



Toward Urban Water Security: Broadening the Use of Machine Learning Methods for Mitigating Urban Water Hazards

Melissa R. Allen-Dumas, Haowen Xu*, Kuldeep R. Kurte and Deeksha Rastogi

Computational Urban Sciences Group, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Climate Change Science Institute, Oak Ridge, TN, United States

OPEN ACCESS

Edited by:

Chaopeng Shen,
Pennsylvania State University (PSU),
United States

Reviewed by:

Andrea Cominola,
Technical University of Berlin,
Germany
Xi Chen,
University of Cincinnati, United States

*Correspondence:

Haowen Xu
xuh4@ornl.gov

Specialty section:

This article was submitted to
Water and Hydrocomplexity,
a section of the journal
Frontiers in Water

Received: 14 December 2020

Accepted: 15 May 2020

Published: 29 January 2021

Citation:

Allen-Dumas MR, Xu H, Kurte KR and
Rastogi D (2021) Toward Urban Water
Security: Broadening the Use of
Machine Learning Methods for
Mitigating Urban Water Hazards.
Front. Water 2:562304.
doi: 10.3389/frwa.2020.562304

Due to the complex interactions of human activity and the hydrological cycle, achieving urban water security requires comprehensive planning processes that address urban water hazards using a holistic approach. However, the effective implementation of such an approach requires the collection and curation of large amounts of disparate data, and reliable methods for modeling processes that may be co-evolutionary yet traditionally represented in non-integrable ways. In recent decades, many hydrological studies have utilized advanced machine learning and information technologies to approximate and predict physical processes, yet none have synthesized these methods into a comprehensive urban water security plan. In this paper, we review ways in which advanced machine learning techniques have been applied to specific aspects of the hydrological cycle and discuss their potential applications for addressing challenges in mitigating multiple water hazards over urban areas. We also describe a vision that integrates these machine learning applications into a comprehensive watershed-to-community planning workflow for smart-cities management of urban water resources.

Keywords: urban water security, hazard mitigation, machine learning, watershed modeling, integrated water resource management

1. INTRODUCTION

A recent United Nations report projects that 60% of the world's total population will live in cities by the year 2030 (U.N., 2018). This highly-urbanized population will face vulnerability to water-related hazards in many ways. For example, the combined effect of natural changes and human intervention on the landscape can lead to flooding, drought, and morphologic instabilities (e.g., stream erosion and instability, erosion, and sedimentation at structures) in and around urban areas, as well as deterioration of water quality, riverine ecology, and natural habitats (Crossman et al., 2013; Krajewski et al., 2016). Because of the accelerated pace of anthropogenic activity, hazard frequency, and intensity is exacerbated requiring immediate delivery of science-based solutions for mitigation, resilience, and adaptation that can be quickly deployed in any hazard-prone area. Mitigating these urban water hazards is challenging for watershed management and the urban planning community (Eriksson et al., 2015) due to the following hydro-complexities. First, these hazards exist in a variety of forms (e.g., floods, droughts, increased soil erosion, and water pollution) and are associated with multiple urban risks (e.g., property inundation and infrastructure failure,

water shortage, landslide, and eco-habitat deterioration) (Carson et al., 2018). Second, these urban water hazards may occur separately or in a multi-hazard chain (Kappes et al., 2012; Komendantova et al., 2014), in which the occurrence of one hazard (e.g., urban flooding) may trigger another hazard (e.g., bank erosion and landslide). Third, the occurrences of different urban water hazards are connected through the flow of the water and watershed processes over a range of spatial scales (Souchère et al., 2010; Santelmann et al., 2019), pressing the need for multiscale mitigation strategies that target hazard drivers at both watershed and urban neighborhood scale (Bertolotto et al., 2007; Xu et al., 2019b).

Given these challenges, a holistic approach to water security is articulated by Ait Kadi and Arriens (2012), as one that produces a world in which each community has access to enough water for social and economic development, and for ecosystems in and beyond those communities; and where those communities are protected from floods, droughts, landslides, erosion, and waterborne diseases (Carson et al., 2018; Aboelnga et al., 2019). Additionally, ensuring urban water security is a complex endeavor, as it involves dynamic processes and requires the interaction and participation of multiple planning actors (stakeholders, resource managers, and policy makers) to safeguard the integrity and security of urban water systems and assets in a continuous, physical, and legal manner. Subsequently, these actors must formulate policies and make investments using robust, adaptive, and accessible strategies that balance the socioeconomic and ecological benefits and urban sustainability with the cost of mitigation measures and management practices, and increase the resilience and preparedness of urban communities against extreme weather and natural disasters (Medema et al., 2014; Carson et al., 2018).

Fundamentally, these methods must have the capability of identifying and assessing the risk of multiple interconnected urban water hazards simultaneously (Kappes et al., 2012; Komendantova et al., 2014). Further, these methods must include system-based techniques for providing generalized predictions and acquiring unseen data in order to obtain reliable and accurate depictions of both current and future states of water resources in both urban areas and their associated watersheds. The projections and updates provided through these techniques must be easy to interpret and to understand, so that researchers, decision makers, and communities can readily obtain useful insights that support the planning of urban water resources, including the mitigation of existing hazards and the prevention of future hazards (Carson et al., 2018; Zaidi et al., 2018).

To fulfill these management needs, comprehensive disaster management frameworks are proposed to promote the collaborative planning and management of water, land, and related resources (Selin and Chevez, 1995; Emerson et al., 2012). These frameworks are developed to reduce the risk of multiple water hazards equitably without compromising the sustainability of vital ecosystems. Examples of these frameworks include Integrated Water Resources Management (IWRM), Adaptive Management (AM), and the Ecosystem Approach (EA) (Cardwell et al., 2009; Dörendahl, 2013; Palmer et al., 2013;

Carson et al., 2018). In general, these frameworks entail a series of planning processes that can be categorized into four major stages (Yu et al., 2018; Sun and Scanlon, 2019):

1. Long-term planning and mitigation
2. Early warning and prediction of hazards
3. Rapid response and rescue
4. Recovery and restoration.

