoT system.

AI systems

AI and its sub fields, machine learning and deep learning approaches, have been recently used to solve many

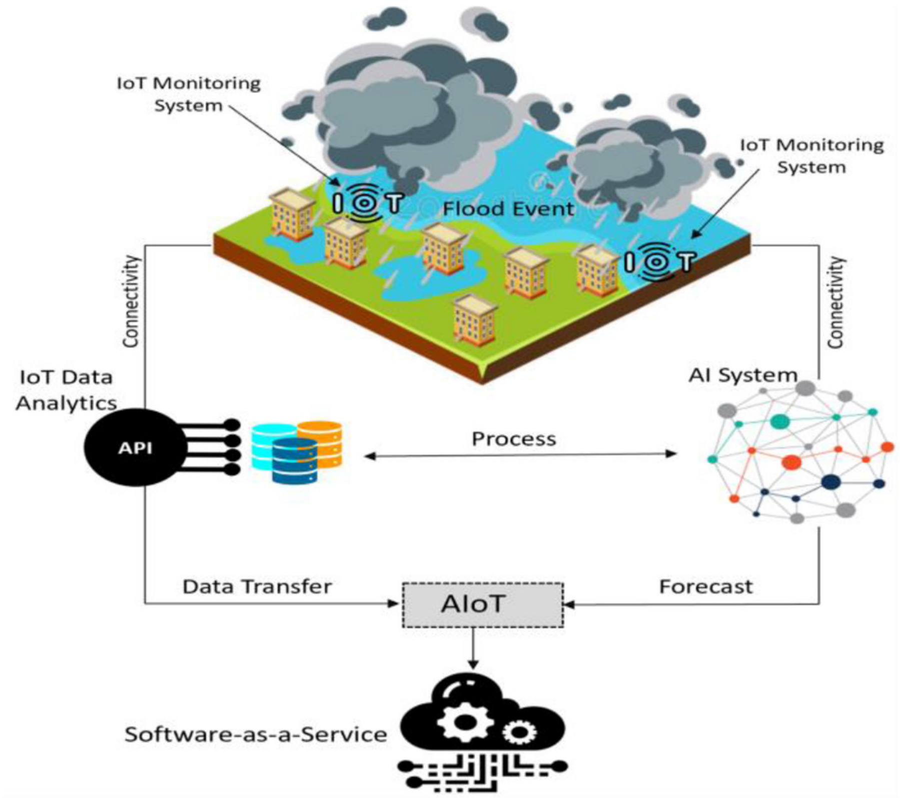


FIGURE 1
Different components of an AIoT system.

environmental related modeling problems. These intelligent systems provide new possibilities for flood studies as more data becomes available and computing power increases. New advances in these intelligent systems provide new methods to simulate time series data (see Samadi et al., 2021) as well as image-based data (Pally and Samadi, 2022). Among many AI algorithms, deep learning methods have been recently applied to flood analytics research such as flood image label detection and recognition. In flood image processing research, deep learning methods such as CNNs can be used to extract visual information from images. CNNs are formed by layers of convolutional filters, fully connected layers, and non-linear operations (Pally and Samadi, 2022). Typically, the first layers contain filters that detect simple features like edges or color, while filters in deeper layers detect more complex features (e.g., complex shapes, objects, etc.).

FAIS structure

FAIS is an example of AIoT application for flood analytics research. The prototype is inspired by the needs to assess the impacts of successive hurricane events in the south and southeast United States. Given the massive amount of data, an automated big data and crowd sourced information system is needed to collect real-time data and create a map-based dashboard to better determine at-risk locations to flooding. These needs and discussion along with existing difficulties in gathering massive data across various Web services provided a comprehensive roadmap for the FAIS development.

FAIS is designed as both a Python package as well as a Web Python platform to collect the data during historical and real-time flood events and visualize impacted areas. FAIS uses IoTs-APIs and various machine learning approaches for transmitting, processing, and loading big data through which the application gathers information from various data servers and replicates it to a data warehouse (IBM database service). Users are allowed to directly stream and download flood related images/videos from the USGS and DOT and save the data on a local storage. The outcomes of river measurement, imagery, and tabular data are displayed on a web based remote dashboard and the information can be plotted in real-time. Several tools and algorithms are integrated within the FAIS application that are discussed below:

Data analytics approaches

To perform data analytics approaches, a new Python package called "FloodImageClassifier" was developed and integrated within the FAIS application for flood image annotation and classification. "FloodImageClassifier" uses various deep learning algorithms (i.e., CNNs) for flood image labeling, inundation area calculation, and flood level classification. Several CNN algorithms such as YOLOv3 (You look only once version 3; Redmon et al., 2016), Fast R-CNN (Region-based CNN; Girshick, 2015), Mask R-CNN (He et al., 2017), and SSD MobileNet (Single Shot MultiBox Detector MobileNet; Liu et al., 2016) were developed and integrated within "FloodImageClassifier" to detect flood labels and estimate floodwater depth. These algorithms encapsulate CNNs in an API to classify images into some categories and assign them sensible labels and scores. MAP (Mean Average Precision) index was used to determine the performance of object detection approaches which is a popular performance metric to evaluate algorithms that involve predicting the object location as well as classifying the probability of occurrence. MAP evaluates the correctness of bounding box prediction using a metric called Intersection over union (IoU). IoU is a ratio between the intersection and the union of the predicted boxes, and the ground truth boxes. Canny edge detection and aspect ratio concept (see Canny, 1986) were then programmed in the package for floodwater level estimation and risk classification. These approaches were performed in order to detect the edges of the water surface as they calculated the surface areas of water which in turn were used to determine the water level.

CNNs algorithms include non-linear approaches for label detection and segmentation such as Rectified Linear Unit (ReLU), a layer that outputs the positive part of the input, and Max-pooling, a layer that performs a non-linear down-sampling. Typically, CNNs include fully connected layers that apply a set of linear and non-linear operations to the extracted features and output a set of probabilities for each class. The multiple layers contain weights and parameters that are learned from training samples using backpropagation and gradient-based optimization methods such as Stochastic Gradient Descent (SGD) or Adaptive Moment Estimation (Adam; Kingma and Ba, 2014). "FloodImageClassifier" package uses "keras" deep learning library to build the CNNs classifier. A schematic figure of developed CNNs is illustrated in Figure 2 that consists of the following layers:

Input layer: The first layer of the CNN is the input layer which takes an image as input, resizes the image and passes the image to the next layer for feature extraction.

Convolutional layers: Three convolutional layers were designed in the model to apply small filters on each part of the image, match the feature points within the image and extract features from the image.

Pooling layer: The extracted features are passed onto the pooling layer, which helps in reducing the special dimensions by shrinking the images down while preserving the most important information within them. It picks the highest values from each region that is retained the best fits for each feature within that region.

Rectified linear unit layer: This layer normalizes the pooling layer obtained values by replacing the negative values with zeros to help the CNN stay mathematically stable.

Fully connected layers: This is the final layer which takes the filtered images as input and then divides them into categories along with their respective labels and scores.

Based on the CNN structure presented in Figure 2, four CNN algorithms including YOLOv3, Fast R-CNN, Mask R-CNN, and SSD MobileNet were developed and integrated

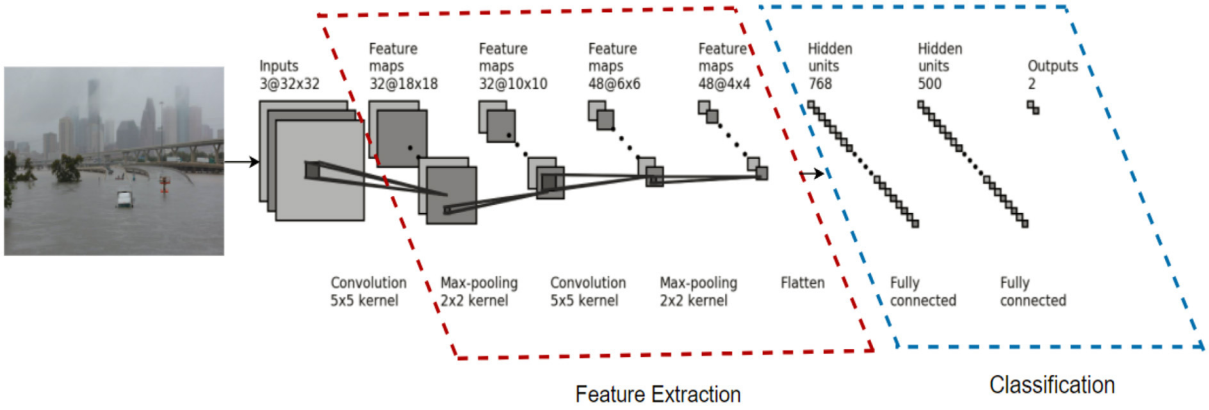


FIGURE 2
The architecture of floodimageclassifier package. Features within the red dash line reflect pooling layers while the features within the blue dash line represent the connected layers.

within the FAIS application. Canny edge detection and aspect ratio concept were also programmed in the package to calculate floodwater level and risk classification. Readers are referred to Girshick (2015), Redmon et al. (2016), He et al. (2017), and Liu et al. (2016) for more information on these algorithms.

FAIS APIs

FAIS was initially developed as a python package targeting two sources of data i.e., USGS and Twitter. The package was then transferred to a web Python platform to collect the data during historical and real-time events and visualize flooding impacts. USGS collects and stores multiple water data including river flow data (flood, streamflow, gauge heights, etc.), water quality, ground water levels, and precipitation at defined gauging stations which are strategically placed at the outlets of rivers and lakes. These placements allow the USGS to correctly monitor and collect the data and compute several statistical indices related to the river flow across the nation. USGS provides two different types of flow data including real-time and historical records based on date and time. FAIS uses the USGS API to document the available service endpoints and leverage the various water data sources. Readers are referred to Donratanapat et al. (2020) for more information on the USGS API development and implementation for the FAIS application.

FAIS also gathers social media data through Streaming and Search APIs. The increasing interaction of federal agencies with social media during the flood crisis has shown that social media platforms are important tools to assess flood situations in real-time. Twitter, one of the biggest social media platforms, contains many fast-paced real-time data and a wide range of historical information regarding news and events that occur in the local and global scales. Being a famous social media, local residents, and government agencies use Twitter to provide information during catastrophic events. For example, during Hurricane Florence (2018), many residents in SC tweeted real-time information about the local damage, road closure and shelter information. The government agencies such as NWS, NHC, SCDOT, and USGS also used Twitter to provide updates about the damaged infrastructure, emergency situations, and resources in real-time and during post-event. These data are valuable because they can provide location specific information about flooding conditions in real-time.

FAIS application collects tweets using two different approaches, i.e., Search and Streaming APIs. These two approaches are based on two use cases of the applications which are historical data gathering as well as streaming/real-time data collection. Search API is more suited to singular and specific queries of tweets, whereas the Streaming API provides a real-time stream of tweets. The *Tweepy* Python package is first integrated within the FAIS application and implemented for accessing and collecting the Twitter data. Twitter developer account was then used to access Token, Token Secret, Consumer Key, and Consumer Secret to manipulate Twitter functionalities.

In addition, a bot software was developed and integrated into the FAIS prototype as part of real time crowd intelligence mechanism for Twitter data gathering. The developed Twitter bot allows users to monitor every tweet and automates all or part of Twitter activities. Specifically, the developed bot (see Figure 3) allows the user to query tweets from Twitter by a specific user and/or keyword using both Search and Streaming APIs. A reusable Python module or a module configure was created that contains logic steps for implementing the bot functionalities. This module reads the authentication credentials from environment variables and creates the Tweepy API object. The bot reads the credentials from four environment variables including CONSUMER_KEY, CONSUMER_SECRET, ACCESS_TOKEN, and ACCESS_TOKEN_SECRET. After reading the environment variables, the bot creates the *Tweetpy* authentication object that eventually can be used to create an API object. The administrator can choose to activate or deactivate the bot and change the keywords for Streaming purposes. FAIS Twitter bot contains three main components notably:

1. Twitter Client: This component communicates with the Twitter API and authenticates the connection to use its functionality. Twitter Client also hosts a function called tweets_listener, that continuously stream tweets and search for the matched keywords. Once tweets_listener finds the match it will then communicate to the other two components.
2. Tweet Analyzer: It analyzes the tweets and gives the result a score after a match is found.
3. Twitter Streamer: This module streams tweets from pre-specified eight keywords, analyzes the data, and organizes them into a data frame. The collected tweets then store in a MongoDB database and display on the FAIS Twitter Streaming section as tabular data. The workflow explaining how the bot gathers real time tweets using the Twitter bot is illustrated in Figure 3.

Flood frequency analysis approach

FAIS provides frequency analysis to estimate extreme flood quantiles that combines elements of observational analysis, stochastic probability distribution and design return periods. FFA techniques predict how flow values corresponding to specific return periods or probabilities along a river could change over different design periods. FFA can be used to estimate the design flow values corresponding to specific return periods for designing structures such as dams, bridges, culverts, levees,

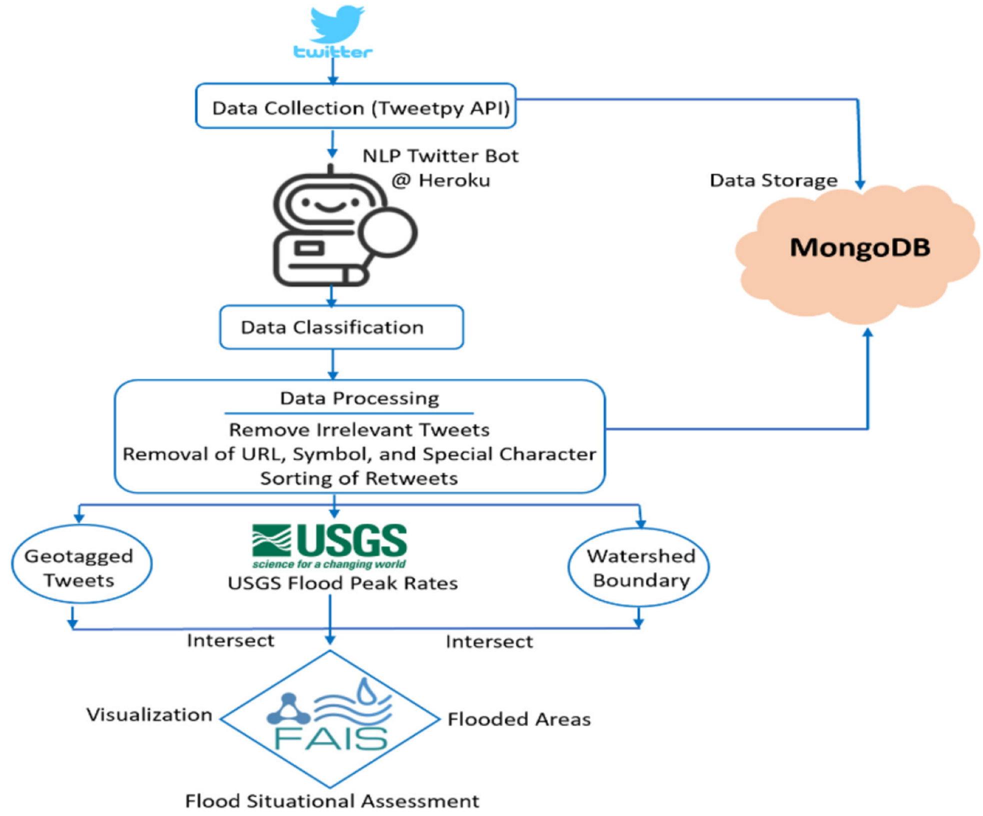


FIGURE 3
The workflow of Twitter APIs and Twitter bot designed in the FAIS application.

highways, sewage disposal plants, waterworks, and industrial buildings. FFA is useful in providing a measurement parameter to assess the damage corresponding to specific flow during flooding event. Accurate estimation of flood frequency not only helps engineers in designing safe infrastructure but also in protection against economic losses due to maintenance of structures. FAIS uses several Python packages such as "scipy.stats" (Bell et al., 2021) and "reliability" (Reid, 2021) to retrieve annual peak flow rates for provided years and calculate distribution parameters to create frequency distribution graphs. The tool fits various probability distributions including Normal, Lognormal, Gamma, Gumbel, Pearson Type III, Weibull, and Loglogistic distributions. These distributions and their probability density functions, and mathematical formulations are presented in Table 1. However, the accuracy of FFA estimates may vary using different probability distributions. Hydrologists typically recommend using Pearson Type III, but other distributions such as Gumbel function can be also used for a river system with less regulation and less significant reservoir operations, diversions, or urbanization effects. It is also important to quantify the precision of estimates, so the tool calculates flood frequencies within 95% confidence interval and provides the accuracy of the calculation. A 95% confidence level indicates that the interval between lower bound and upper bound contains the true value of the population parameter (flood data).