Within the long-term planning and mitigation stage, we summarize here a list of common planning processes from several planning frameworks (Yoe and Orth, 1996; NRCS, 2003; USEPA, 2012), and we address machine learning (ML) methods for application to these processes throughout the paper. These steps are as follows:

1. Identification and assessment of multi-hazard risk in urban water systems.
2. Determination of the objectives of urban water planning and hazard mitigation.
3. Inventory of useful data resources that can define urban water hazards and risks, indicate the performances of existing urban water systems, and reflect the current state of the urban water system and the watershed to which it pertains.
4. Identification, evaluation, and selection of Best Management Practices (BMPs) from a variety of planning alternatives for water quality improvement, stormwater management, and erosion controls (NRCS, 2011; USEPA, 2018).
5. Evaluation of the performance and effectiveness of the implemented plan by examining information and monitoring data collected from pilot studies.
6. Identification, evaluation, and selection of proposed modifications for ongoing or existing plans and implementation schedules based on the future scenarios of urban water.

Despite the usefulness of these planning directives, the implementation of these processes is sophisticated and faces both methodological and technical challenges. Methodological challenges are associated with the long-term planning and mitigation processes and include: (a) assessing the multi-hazard risk and vulnerability of a municipal water system (Kappes et al., 2012; Jetten et al., 2014; Lambert, 2014), and (b) optimizing the selection of the BMPs from a variety of mitigation alternatives based on multiple criteria and objectives (FHWA, 2000). Technical challenges are associated with the implementation of multiple planning processes. One of the major technical challenges is related to the discovery and integration of a large volume of interdisciplinary data and simulation models (Adamala, 2017), which is essential for supporting the multi-hazard risk assessment in the long-term planning and mitigation process, as well as for informing rapid response and rescue during a hazardous event. These information resources can provide data-driven and model-driven insights for informing the current and future state of urban water systems and watersheds. Another major technical challenge is related to the accurate and timely prediction of hazardous events, which help facilitate early warning and prevention of hazard.

Conventionally, these challenges are approached using domain models and human justification of decision-makers, and therefore require computation- and labor-intensive efforts for coupling multiple models and investigating the underlying physical processes of different hazards. In recent decades, developments in advanced ML techniques has offered a more time efficient method for overcoming these challenges in an intelligent manner. Many review papers have enumerated ML and big data applications for enhancing various water resources management related applications and hydrological analysis (Adamala, 2017; Holzbecher et al., 2019) and for mitigating a specific water hazard, such as flooding (Mosavi et al., 2018), water pollution (Haghiabi et al., 2018), and erosion (Abdulkadir et al., 2019). In this paper, we explore and discuss benefits and potential opportunities of the ML applications for enhancing the mitigation of multiple urban water hazards. Herein, we review a selection of successful studies that apply various ML techniques and hybrid modeling techniques (i.e., the fusion of ML methods with process-based domain models) to overcome challenges encountered by different planning processes for integrated urban water management. Hybrid models are a mixture of inductive (data-driven) and deductive (process-based) approaches (Goldstein and Coco, 2015; Hajigholizadeh et al., 2018; Frame, 2019) and are referred to by Goldstein and Coco (2015) as the use of empiricisms built from ML in process-based models. Other researchers (e.g., Karpatne et al., 2016) approach hybrid modeling from the opposite direction—as “theory-guided data science,” in which data analysis, given sufficient grounding in physical principles, can represent causative relationships among parameters.

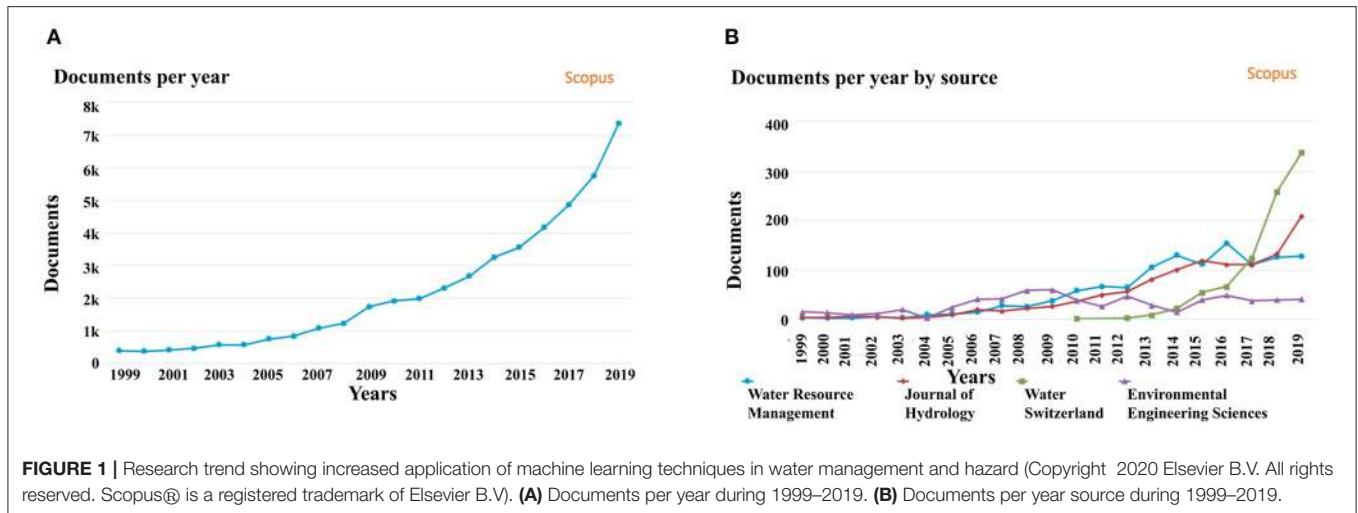
Additionally, we provide a vision for ways in which ML techniques can be used to facilitate different processes in the planning framework for the future. Different from previous review articles that focus on the machine learning application in the water management sector (Sun and Scanlon, 2019; Chen et al., 2020), we review innovative and application-ready machine learning solutions to facilitate urban water hazard mitigation from the practical aspect of addressing technical and methodological challenges in water resources and disaster management frameworks. The target audience of this paper includes watershed management authorities (WMAs), urban and regional planners, and research professionals in the water resources management sectors. To retrieve the relevant literature in this field that applies various ML techniques for urban water management, we conducted searches using tools such as Google scholar (<https://scholar.google.com>) and Scopus (<https://www.scopus.com>). **Figure 1** shows the result of the query: (“*Random Forest*” OR “*Artificial Intelligence*” OR “*ANN*” OR “*Support Vector Machine*” OR “*ANN*” OR “*Artificial Neural Network*” OR “*Neural Network*” OR “*SVM*” OR “*Machine Learning*”) AND (“*water management*” OR “*water resources management*” OR “*watershed management*” OR “*watershed planning*” OR “*urban water systems*” OR “*multi-hazard*” OR “*water hazard*” OR “*flood disaster*” OR “*water pollution*”) AND [EXCLUDE (PUBYEAR, 2020)]. We executed the query for years 1999–2019, and excluded year 2020. The above query retrieved a total of 46,145 documents from Scopus such that either article