The usual domains for these seven distributions are the whole real line for the normal density functions and values larger than ε for the Pearson Type III density function, which can, in principle, be any real number. The Pearson type III distribution has been adopted in many FFA assessment as the standard distribution because of its better fit and performance (Sumioka et al., 1997). If $\varepsilon = 0$, then Pearson type III distribution reduces to the gamma distribution. The default estimation parameter method for these distributions is maximum likelihood estimation (MLE), although other approaches such as least-squares, weighted moments, linear moments, and entropy to compute parameter values are also used in the prototype. Readers are referred to Vogel et al. (1993) and Singh (1996) for more information on the probability distributions for FFA and parameter estimation methods.

TABLE 1 Probability distributions and their density functions used in the FFA section of FAIS application.

Distribution	PDF	Assumption	Domain
Gamma	$\frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\frac{x}{\beta}}$	$\alpha > 0, \beta > 0$	$x > 0$
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$\mu = 0,\ \sigma = 1$	$x\epsilon(-\infty, +\infty)$
Log-normal	$\frac{1}{\sqrt{2\pi}\sigma x}\ e^{-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2}$	$\alpha > 0$	$x > 0$
Weibull	$\frac{\alpha x^{\alpha-1}}{\beta^{\alpha}} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$	$\alpha > 0,\ \beta > 0$	$x > 0$
Pearson type III	$\frac{1}{\Gamma(\alpha)\beta}\left(\frac{x-\varepsilon}{\beta}\right)^{\alpha-1}\ e^{-\frac{x-\epsilon}{\beta}}$	$\alpha > 0,\ \beta > 0$	$x > e^{\varepsilon}$
Gumbel	$\frac{1}{\beta}\ \exp\left[\frac{x-\alpha}{\beta} - \exp\left(\frac{x-\alpha}{\beta}\right)\right]$	$-\infty\ \leq\ x\ \leq\ \infty$	$\alpha = 0,\ \beta = 1$
Loglogistic	$\frac{\alpha x^{\alpha-1}}{\beta^{\alpha}\left[1+\frac{x}{\beta}\right]^2}$	$\alpha > 0,\ \beta > 0$	$0 \leq \pi\ < +\infty$

Application

FAIS implementation

Different components of the FAIS prototype were implemented for collecting data and assessing flood risk and identifying at-risk areas to flooding in real-time. These executions are explained below.

USGS data gathering

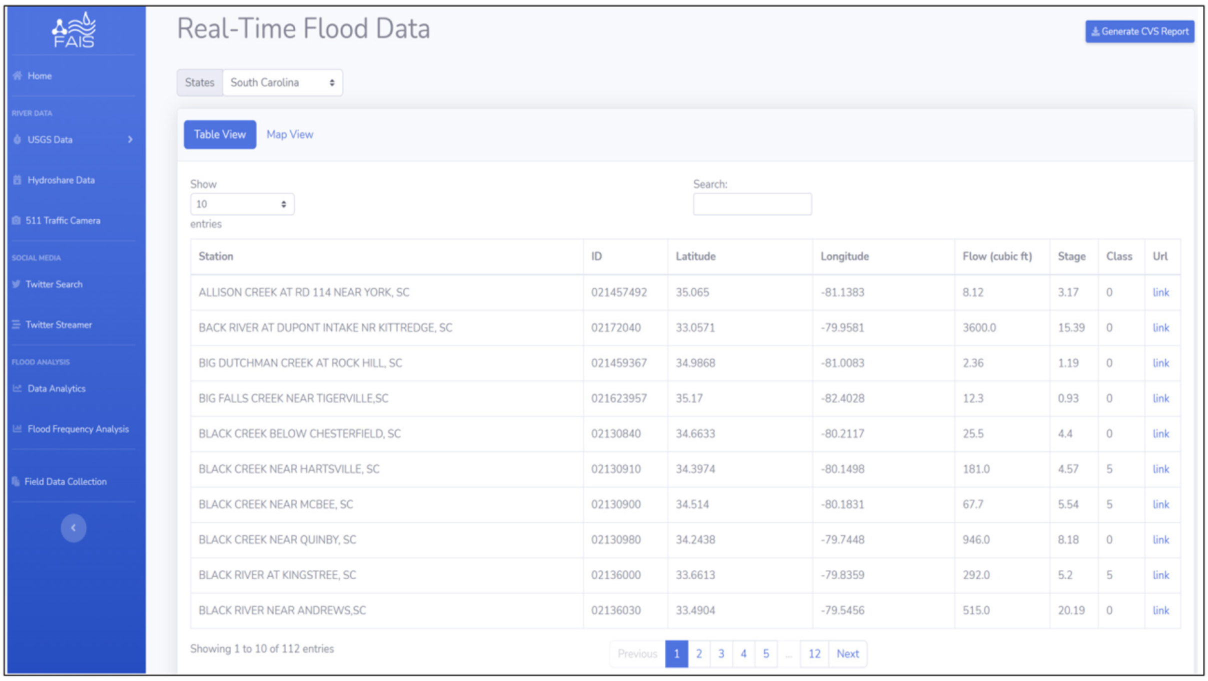
FAIS uses USGS APIs for real-time and historical data collection as well as images collection from the USGS river cameras. Users can collect real time and historical discharge (cfs) and gauge height (ft) data and display the gauge geolocation using the FAIS visualization dashboard. A summary of data including maximum gauge height, latitude, longitude, and elevation can also be retrieved on the map dashboard by clicking on each USGS station. Gathering USGS historical data involves selecting the target state, the interested station, and the date. After the query criteria is entered, FAIS creates a request URL and sends it to the USGS server for collecting the data. The prototype displays the data as "Table View" as well as "Map View" and plots the results (see Figures 4–6). Users can also upload a .csv format file of all the collected data that contains station name, ID, latitude, longitude, discharge, gauge height, and the USGS original URL. The prototype also gathers the USGS flow data through HydroShare Web services (results not shown here). It uses HydroShare Representational State Transfer (REST) API to access the data through the Web user interface programmatically.

Flood risk assessment using Twitter geotagged data

During flooding events, citizens use Twitter to share flood information such as damages, road closure, shelter information, etc. Government agencies such as NWS, NHC, DOT, and USGS also use Twitter to disseminate data and updates about flooding conditions, damaged infrastructure, emergency situations, evacuation route, and other resources. A tweet can provide a variety of information, such as text, images, videos, audio, and additional links. In addition, there is also a significant number of metadata that is attached to each tweet. This metadata includes information regarding geolocation (either a place name or coordinates), the author's name, a defined location, a timestamp of the moment the tweet was sent or retweeted, the number of retweets, the number of favorites, a list of hashtags, a list of links, etc. This information is valuable and has the potential to provide reliable information and actionable intelligence when attempting to extract tweets and use themto assess flood situations.

FAIS uses NLP to cleanse, filter, and group flood related tweets including tweet geolocation information, related images, etc. To accomplish this need, a Twitter Streaming bot (functions on both iOS and Mac) was developed and deployed at Heroku cloud platform outside of the application access which is controlled by the Heroku User Interface. Heroku is a cloud platform as a service (PaaS) that enables system-level supervision and coordination of Twitter APIs, crowd sourced data, and tweets. Twitter bot automates tweet gathering and continuously cleans and monitors all Twitter activities during real time implementation. During a real-time flooding event, the bot gets notification when new content, such as tweets that match certain criteria (keywords) is created. Overall, eight keywords including "Flood Damage," "Road Closure," "Emergency Management and Response," "Flooded Neighborhood," "Infrastructure Damage," "Evacuation Route," "Shelter and Rescue," and "Storm Surge" are incorporated into the Twitter data gathering section.

FAIS was used to identify at-risk areas to flooding during Hurricane Laura. Hurricane Laura was a deadly and destructive category 4 hurricane that made landfall in the U.S. state of Louisiana (LA) in August 2020. While impacted many countries and people on its path, Laura prompted many warnings and watches across LA coastal communities. Early on August 27, Laura made landfall near peak intensity on Cameron, LA and was recorded as the tenth-strongest U.S. hurricane that made landfall by wind speed on record. When Laura approached LA, FAIS application was used to constantly stream the tweets and



Real-Time Flood Data

States: South Carolina

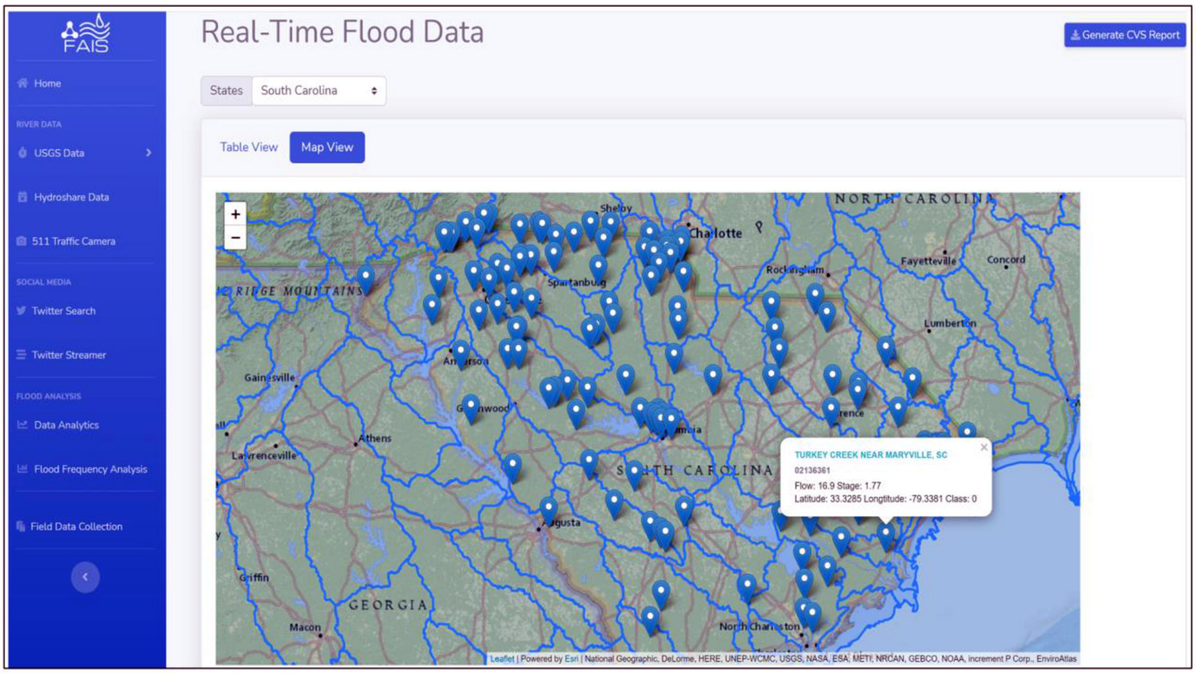
Show 10 entries

Search:

Station	ID	Latitude	Longitude	Flow (cubic ft)	Stage	Class	Url
ALLISON CREEK AT RD 114 NEAR YORK, SC	021457492	35.065	-81.1383	8.12	3.17	0	link
BACK RIVER AT DUPONT INTAKE NR KITTREDGE, SC	02172040	33.0571	-79.9581	3600.0	15.39	0	link
BIG DUTCHMAN CREEK AT ROCK HILL, SC	021459367	34.9868	-81.0083	2.36	1.19	0	link
BIG FALLS CREEK NEAR TIGERVILLE,SC	021623957	35.17	-82.4028	12.3	0.93	0	link
BLACK CREEK BELOW CHESTERFIELD, SC	02130840	34.6633	-80.2117	25.5	4.4	0	link
BLACK CREEK NEAR HARTSVILLE, SC	02130910	34.3974	-80.1498	181.0	4.57	5	link
BLACK CREEK NEAR MCBEE, SC	02130900	34.514	-80.1831	67.7	5.54	5	link
BLACK CREEK NEAR QUINBY, SC	02130980	34.2438	-79.7448	946.0	8.18	0	link
BLACK RIVER AT KINGSTREE, SC	02136000	33.6613	-79.8359	292.0	5.2	5	link
BLACK RIVER NEAR ANDREWS,SC	02136030	33.4904	-79.5456	515.0	20.19	0	link

Showing 1 to 10 of 112 entries

FIGURE 4
USGS real-time flood data gathering interface. Users can download a .csv file of the data and visit the original data sources at the USGS portal.

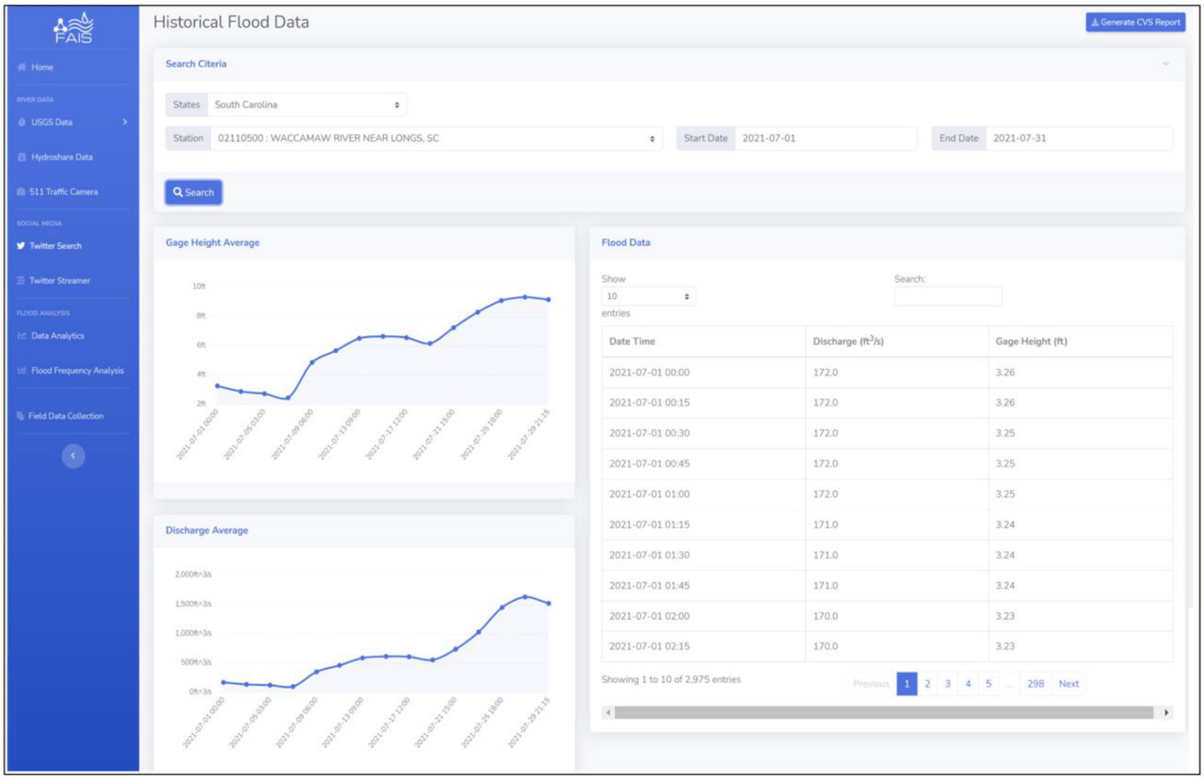


Real-Time Flood Data

States: South Carolina

TURKEY CREEK NEAR MARYVILLE, SC
02136361
Flow: 16.9 Stage: 1.77
Latitude: 33.3285 Longtitude: -79.3381 Class: 0

FIGURE 5
Map view interface of the USGS flood gauges. A summary of data for each gauging station can be retrieved from Map View section.