title, list of keywords or abstract satisfies the query. It is clear from **Figure 1A** that there is a significant growth in ML based approaches for water related areas such as water management and urban water hazards. **Figure 1B** shows the top four scientific journals which receive research on ML application to water related areas. The graph in **Figure 1B** also confirms the increasing trends in the applications of ML techniques in water management and hazards.

Among the thousands of literature identifies from the Scopus, we select a handful of studies that are either published in recent years or are most relevant to and practical for improving specific processes and steps in the generic hazard mitigation stages and long-term water planning frameworks that are discussed early in the introduction section. We also consider the diversity and novelty of the machine learning techniques during the selection of studies for more detailed reviews and discussions. Based on the challenge and planning process targeted by these studies, we divide our review here into the following sections. Section 2 reviews the predictive data analytics powered by various ML techniques that help planners predict water-related hazards (e.g., flood, drought, water quality, and soil erosion and sediment transport). Multiple applications of hybrid modeling are also discussed in this section. Additionally, a subsection reviewing innovative combinations of ML and remote sensing technologies for disaster management is included, as remote sensing technologies are increasingly applied for improving the discovery and extraction of useful information and features (e.g., land use and land cover, flood inundation extent, and reservoir storage from satellite imagery) that are critical for early warning of hazards and rapid response and rescue during hazardous events (Hodgson et al., 2010). Section 3 presents the ML applications for the identification and assessment of water-related multi-hazard risks and vulnerability (e.g., building inundation, infrastructure failure, and economic loss) in urban water systems. In section 4, we review a few case studies that utilize ML algorithms to optimize the selection of urban BMPs, which can improve long-term planning and mitigation and recovery and restoration processes. Finally, in section 5, we present our vision for the application of next-generation ML techniques to efficient generation of mitigation strategies in response to urban water hazards. ML methods and their performance as applied to each issue are summarized in **Table 1**.

2. EARLY WARNING AND PREDICTION OF URBAN WATER HAZARDS

The capability to predict timely and accurate occurrence, intensity, and frequency of natural hazards is essential to every planning process that develops disaster preparedness and response to ensure public safety and mitigate unfavorable consequences associated with hazardous events (de Goyet et al., 2006). Traditionally, hydrological processes that contribute to water-related hazards have been analyzed using probabilistic modeling and physics based modeling approaches. The probabilistic approaches are devised to estimate the available stock over relatively short future time horizons (Philbrick and



Kitanidis, 1999). However, since the overall global climate is changing, rainfall data in any given area are non-stationary; thus the past does not necessarily predict the future, and the information given in recent data points may be more predictive than that of the data points from the more distant past (Tay and Cao, 2002). Limitations of probabilistic methods to produce realistic and specific results for water security planning have required the employment of physics based models for these predictions. Modeling hazardous events using physics based approaches requires the theoretical understanding of the atmospheric, land, and human processes and their interconnections; along with dynamics behind multiple hazards. However, many physics based models are designed to simulate pristine watersheds where hydrology is assumed to behave in a “pure” way, untainted by human interference (Joslin, 2016); therefore these physics based models are not suitable alone for predicting water-related hazards in urban watersheds. In addition, physics based models require large parallel machines and long periods of time for computation, neither of which may be available to water managers. Compared with the traditional modeling approaches, predictive data analytics powered by ML models can directly extract knowledge of natural disaster processes based on previous disaster occurrences and geo-environmental factors without prior knowledge (Pham et al., 2016; Rahmati et al., 2019). Unlike physics based modeling approaches, ML techniques can provide a bridge between physics based and probabilistic models because they can highlight patterns, trends, and regularities in data without requiring detailed understanding of the physical processes (Dibike and Solomatine, 2000; Rahmati et al., 2019), even when data are sparse, and with less complexity of construction and at relatively low computational cost (Mekanik et al., 2013).

Based on the scientific reasoning behind them, ML applications for predicting water-related parameters can be categorized either as inductive, whereby classifications are made based on statistical similarity in the hydrologic data directly; or deductive, whereby environmental variables (e.g., watershed characteristics) are analyzed as key drivers

of hydrology to create classification (Wagener et al., 2007, 2010; Olden et al., 2012; Auerbach et al., 2015). Because the inductive approach requires abundant hydrologic data (although all watersheds are ungauged at some point with unavailable or insufficient measurements; Joslin, 2016) many studies have favored the deductive approach, which classifies rivers and watersheds based on readily available environmental data that reflect the main drivers of hydrologic processes (Auerbach et al., 2015). Many researchers have utilized the deductive approach to relate stream condition (e.g., flow regimes, biodiversity, streamflow) with upstream watershed characteristics for different water resource management purposes (Poff and Allan, 1995; Snelder and Biggs, 2007; Carlisle et al., 2008; Reidy Liermann et al., 2012; Rice et al., 2015). The rationale for deductive classification methods, such as hydrologic regionalization, environmental regionalization, and environmental classification is to group river hydrological characteristics by spatial representation (e.g., river basin, region, catchment) based on environmental, hydrological, physical, and climatic similarity (Olden et al., 2012) to develop reliable class and empirical relationships between predictor and watershed characterizations.

2.1. Floods

Long term processes of change, including changes in climate, shifts in population, and increases in urbanization, will likely increase future urban flood risk changing the assumptions upon which flood risk analysis and management has long been based (Gangrade et al., 2019), and requiring new tools for risk assessment (Milly et al., 2008). In order to understand how to predict floods and to mitigate their effects on urban areas using new tools, it is important to understand the events that lead to flooding. The locations and processes that contribute to floods include atmospheric processes, catchment-level floods, river flooding, and accumulation of water in flood-prone urban areas (Merz et al., 2010). We discuss next the ML methods applied to each of these processes.

TABLE 1 | Machine learning methods discussed in each section.

Topic	Machine learning	Summary	References
Atmosphere (section 2.1.1)	ANN	Use of ANN has proved to be efficient in analyzing and representing complex, non-linear relationships between multiple atmospheric and hydrological parameters. Compared with the traditional modeling approach, ANNs have varying performance and are less time consuming.	Sahoo et al., 2017; Zaidi et al., 2018
	SVM	The performance of SVMs for forecasting regional rainfall varies across the geographic area. For non-stationary time series forecasting, DSVMs generalize better than the standard SVMs. The evaluation of these methods is conducted using both real data and simulated data.	Cao and Gu, 2002; Mohanty and Mohapatra, 2018
	Anomaly detection	Various anomaly detection algorithms have been proposed for detecting point anomalies to improve hydrological and climate data quality, as well as to mine potentially meaningful pattern anomalies within a given time series or spatio-temporal data.	Chandola et al., 2009a; Das and Parthasarathy, 2009; Sun et al., 2017
Catchment (section 2.1.2)	Evolutionary algorithms	Genetic programming approach performs better than the traditional hydrological models during scenarios where surface water movement and water losses are poorly understood.	Whigham and Crapper, 2001
	Cellular automata (CA)	The CA technique provides a versatile approach for modeling complex physical systems using a simplified 5-feature cell-based system. Compared with physically-based models, CA can dramatically reduce computational load, while providing a minimum required accuracy for rapid flood analysis in large-scale applications.	Guidolin et al., 2016
Rivers (section 2.1.3)	ANN	Compared with statistical models, ANNs tend to perform better for simulations containing non-linear patterns. Among popular ANNs, FFBPNN has been proven to have the best performances by many studies, and GRNN performs better than RBFNN in most cases. Thus, ANNs can serve as helpful tools for predicting river floods, as well as for mitigating missing flow data records.	Shamseldin, 2010; Badrzadeh et al., 2013; Tayyab et al., 2016
Urban Flood (section 2.1.4)	ANN, Bayesian linear regression, boosted decision tree regression, decision forest regression, linear regression	Noymanee et al. (2017) compared multiple ML techniques for predicting urban flood peak using a list of error metrics. The performance of different ML techniques varies for predicting urban flood stage at different timestamps. The study demonstrated that for predicting flood peaks, ANNs and Boosted trees performed best.	Noymanee et al., 2017
Indirect effects (section 2.1.5)	Reinforcement learning with agent-based models	Yang S. et al. (2019) used two studies to demonstrate the effect of Reinforcement learning with agent-based models in supporting the decisions of recovery actions after a flood disaster. The case that adopts the ML technique outperformed the other case and achieved a shorter recovery time.	Yang S. et al., 2019
Drought prediction (section 2.2.1)	ARF, BRT, Cubist	Model performance varies by drought type and across different regions. Park et al. (2016) demonstrated a case study showing that boosted random forests generally produced better results than the other two models for both arid and humid regions.	Park et al., 2016
	CART, random forests	Kuswanto and Naufal (2019) used a case study (based on both the TRMM and MERRA-2 datasets) to demonstrate that random forests perform well for prediction of droughts in East Nusa Tenggara, Indonesia.	Kuswanto and Naufal, 2019
	CART, BRT, random forests, MARS, FDA, SVM	Rahmati et al. (2019) demonstrated that the performance of different models varies when predicting the risk for different types of hazards. For example, the SVM model showed the highest accuracy for avalanches, while BRT demonstrated the best performance for flood hazards.	Rahmati et al., 2019
	ANN, SVR	For predicting the Standardized Precipitation Index (SPI) (in this case SPI 3, SPI 12, and SPI 24), a meteorological drought index, the wavelet boosting ANN (WBS-ANN) and wavelet boosting SVR (WBS-SVR) models produced better prediction results compared to the SVM.	Belayneh et al., 2016
	XGBoost	Zhang R. et al. (2019) demonstrated that the incorporation of non-linear and lag effects of predictors into the XGBoost method can significantly improve prediction accuracy of Standardized Precipitation Evapotranspiration Index (SPEI) and drought, providing a new modeling strategy for drought predictions based on multistation data.	Zhang R. et al., 2019

(Continued)