Historical Flood Data

Search Citeria

States: South Carolina

Station: 02110500 : WACCAMAW RIVER NEAR LONGS, SC

Start Date: 2021-07-01

End Date: 2021-07-31

Flood Data

Show 10 entries

Search:

Date Time	Discharge (ft^3/s)	Gage Height (ft)
2021-07-01 00:00	172.0	3.26
2021-07-01 00:15	172.0	3.26
2021-07-01 00:30	172.0	3.25
2021-07-01 00:45	172.0	3.25
2021-07-01 01:00	172.0	3.25
2021-07-01 01:15	171.0	3.24
2021-07-01 01:30	171.0	3.24
2021-07-01 01:45	171.0	3.24
2021-07-01 02:00	170.0	3.23
2021-07-01 02:15	170.0	3.23

Showing 1 to 10 of 2,975 entries

FIGURE 6
USGS historical flood data collection and visualization interface. Users can display the data for any state and any USGS station by selecting the state and USGS gauging station. A .csv format file of the data can be also downloaded directly from the interface.

identify at-risk areas to flooding. Georeferenced tweets were gathered and filtered by eight keywords (mentioned above) and queries across the shortlisted areas in LA. The retrieved tweets were then used to intersect a maximum of seven locations with the USGS peak flood rates and watershed boundary. Geotagged tweets coordinates considered as a center-point for ∼16 km wide square boxes. This size is arbitrarily chosen to cover the areas nearby to each gauge. These seven at-risk locations were monitored during Hurricane Laura for any updated tweets whose geotag intersected a bounding box, which constituted a "match." The retrieved tweets were then stored in a MangoDB database which is widely used as an open-source database to store JSON format files. Due to the size of queried data, the Twitter bot filtered the data (>95% of uninterested/mismatched tweets) and only kept those flood related tweets that match the eight keywords mentioned above along with the text, geolocation, author ID, and date.

To identify at-risk locations, FAIS first cycles through a set of USGS web addresses to find river gauge height readings, parsing these flat files using Python web scraping technique and obtaining all the latest river levels. Each river level reading is compared with its respective long term cached average level to identify the highest relative river levels in real-time. The highest river level will then be intersected with watershed polygons as well as geotagged tweets (filtered by Streaming bot) to identify flooded locations in real-time. To update at-risk locations in real-time, a shell script in Python runs on a local computer server, the script is reset every 3 h to update at-risk areas to flooding from the latest national and environmental data sources as well as geotagged data. A period of 15 min was initially chosen as the intended trade-off between tracking the latest at-risk location forecasts. API updates can be varied between 15 min to several hours but based on real-time testing the period extended to 3 h to allow some reaction time from those areas on Twitter. However, the choice of time period depends solely on the project requirements as well as the flooding impact and severity. During Hurricane Laura, the time between a tweet appeared online and visually plotted in the FAIS as being potentially relevant (in terms of location and content) was in the order of few seconds to 2 min, thereby this rapid analysis could provide an early information channel to the stakeholders for emergency management, asset allocation and rescue plan. A web-based console and a visualization tool—GeoJSONLint13—are used in the FAIS application to view results and inspect the polygons.

In addition, the prototype was operationally tested during Hurricane Dorian (September 04-06, 2019) in the Carolinas and georeferenced tweets were also gathered in real-time to identify at-risk locations to flooding (see Donratanapat et al., 2020). A list of real-time tweets is presented in Table 2. The tweets were then used to intersect with the USGS gauging stations and watershed boundary to identify at-risk areas (see Figure 7).

Big data analytics

Data analytics section of the FAIS application uses "FloodImageClassification" package for flood image labeling, inundation area calculation, and flood level classification. Four CNN algorithms such as YOLOv3, Fast R-CNN, Mask R-CNN, and SSD MobileNet were programmed in the "FloodImageClassification" package to detect flood labels and estimate flood depth and inundation area. These algorithms encapsulate CNNs in an API to classify images into some categories and assign them sensible labels and scores. Less than eight thousands flood images were first collected from the USGS river web cameras, DOT traffic images, and search engines such as Google and Bing. The data were formatted and split into training and test datasets. The models were trained for 27 epochs with a batch size of 72. The images were partitioned into training and testing sets in the ratio of 9:1. The images present within the training set were resized and normalized before training. Once all the images were resized and normalized, "labelIme" Python package was used to annotate these images with eight different object categories such as "vehicle," "forest," "tree," "traffic sign," "water vessels," "residential areas," and "critical infrastructure." Annotation of images involves highlighting each of the objects within an image manually using bounding boxes and labeling them appropriately. MAP index was used to determine the performance of each object detection algorithm. Canny edge detection and aspect ratio concepts were also used to calculate flood depth and inundation areas.

Given an input image, images were first resized and converted into a grayscale. Next, the skyline was identified and eliminated from the images. Since both water and skyline have the same color gradient, it is possible the skyline could be detected as a water surface. Therefore, the skyline was eliminated and a portion of images consisting of the water surface was taken into consideration. For each image, the edges of the water surface were first detected, and the associated contour was automatically drawn around the water surface and the area of the contours (i.e., the area of water surface) was calculated. Next, based on the aspect ratio which is calculated by taking into consideration the area of water surface, the floodwater level was estimated. The estimations were then categorized into "shallow," "moderate," and "deep" to reflect flood risk conditions. This process is automated in Python to provide the user a seamless procedure to perform data analytics in the FAIS application. This automatics procedure takes on average about 2–3 min to complete. As illustrated in Figures 8–11, it is evident that the CNN models were capable of detecting multiple objects within a single image as it almost detected 90% of the objects precisely. However, object detection models produced different outcomes. Segmentation models such as Mask R-CNN identified the foreground shapes and highlighted the objects using bounding boxes as well as by drawing a mask on the object. This helped in clearly segmenting one object from another one whereas other object detection models such as Fast R-CNN, YOLOv3, and SSD MobileNet highlighted the detected objects using a single bounding box only.

The prediction scores (MAP) of different CNN models were calculated for different object categories namely vehicle, person, forest, tree, traffic sign, residential area (i.e., houses), water vessels (i.e., boats, ships, etc.) and critical infrastructure (bridges, dams, etc.) by passing the same set of test images to each of these four CNN models (Table 3). This determined which of these models is the best object detection approach for flood images. Results revealed that Mask R-CNN and Fast R-CNN were particularly skillful in detecting labels and segmenting them, although Mask R-CNN showed better performance in detecting labels. The processing to perform object detection for each CNN algorithm varies between 2 and 4 min depending on the system memory and processor.

Once the detection results were generated by CNN models, each detected object was removed from the images and the images were reconstructed by filling the void spaces in a plausible manner using exemplar based inpainting (see Criminisi et al., 2003). We used partial convolutions with an automatic mask update to fill in the voids within an image. Image inpainting model substituted convolutional layers with partial convolutions and masked the updates. This algorithm successfully identified the target region which was filled using the surrounding areas of the target region as reference. The filled region was used to calculate floodwater depth and inundation area. As shown in Figure 12, the edges of the water surface (i.e., the drawn contours) was first detected and then the area of the water surface was calculated. To isolate the water surface from the other objects that were highlighted, OpenCV's *findContours()* function was used to identify the foreground mask shapes and to draw contours around them. Next, the area of each of the contours was calculated, the contour areas were then sorted and only the largest contour was printed over the original image, allowing us to accurately segment and calculate the water surface area (flood inundation area). The area is also calculated based on the pixel dimensions and the number of pixels located in the flooded area. After highlighting the water surface using contours, a bounding box was drawn around the contour to calculate the aspect ratio. The aspect ratio value is used to determine the floodwater level as shown in Figure 12. If the aspect ratio is in a range of 1.27–1.8 the floodwater level is then considered to be low, if the aspect ratio is in a range of 0.54–1.26 the floodwater level is considered to be moderate, and if the

TABLE 2 The relevant tweets and tweet geolocation and time for 12 at-risk locations in LA during Hurricane Laura flooding.