TABLE 1 | Continued

Topic	Machine learning	Summary	References
	W-QEISS	Zaniolo et al. (2018) applied a variable subset selection algorithm to improve the FRIDA's (FRamework for Index-based Drought Analysis) capability for automating the design of basin-customized drought indexes across different types of basins. The algorithm is based on a Wrapper for Quasi-Equally Informative Subset Selection (W-QEISS) and is capable of maximizing the wrapper accuracy, minimizing the number of selected variables, and optimizing relevance and redundancy of the subset.	Zaniolo et al., 2018
Water quality prediction (section 2.3)	ANN	Several studies have shown the ability of ANNs to simulate water quality variables and to produce simulated values for un-gauged locations.	Palani et al., 2008; Singh et al., 2009; Garcia-Alba et al., 2019
	XGBoost, RF	Lu and Ma (2020) evaluated the prediction performances of two novel hybrid decision tree-based ML models (based on XGBoost and RF) using the absolute percentage errors. The RF-based model has the best performance for predicting temperature, dissolved oxygen, and specific conductance, and the XGBoost-based model is best for predicting the pH value, turbidity, and fluorescent dissolved organic matter.	Lu and Ma, 2020
Soil erosion (section 2.4)	Random forests, M5P, RT, REPT	Bui et al. (2020) demonstrated the capability of hybrid algorithms to improve the predictive power of several standalone ML models. Among these models, the Hybrid BA-RT showed the best performance.	Bui et al., 2020
	Tree-based ML methods	Rahmati et al. (2017) found that many tree-based models (e.g., RF, RBF-SVM, BRT, and P-SVM) performed excellently both in the degree of fit and in performance for predicting gully headcuts. Hosseinalizadeh et al. (2019) proved that random forests were the most effective of these models for predicting and mapping gully headcuts in the future.	Rahmati et al., 2017; Hosseinalizadeh et al., 2019
	SVM	Mustafa et al. (2018) demonstrated that SVMs with different kernel functions have different performance levels for predicting soil erosion. They found that the polynomial kernel function had the highest performance, followed by linear and radial basis functions. Pourghasemi et al. (2017) explored multiple individual and ensemble ML methods (e.g., ABB, SVM, maximum entropy) for soil erosion prediction and concluded that the ANN-SVM ensemble performed best.	Pourghasemi et al., 2017; Mustafa et al., 2018
	ANN	Rahmati et al. (2017) demonstrated that the ANN could be applied to produce accurate and robust gully erosion susceptibility maps for decision-making and soil and water management practices, even though the random forests outperform ANN in many cases. Abdollahzadeh et al. (2011) demonstrated that ANN outperforms Multi Linear Regression (MLR) for predicting soil erosion.	Abdollahzadeh et al., 2011; Pourghasemi et al., 2017; Rahmati et al., 2017
Sediment transport (section 2.4.1)	ANN	The performance of ANN varies based on the training dataset (e.g., the time span and data quality) and the type of sediment for prediction. These ANN predictions are often tested against domain models and theories.	Tayfur, 2002; Lin and Montazeri Namin, 2005; Bhattacharya et al., 2007; Yang et al., 2009
	Adaptive-network-based fuzzy inference system (ANFIS)	Wieprecht et al. (2013) demonstrated that the ANFIS approach could be a useful alternative technique for predicting both bedload and total bed-material load. Lin and Montazeri Namin (2005) found that the method can be used to model both uniform and non-uniform suspended sediment. Bakhtyar et al. (2008) revealed that the ANFIS model provides higher accuracy and reliability for longshore sediment transport techniques than other methods, such as Fuzzy Inference System and CERC.	Lin and Montazeri Namin, 2005; Bakhtyar et al., 2008; Wieprecht et al., 2013
	M5 model trees	Goyal (2014) presented a comparative evaluation of the performance of M5 Model Tree and wavelet regression vs. ANN clearly demonstrating that M5 Model Tree and wavelet regression outperform ANN models in estimation of sediment yield. Onderka (2012) compared the M5 model tree with the conventional power-law rating curves, and concluded that the M5 model has better performance for modeling suspended sediments in a headwater catchment.	Onderka, 2012; Goyal, 2014

(Continued)

TABLE 1 | Continued

Topic	Machine learning	Summary	References
Sediment load (section 2.4.1)	Random forests	Francke et al. (2008) demonstrated that Random forests and quantile regression forests, compared with generalized linear models, are more accurate and favorable for reproducing sediment dynamics.	Francke et al., 2008; López-Tarazón et al., 2012
	Genetic algorithms (GA)	Yadav et al. (2019b) suggested that GA models outperform other models, such as ANN and SVM, for estimating suspended sediment yield. Altunkaynak (2009) found that GA models outperform the regression method for predicting sediment loads.	Altunkaynak, 2009; Yadav et al., 2019b
	Unsupervised techniques	These methods are self-organizing, and their results are often validated using domain models and knowledge. For example, Xu et al. (2019a) used the concept of geological landform regions to verify the clustering results of sedimentation potential from a self-organizing map.	Ahmed et al., 2018; Xu et al., 2019a
Urban infrastructure (section 2.4.1)	Random forests	Xu et al. (2019a) demonstrated decent performance of random forests for forecasting sedimentation risks at culverts by validating the results with field inspection.	Xu et al., 2019a
	ANFIS	Azamathulla et al. (2011, 2012) demonstrated that the ANFIS approach can give more satisfactory results for predicting both the scour depth at culvert outlets and sediment transport in clean sewers compared with other methods (regression equations and ANN).	Azamathulla et al., 2011, 2012
Flood management with RS (section 2.5.1)	ANN	Tsintikidis et al. (1997) used a shallow neural network with one hidden layer to estimate rainfall from a passive microwave radiometer SSM/I data. The network considered brightness temperature and associated polarization information as inputs and it output the rainfall rates.	Tsintikidis et al., 1997
	Random forests	Kühnlein et al. (2014) performed a precipitation estimate using random forests with satellite-derived information on cloud-top height, cloud-top temperature, cloud phase, and cloud water path retrieved from Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI). Feng et al. (2015) developed a random forest based approach to map accurately a flooded area using high-resolution (0.2 m) imagery obtained from Unmanned Aerial Vehicle (UAV) imagery.	Kühnlein et al., 2014; Feng et al., 2015
	K-NN	Shahabi et al. (2020) employed a ML ensemble method with four different k-nearest neighbor (kNN) algorithms for flood detection and susceptibility mapping using Sentinel-1 images to generate the flood inventory and SRTM DEM to obtain various flood-related conditioning factors.	Shahabi et al., 2020
	LSTM	A spatio-temporal sequence forecasting using Convolutional Long-Short Term Memory (ConvLSTM) for precipitation nowcasting. RADAR echo data in 2D from a ground-based RADAR was used in this study. ConvLSTM forecasted the echo data.	Shi et al., 2015
	CNN	Pan et al. (2019) used CNN to improve the precipitation estimates from NWP models. Based on the work by Hong et al. (2004), Hayatbini et al. (2019) proposed a CNN model to estimate precipitation using geostationary satellite data GOES-16. Jain et al. (2020) used water indices with CNN to detect flood water. Potnis et al. (2019) used a CNN based architecture called ERFNet to detect flooded urban regions from high resolution Worldview-2 satellite imagery. Jiang et al. (2020) proposed an approach to obtain waterlogging depth from video images using CNN.	Hong et al., 2004; Hayatbini et al., 2019; Pan et al., 2019; Potnis et al., 2019; Jain et al., 2020; Jiang et al., 2020
	Knowledge-based approaches	Kurte et al. (2017) used a semantics-driven framework to enable spatial relationships based semantic queries to detect flooded regions from satellite imagery and further extended the framework (Kurte et al., 2019) to accommodate temporal dimension that enabled spatio-temporal queries over flooded regions. In a similar approach, Potnis et al. (2018) developed a flood scene ontology (FSO) which formally defines complex classes such as <i>Accessible Residential Buildings</i> , to classify flooded regions in urban area from satellite imagery.	Kurte et al., 2017; Potnis et al., 2018; Kurte et al., 2019