Text	Location	Username	Date/time	Latitude	Longitude
Watches for flash flood, hurricane, and surge. I should not get direct surge from pontchartrain, at 2–4 feet, but it will have an impact on drainage since I am so close to the lake. This was the same fear I had when it looked like Laura was gonna hit this area	South Louisiana, Plain Dealing, Bossier Parish, Louisiana, 00710, USA	BonnieBlueTK	Mon Aug 24 21:48:18 +0000 2020	32.9016393	−93.7047546
Man I've been in earthquakes, 93 flood, hurricanes, tornadoes, and wildfires… I just went through hurricane Laura with the eye right over my home	Louisiana, USA	darinnstacy	Wed Aug 26 03:10:16 +0000 2020	30.8703881	−92.007126
Calcasieu Parish officials warning that roughly 70% of area south of I-10 could see flooding from Hurricane Laura and that I-10 could flood in the parish	Baton Rouge, East Baton Rouge Parish, Louisiana, USA	MelindaDeslatte	Tue Aug 25 21:34:10 +0000 2020	30.4459596	−91.18738
Hurricane Laura strikes Louisiana as Category 4 storm, battering Lake Charles area and bringing flood threat—The Washington Post	Louisiana, USA	StateStatus_LA	Thu Aug 25 10:39:32 +0000 2020	30.8703881	−92.007126
Hurricane Marco and Tropical Storm Laura could bring heavy rains to this area causing flooding. Stay safe! Do not venture…	New Orleans, Orleans Parish, Louisiana, USA	kingfrost_24	Mon Aug 24 19:26:47 +0000 2020	29.9499323	−90.0701156
Calcasieu Parish officials warning that roughly 70% of area south of I-10 could see flooding from Hurricane Laura…	Baton Rouge, East Baton Rouge Parish, Louisiana, USA	CEStephens	Tue Aug 25 21:35:08 +0000 2020	30.4459596	−91.18738
Under a Hurricane Watch for my area now. Meteorologists are saying get ready for a "one two punch" hurricane Marco on Monday, Hurricane Laura on Wednesday. Flood gates were being closed this evening to protect the City from storm surge	New Orleans, Orleans Parish, Louisiana, USA	DanielBailleau	Sun Aug 23 01:24:19 +0000 2020	29.9499323	−90.0701156

aspect ratio is in a range of 0.18–0.54 and less than 0.18 (<0.18) the floodwater level is considered to be high (sever flood risk; see Table 4).

FFA for the USGS 02147500 Rocky Creek at Great Falls, SC

FAIS uses multiple probability distributions such as Normal, Lognormal, Gamma, Gumbel, Pearson Type III, Weibull, and Loglogistic distributions to compute FFA for any given flood gauging station in US. Figure 13 shows FFA results for the USGS02147500 Rocky Creek at Great Falls, SC. Among several probability distributions, Loglogistic distribution was a perfect fit followed by the Weibull distribution. Both log-Pearson type III and log-normal distributions were also showed appropriate fits when the maximum likelihood method was used for parameter estimation. As illustrated, all annual flood data points (overall 64 data points) were bracketed within 95% uncertainty bound using the Loglogistic distribution while the uncertainty bound for the Weibull distribution skillfully bracketed moderate flood rates and showed less performance with respect to low and maximum flood data. Analysis suggests that annual flood peak of ∼15,000 cfs for the Rocky Creek represents a design return period of 25-year. The design discharge of a 100-year flood (1% probability) occurring at the USGS02147500 gauging station is >30,000 cfs. This gauge is associated with Lower Catawba basin (hydrologic unit code 03050103) Chester County, SC with a drainage area of 194 square miles which is mostly a

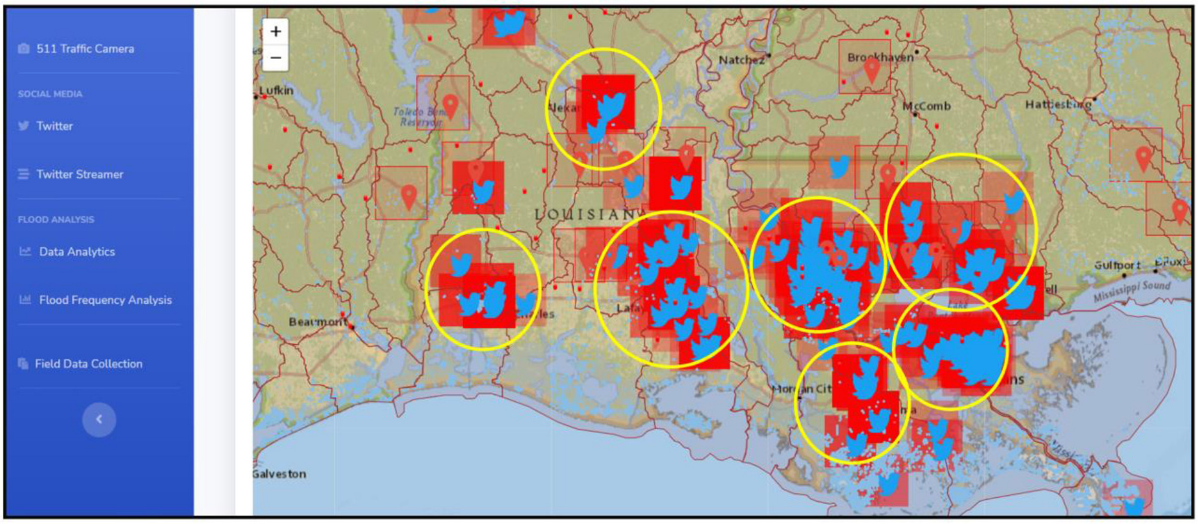


FIGURE 7
At-risk locations to Hurricane Laura in LA. Yellow circles indicate at-risk locations identified by using collected tweets in real-time that were intersected with the USGS peak flood data as well as the watershed boundary. Note that FAIS streaming approach intersects ≥ 3 tweets with the USGS peak flow rates and disregards <3 tweets for intersection.

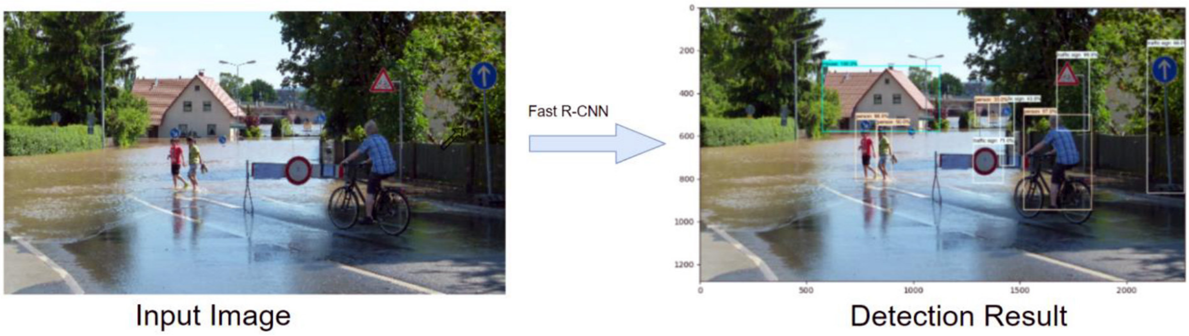


FIGURE 8
The Fast R-CNN detection results with bounding boxes.

rural basin. FFA for this location proved that high peak values made critical contributions to the upper tail of the Loglogistic probability distribution.

FAIS application also provides a .CSV format file of FFA that can be downloaded from the interface. The use of FAIS application for FFA provides an easy assessment for the design of engineering structures such as culverts, bridges, and dams. With the many challenges facing existing probability distribution fitting and performance calculation for a given design problem, FFA functions, numerical estimation and uncertainty calculation, and graphical capabilities together with its flexibility to fit multiple distributions, can go a long way. This makes FAIS application an ideal tool to assess flood frequencies for any USGS gauging station. It should be noted that FFA builds a predictive model based on existing flood data, so the accuracy is greatly enhanced with a larger dataset (>10 years).

Discussion and future works

This article elaborated on the latest concepts related to AIoT in flood related research, emphasizing the architectures and functionalities of this approach and how these capabilities were used to design FAIS application. AIoT brings AI capabilities and IoT power to flood related research, enhances data management and analytics, and aids in intelligent decision-making process. As AIoT continues to evolve into a computing paradigm endowed with a high computing rigor, it has become

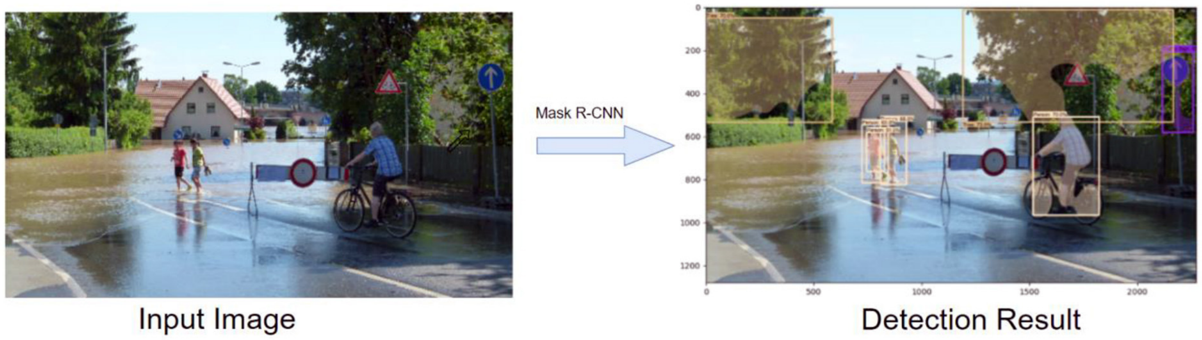


FIGURE 9
The Mask R-CNN detection results with bounding boxes.

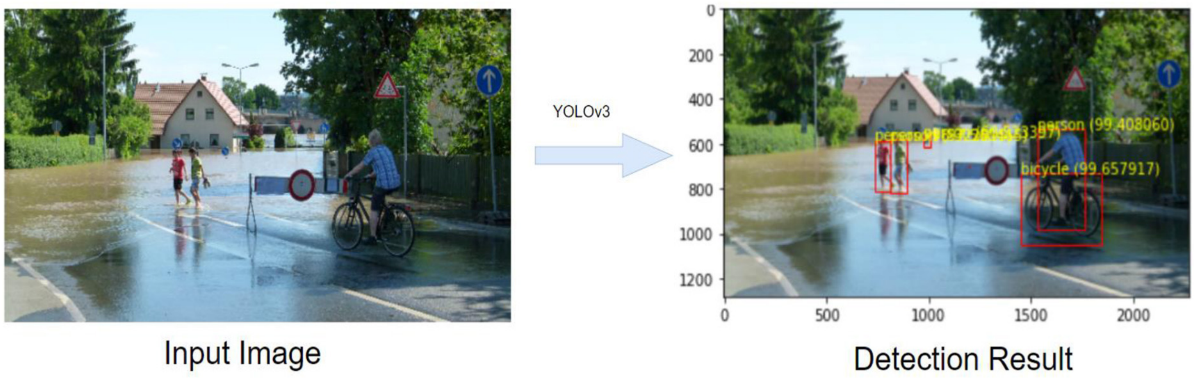


FIGURE 10
YOLOv3 detection results with bounding boxes.

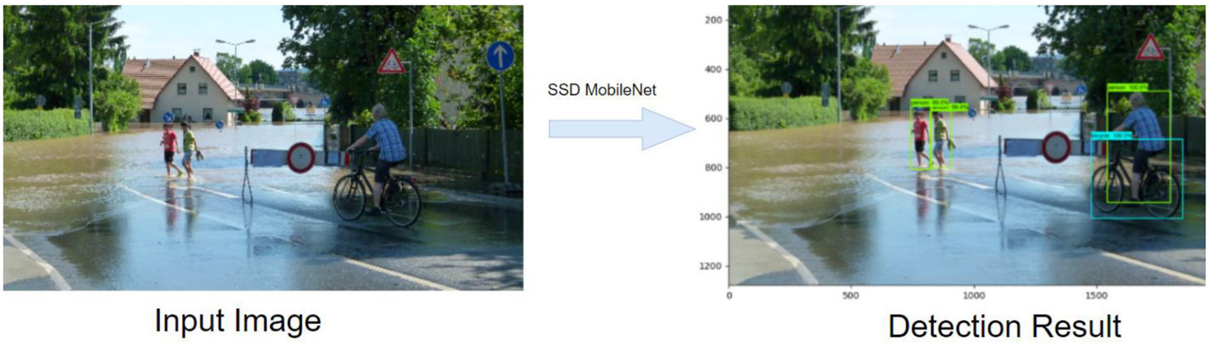


FIGURE 11
SSD MobileNet detection results with bounding boxes.

apparent that traditional solutions for training, validation, and testing models are no longer appropriate for flood computational challenges brought about by the big data and catchment complexities.

As an AIoT paradigm, FAIS enables real-time flood data gathering and analytics across different domains. This prototype has been driving the confluence of IoT, deep learning, and big data to enable a real-time assessment of domain-inspired

TABLE 3 The prediction score of different labels using multiple CNNs.

Models/object categories	Vehicle (%)	Forest (%)	Traffic sign (%)	Tree (%)	Residential area (%)	Person (%)	Water vessel (%)	Critical infrastructure (bridge, dam, road, storm water facilities, and railroad) (%)
SSD MobileNet	92	53	60	97	51	74	98	95
Fast R-CNN	99	56	99	89	100	99	48	100
Mask R-CNN	85	70	87	73	96	91	68	89
YOLOv3	99.9	*	*	*	*	98.3	95	*

Mask R-CNN was a superior segmentation model followed by the Fast R-CNN.
*YOLOv3 was not able to compute performances for several labels.

FIGURE 12
Floodwater depth estimation (Level 3) using canny edge detection and aspect ratio approach along with the calculated area (75862.0 based on pixel width and length dimension).

TABLE 4 Water levels associated with Aspect Ratios and flood severity and risk estimation.

Water level	Aspect ratio	Flood severity and risk
Level 1	>1.8	Low
Level 2	1.62–1.8	
Level 3	1.44–1.62	
Level 4	1.26–1.44	
Level 5	1.08–1.27	Moderate
Level 6	0.90–1.08	
Level 7	0.72–0.90	
Level 8	0.54–0.72	
Level 9	0.36–0.55	Severe
Level 10	0.18–0.36	
Level 11	<0.18	

intelligent architectures for AI-driven decision making and discovery. FAIS was developed as a national scale prototype for flood data analytics assessment, based on both historical and real-time flood warning and river level information, as well as crowdsourced processing of tweets. This included automated selection and analysis of large volumes of geotagged and relevant social media data, and recent advancement in data analytics algorithms. FAIS application intelligently identifies at-risk areas to flooding in real time and defines the geospatial footprint of a flood event using georeferenced tweets. The application also uses various image processing algorithms to detect labels and calculate flood depth and inundation areas. Overall, FAIS pipeline proved to be a robust and user-friendly AIoT prototype for both real-time and post-event analysis of flooding data at local scale that could help stakeholders for rapid assessment of flood situation and damages. Improved data collection and timely assessment of at-risk locations allow more efficient mutual aid in the operational theater for warnings and evacuations, and more effective search and rescue plans while enabling automatic dispatching of relief resources and evacuation plans.