(Continued)

TABLE 1 | Continued

Topic	Machine learning	Summary	References
Water quality monitoring with RS (section 2.5.2)	ANN	Dogan et al. (2009) used ANN to improve the accuracy of biological oxygen demand (BOD) estimation from RS imagery. Wu et al. (2014) used ANN for TSS turbidity estimations to analyze data measured with a hyperspectral spectroradiometer. Hafeez et al. (2019) compared various ML techniques for estimating water quality indicators from RS imagery and found that ANN worked well. Govedarica and Jakovljević (2019) found that the SVM algorithm worked better than ANN with Landsat 8 data and ANN worked better than SVM when Sentinel-2 data was used for water quality monitoring.	Dogan et al., 2009; Wu et al., 2014; Govedarica and Jakovljević, 2019; Hafeez et al., 2019
	SVM	Wang et al. (2011) used the support vector regression (SVR) method to retrieve various water quality estimators from SPOT-5 satellite data. Huo et al. (2014) used genetic algorithms combined with support vector machines (GA-SVM) to build an inversion model for eutrophic indicators such as Chl-a from Landsat ETM imagery.	Wang et al., 2011; Huo et al., 2014
Impervious surface detection with RS (section 2.5.3)	Random forests	Bian et al. (2019) used a random forest algorithm and time-series data from multiple satellites HJ-1A/B and GF-1/2 to estimate the changes in the impervious surface percentage over the years 2009–2017.	Bian et al., 2019
	PBL, PUL, SVM	Yao et al. (2017) adopted a one-class classification approach to detect impervious surfaces using high-resolution GF-1 satellite images, and found that Presence and Background Learning (PBL) and Positive Unlabeled Learning (PUL) outperformed SVM models.	Yao et al., 2017
	CNN	Zhang H. et al. (2019) used a deep CNN approach with data fusion from optical and SAR satellites WV-3, Sentinel-2, and Radarsat-2. Similar other works, Sun et al. (2019) (used 3D CNN with WV-3 and LiDAR), McGlinchy et al. (2019) (used UNet with WV-2), show increasing trends of using deep learning based approaches with multi-satellite data fusion.	McGlinchy et al., 2019; Sun et al., 2019; Zhang Z. et al., 2019
Multi-hazard assessment (section 3)	BRT, GAM, SVM	Rahmati et al. (2019) investigated and mapped multi-hazard exposure using a combination of ML models. They found that the different ML models differed in their accuracy in predicting the different hazards, but that the applied ML models were nevertheless useful and generalizable for multi-risk mapping.	Rahmati et al., 2019
	Random forests, RBF neural network	Chen et al. (2019) evaluated the risk of regional flood disaster in the Yangtze River Delta (YRD) region. They discovered that the level of urban flood disaster is closely related to rainfall, topography, economic development, land use, soil erosion, urban flood control investment, and disaster emergency response capability.	Chen et al., 2019
	Random forests, SOM	Xu et al. (2019a) showed that ML application can be used not only for multi-risk assessment and hazard prediction but also for exploring the complex and interconnected processes behind multiple hazards.	Xu et al., 2019a
	Random forests	Pourghasemi et al. (2020) developed the Sendai framework, which used random forests to produce a reasonable understanding of the factors controlling flood, forest fire, and landslide occurrence, and to produce a multi-hazard probability map for facilitating integrated and comprehensive watershed management and land use planning.	Pourghasemi et al., 2020
	LSTM	Yang T. et al. (2019) used long short-term memory units (LSTM) to improve the timing component of the amplitude of peak discharge for flood simulations generated by global hydrological models over different climate zones.	Yang T. et al., 2019
Best management practices (BMP)	GA/adaptive search	Hadka and Reed (2013) developed a high-performance adaptive search “Borg” algorithm, which was shown to be the most scalable and the best performing of five best performing multi-objective optimization algorithms applied to rainfall-runoff calibration, long-term groundwater monitoring, and risk-based water supply portfolio planning. Others applied GA-based optimization models to find solutions to water quality problems for several watersheds in the United States by connecting non-point pollution reduction models with economic components.	Srivastava et al., 2002; Hadka and Reed, 2013; Limbrunner et al., 2013; Reed and Kollat, 2013; Chen et al., 2015

2.1.1. Atmospheric Process Methods

One ML method that is used to capture the underlying relationship between independent and dependent variables in atmospheric processes is Artificial Neural Networks (ANNs). ANNs are interconnected networks comprising an input layer, some number of hidden layers, and an output layer. Each layer contains several processors, or nodes, referred to as artificial neurons. The neurons in each layer are connected to the neurons in the previous and next layers, and they transfer information from one layer to the next. Synaptic weights and biases, along with activation functions applied to the input layer, modulate the input signals sent from one layer to the next. The processed information is then sent as output to the connected neurons in the output layer (Zounemat-Kermani et al., 2020). The power of ANNs is their ability to learn functional relationships, with minimal empirical error, between these variables. Additionally, the use of activation functions with ANNs allows them to handle non-linear data effectively (Zaidi et al., 2018). In fact, many water related studies (e.g., Sahoo et al., 2017) using ANNs have shown that complex, reproducible, non-linear relationships exist among, for example, precipitation, temperature, streamflow, climate indices, irrigation demand, and groundwater levels.