The versatility of the AIoT approaches brings the intelligence and a new paradigm in flood computing and

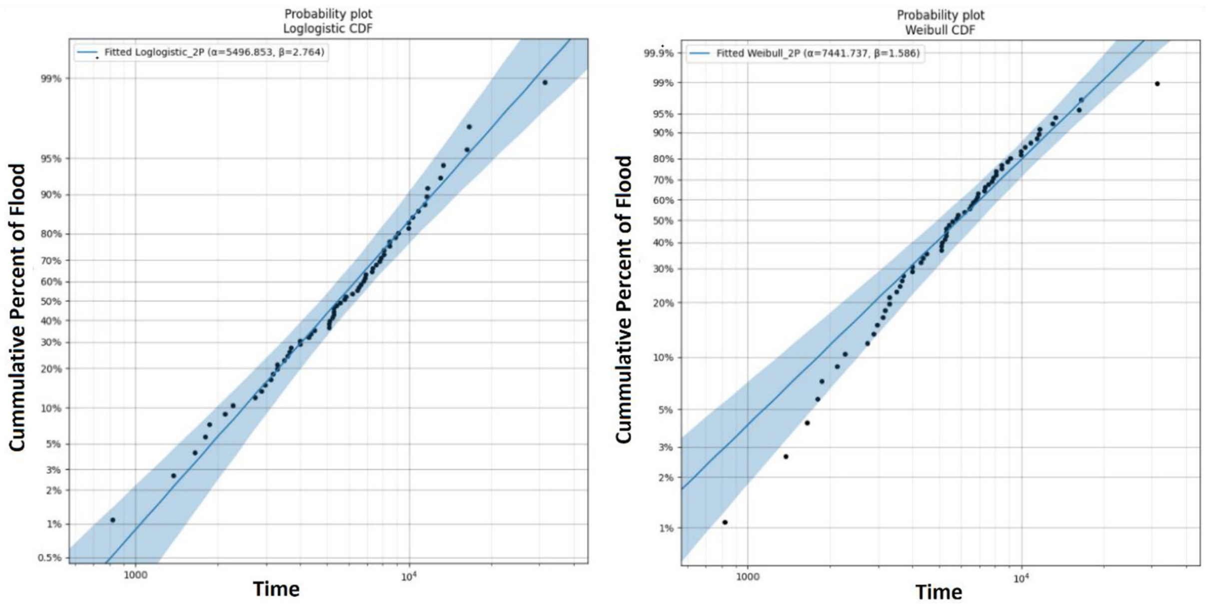


FIGURE 13
FFA for the USGS 02147500 using Loglogistic and Weibull distributions.

modeling. AIoT enables data provision as a service and provides controlled access to this data through APIs that ensures fast and accurate analysis of data across catchment scales. As the technologies reach different application sectors due to specific domains and data, the AIoT paradigm will evolve and expand in future heading to significant developments in terms of research and innovation. However, there are several challenges that need to be addresses to see AIoT full potential in flood research and applications. Given the huge number of devices, AIoT requires simultaneous connectivity and functional scalability that have become a concern for real-time application. Other deficiencies such as fault tolerance and data storage and complexities add more challenges to the AIoT application for real-time decision making. Moreover, the current data services do not support the reuse of data exploration processes and the data derived from analytics systems. Limitless analytics service combined with deep learning, data lake and data warehouse to process streaming data, and automated data lifecycle management are promising ways forward. Although, enabling AIoT interoperability for connecting data services and analytics requires solutions that must be realistic and scalable to multiple data platforms with the possibility to plug and play dynamically new analytics tools and AI algorithms.

More effort should be made to leverage AIoT application to better manage flood risk, make timely and effective decisions and forecast, and automate flood forecasting operations. For instance, improving the intelligence of AIoT systems to handle dynamic and complex environments using Edge computing could potentially handle hybrid models and reduce forecasting time by bring decentralized computing power as close as possible to the origin point of the data. In the future, empowered by rapidly developing AI technologies and big data analytics, many fast, smart, and safe AIoT applications are expected to deeply reshape flood related research and applications.

Software and data availability

FAIS-v4.00 application is publicly available at Clemson-IBM cloud service. FAIS python package is also available at the author GitHub account. The data and flood images used to perform this research is available upon request.

Data availability statement

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

Author contributions

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

Funding

This work was supported by the U.S. National Science Foundation (NSF) Directorate for Engineering under grant CBET 1901646.

Acknowledgments

The author acknowledges IBM company for providing free credits to deploy and sustain FAIS application.

Conflict of interest

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

Publisher's note

Author disclaimer

Any opinions, findings, and discussions expressed in this study are those of the authors and do not necessarily reflect the views of the NSF.

References

Barker, J. L. P., and Macleod, C. J. A. (2019). Development of a national-scale real-time Twitter data mining pipeline for social geodata on the potential impacts of flooding on communities. *Environ. Model. Softw.* 115, 213–227. doi: 10.1016/j.envsoft.2018.11.013

Bell, P., Evgeni, B., Justin, C., Ralf, G., Matti, P., Tyler, R., et al. (2021). *SciPy User Guide*. Available online at: https://docs.scipy.org/doc/scipy/reference/

Canny, J. (1986). "A computational approach to edge detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8*, 679–698.

Criminisi, A., Perez, P., and Toyama, K. (2003). "Object removal by exemplar-based inpainting," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conferenceon*, Vol. 2 (Wisconsin, DC: IEEE), II.

De Albuquerque, J. P., Herfort, B., Brenning, A., and Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *Int. J. Geogr. Information Sci.* 29, 667–689. doi: 10.1080/13658816.2014.996567

Demir, I., Yildirim, E., Sermet, Y., and Sit, M. A. (2018). FLOODSS: Iowa flood information system as a generalized flood cyberinfrastructure. *Int. J. River Basin Manag.* 16, 393–400. doi: 10.1080/15715124.2017.1411927

Donratanapat, N., Samadi, S., Vidal, M. J., and Sadeghi Tabas, S. (2020). A national-scale big data prototype for real-time flood emergency response and management. *Environ. Model. Softw.* 133, 104828. doi: 10.1016/j.envsoft.2020.104828

Feng, Y., and Sester, M. (2018). Extraction of pluvial flood relevant volunteered geographic information (VGI) by deep learning from user generated texts and photos. *ISPRS Int. J. Geo Information* 7, 39. doi: 10.3390/ijgi7020039

Fries, K.J. and Kerkez, B. (2018). Using Sensor Data to Dynamically Map Large-Scale Models to Site-Scale Forecasts: A Case Study Using the National Water Model. *Water Resources Research*, 54, 5636–5653.

Girshick, R. (2015). "Fast r-cnn" in *ICCV. Computer Vision and Pattern Recognition* (Santiago).

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask r-cnnR-CNN," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice: IEEE), 2961–2969.

Kingma, D. P., and Ba, J. (2014). "Adam: a method for stochastic optimization," in *Proceedings of the International Conference on Learning Representations* (San Diego, CA).

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, Y.-C., et al. (2016). "Ssd: single shot multibox detector," in *Computer Vision - ECCV 2016*, eds B. Leibe, J. Matas, N. Sebe, and M. Welling (Cham: Springer International Publishing), 21–37.

Pally, R., and Samadi, S. (2022). Application of image processing and convolutional neural networks for flood image classification and semantic segmentation. *Environ. Model. Softw.* 148, 105285. doi: 10.1016/j.envsoft.2021.105285

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 779–788. doi: 10.1109/CVPR.2016.91

Reid, M. (2021). *A Python Library for Reliability Engineering*. Available online at: https://reliability.readthedocs.io/en/latest/About%20the%20author.html

Samadi, S., and Pally, R. (2021). *The Convergence of IoT, Machine Learning, and Big Data for Advancing Flood Analytics*. European Geophysical Union. Oral Presentation. Available online at: https://meetingorganizer.copernicus.org/EGU21/EGU21-7782.html

Samadi, S., Sadeghi Tabas, S., Wilson, C., and Hitchcock, D. (2021). "Big data analytics for rainfall-runoff modeling: a case study of the cape fear catchment, North Carolina. Book chapter for the AGU/Wiley edited volume," in *Advances in Hydroinformatics: Artificial Intelligence and Optimization for Water Resources*, eds D. Solomatine and G. Corzo (IHE Delft for Water Education) (in press).

Singh, V. P. (1996). *Entropy-Based Parameter Estimation in Hydrology.* Dordrecht: Kluwer Academic Publishers.

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., and Gibson, C. (2017). Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. *J. Hydrol.* 548, 801–817. doi: 10.1016/j.jhydrol.2017.03.019

Sumioka, S. S., Kresch, D. L., and Kasnick, K. D. (1997). *Magnitude and Frequency of Floods in Washington*. USGS Water Investigations Report 97-4277. Washington, DC.

Vogel, R. M., Thomas, W. O., and McMahon, T. A. (1993). Flood flow frequency model selection in the Southwestern United States. *J. Wat. Resour. Plan. Manage. ASCE* 119, 353–366. doi: 10.1061/(ASCE)0733-9496(1993)119:3(353)

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: frontiersin.org/about/contact

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