Another ML method that has been used for predicting average rainfall is a classification algorithm known as Support Vector Machines (SVM) (e.g., Mohanty and Mohapatra, 2018). This method, developed by Vapnik (1995), is based on *Structural Risk Minimization*, which, rather than minimizing empirical error, as ANNs do, minimizes an upper bound of the generalization error ϵ . Dynamic Support Vector Machines (DSVMs), a modified version of the SVM, can be used to accommodate the structural changes in non-stationary rainfall data because it uses, instead of a static ϵ and static regularization constants, an exponentially decreasing ϵ , and exponentially increasing regularization constants (Cao and Gu, 2002) to allow room for analysis of changing patterns in the data.

The probabilities of hydrological extreme events such as floods and drought are modeled using different distributions from those that predict future average values. Traditionally, these events and their return periods are estimated with distributions associated with Extreme Value Theory (e.g., Kao and Ganguly, 2011). However, ML techniques for anomaly detection have begun to be applied to hydrological extremes problems. Anomaly detection is the identification of outliers in the data, or items that differ significantly from the overall trend of the data. Typically, anomalous data is related to issues such as measurement equipment failure or an extreme hydrological event. For example, Das and Parthasarathy (2009) used unsupervised spatio-temporal distance-based and neighborhood-based anomaly detection method with global climate data to identify extreme drought and heavy rainfall at specific locations. Characterization of short-term and long-term future extreme events have also been made with anomaly detection using trends found in historical time series. For these analyses, techniques such as kernel-based (rule-based classification), window-based (examination of the data in smaller “windows” in space or time), predictive, and segmentation (partitioning data into even smaller, possibly unequal, segments) algorithms are employed along with anomaly detection for

locating extremely low and extremely high temperature and precipitation events (Chandola et al., 2009b). In the case of the research by Sun et al. (2017), a density-based method was applied to anomaly detection in a hydrological time series. That is, the data were transformed to a piecewise linear representation through the important feature points of the data before mapping their slope, length, and mean to three-dimensional space for examination.

2.1.2. Catchment-Level Methods

Flood models at the catchment level analyze mainly issues of runoff generation and concentration leading to flood discharge. Because flood flow predictions are complex, non-linear, and not well-understood, ML may be required to *evolve* algorithms to derive characteristics of a particular flow. One way of evolving these algorithms is with the use of genetic programming, or genetic algorithms (GA), which produce, using routines imitating Darwin’s “natural selection,” algorithms directed to perform tasks defined by a set of training examples. Whigham and Crapper (2001) applied a type of genetic programming system to discover rainfall-runoff relationships for two meteorologically and topographically different catchments, one in Wales and one in Australia, and compared the results to those obtained with a traditional deterministic lumped parameter model. While both models did well when rainfall and runoff were correlated, the genetically programmed model performed better on the more poorly correlated data because it was allowed not to assume any underlying relationships, only to demonstrate its “fitness” to solve the problem.

Guidolin et al. (2016) used a two-dimensional cellular-automata-based model employing simple transition rules and a weight-based system to model catchment-level runoff. This diffusive-like method is designed to work with various general grids (rectangular, hexagonal, triangular) and with different neighborhood types (e.g., Moore or von Neumann). It also allows for model parallelization to increase its efficiency in large compute environments. To propagate a flood using this method, ratios of water to be transferred from a central cell to downstream neighbor cells are calculated using a weight-based system, with water volume transferred limited by Manning’s formula (Manning et al., 1890), and the critical flow equation. Water velocity and an adaptive time step are evaluated within a larger updated timestep. The results of the emergent behavior of this process shows good agreement with much more computationally intensive physical methods.

2.1.3. Machine Learning for Analyzing River Floods

Flood hazard in rivers can be characterized by the probability and intensity of large river flows and their consequent inundations, and it depends on the atmospheric and catchment processes preceding river flood generation (Merz et al., 2010). In fact, river floods are generally defined in hydrological terms by their water level or amount of discharge. Thus, Shamseldin (2010) explore the use of ANN for forecasting discharge from the Blue Nile river in Sudan. The type of neural network they chose was that of a multi-layer perceptron (MLP) feedforward network, a non-linear input–output model consisting of a network of

interconnected neurons, or computational units, linked together by connection pathways. The input layer is essentially a set of vectors of independent variable values, whereas the output layer is a set of possible dependent variable vectors of values. Between these two layers is a hidden layer containing an unknown number of neurons which are usually estimated by a trial-and-error procedure based on a mathematical non-linear transfer function (Shamseldin, 2010). Input variables in this case were weighted historical rainfall estimates, weighted seasonal rainfall estimates, and seasonal expectation of discharge; and the output variables were the river discharge values. Results showed strong correlation with observations for the river.

In addition to the multilayer perceptron ANN approach, other types of ANNs have been used to analyze river floods. For example, Tayyab et al. (2016) applied and compared three different types of ANNs to predict stream discharge for the Jinsha River Basin in China. The methods included feedforward back propagation neural networks (FFBPNN), generalized regression neural networks (GRNN), and radial basis function neural networks (RBFNN). The differences among these approaches lies in the hidden layer functions and activation functions that are applied to the problem. Badrzadeh et al. (2013) expanded on these ANN approaches by coupling wavelet (transforms that identify trends in the data normally not revealed by signal analysis approaches and also help to de-noise a dataset) multi-resolution analysis and adaptive neuro-fuzzy interface system (ANFIS) techniques (integration of neural networks and fuzzy logic) as preprocessing techniques to the ANN and show improved daily river flow forecasting over the use of ANNs alone, especially for long lead times. Mosavi et al. (2018) demonstrated the application of ANNs, neuro-fuzzy, SVM, and support vector regression (SVR) (SVM with regression only), in forecasting river floods and predicting the runoff hydrograph. The robustness of these techniques was evaluated and was found to be in good agreement with the observations.

2.1.4. Methods for Addressing Flood-Prone Urban Areas

Building resilience to natural disasters is one of the most pressing challenges for achieving sustainable urban development in flood-prone regions (Chang et al., 2019). River flooding in urban areas can cause high levels of damage, and while a relationship between hydrological characteristics and damaging floods may exist, knowing about an area's hydrological characteristics does not always indicate understanding of its vulnerability to damaging floods (Pielke, 2000). This understanding is imperative for hazard-mitigation planning for urban areas because these areas' responses to rainfall extremes tend to be faster than those for natural surfaces (Rodriguez et al., 2003). Thus, strategies for flood mitigation in these areas such as detention ponds, soakaways, permeable concrete, and green spaces, or upstream solutions such as river training and construction of dams and levees (Shamseldin, 2010) should be evaluated and implemented based on a thorough understanding of flood risks and responses of the area. For example, for predicting urban floods for the city of Pattani south of Thailand, Noymanee et al. (2017) examined the entire Pattani basin, which includes two dams for water

management: a diversion-type, Pattani Dam, and a hydropower plant, Bang Lang Dam. It is known that the most frequent floods are a result of overflow from flash flooding of the Pattani Dam rushing toward the city. The researchers acknowledge that a comprehensive approach to controlling floods in the area must include both structural and non-structural measures such as the development of improved technology for data management of the drainage network, and an increase in the sensors' frequency and extent of coverage. Thus, Noymanee et al. (2017) tested five different ML methods using open data pertaining to the area hydrology, the dam structures, the drainage network, and the technological components of the dams to explain the occurrence of extreme floods estimating dam water levels and cumulative precipitation amounts to forecast flood peaks in the urban area. The five methods tested included an ANN, Bayesian linear regression (statistical inference using Bayes' theorem), boosted decision tree regression and decision forest regression (both similar to random forest analysis discussed in section 2.2.1) and linear regression. Results showed the lowest error and highest correlation with the observations in the urban area from the Bayesian linear regression. This favorable result for that method may have occurred because it was informed by probability distributions drawn from prior data.

Often, in order to understand and manage risks of urban flooding beyond purely hydrological considerations, integration of decision support tools with predictive models is instructive. For example, one study (Rozos, 2019), combined a hydrological model, a demand management model called a network flow programming model (NFP), and an Feed Forward Neural Network (FFNN) to simulate a water supply system in Athens, Greece. The NFP optimizes and simulates the operation of a water supply system given hydrological inputs. FFNNs are the simplest type of ANN, whereby information moves in a forward direction from input nodes to the hidden layer to the output nodes (Mosavi et al., 2018) and they lend themselves to multi-model coupling. In this case, the NFP used synthetic data of a length capable of capturing the risk of each policy. Then the penalty functions of the NFP were selected to reflect the operating policies with different levels of risk acceptance. This process provided a large set of training data over a long period of time that was then used as input to the FFNN. This process allowed optimal decisions to be identified and made for the Athens system.

2.1.5. Predicting Indirect Flood Effects in Urban Areas

Indirect flood effects are those that cause damage to assets outside the flooded area. These assets can be physical, economic, social, or ecological in nature with impacts lasting for days, months, or even years after a large flooding event (Costello et al., 2019). In order to evaluate the extent of these effects, multi-agent-based simulations have been applied. Agent-based models simulate actions and interactions of autonomous agents, which can be individual actors or groups of actors, to assess the effects of these individual actions on the system as a whole. In one study (Yang S. et al., 2019), reinforcement learning, which rewards software agents for actions taken to maximize their cumulative reward, was used with the agent-based simulation for the optimization of post-disaster recovery for both individual companies and

supply chains for Tokyo, Japan. That study showed improved indirect damage estimation accuracy and mitigation potential over statistical methods and rough empirical models.

2.2. Drought

Drought is a prolonged period of precipitation deficit that may occur at varying spatiotemporal scales ranging from local to regional, lasting for weeks, months, multiple years, or even decades (Pendergrass et al., 2020; Hao et al., 2018). Drought may be exacerbated by extreme heat, soil moisture deficit, land atmosphere feedbacks, sea surface temperature anomalies, atmospheric circulation, and human activities such as land use and land cover changes and increased water demand (Cook et al., 2007; Dai, 2011; Kam et al., 2014). Droughts are high-impact weather hazards that affect agriculture, economy, ecosystem, water supply, and human lives (Hao et al., 2018). Over the past two decades, the total cost associated with drought is estimated to be billions of dollars (Huntingford et al., 2019). In a warming climate, the duration and intensity of drought is further projected to increase (Pagán et al., 2016; Pendergrass et al., 2020). Therefore, an advancement in the capability of timely prediction and development of early warning systems is crucial for drought risk management and strategic planning.

2.2.1. Advancement in the Use of Machine Learning Techniques for Drought Prediction

Drought is a complex weather hazard (Van Loon, 2015); therefore, a comprehensive understanding of the physical mechanisms that drive drought is essential to improving drought prediction (Huang et al., 2016). Numerous studies have been conducted to understand the intricate physical processes that lead to the extreme low moisture conditions of drought. Scientists have employed dynamical methods that involve climate and hydrological model simulations, statistical models using a suite of predictors and drought indices, as well as hybrid models for drought prediction (Fernández et al., 2009; Dutra et al., 2014; AghaKouchak, 2015; Mo and Lyon, 2015; Wood et al., 2015; Hao et al., 2017, 2018).

During the last decade, there has been an increase in the use of ML techniques to improve drought predictability (Hao et al., 2018). For instance, random forest ML algorithms have been increasingly used in drought prediction studies (Park et al., 2016; Kuswanto and Naufal, 2019; Rahmati et al., 2020). Random forests are extensions of decision tree analysis that start with classification trees—types of decision trees that can be grown together as a “forest” in a computational system. They provide highly accurate classification and characterization of complex predictor variable interactions while maintaining flexible analytical technique selection (Allen et al., 2018). Random forests also provide the capability to deal with the issue of overfitting and multicollinearity as compared to the traditional linear regression models (Konapala and Mishra, 2020). Park et al. (2016) employed random forests, boosted regression tree, and Cubist ML algorithms (rule-based model trees on which the terminal leaves contain linear regression models) for meteorological and agricultural drought monitoring using 16 remote sensing based drought factors over arid and

humid regions in the United States. Their findings suggest that among the three approaches, random forests provide the best performance for Standardized Precipitation Index (SPI) prediction. Similarly, Kuswanto and Naufal (2019) found the performance of random forests to be optimal when using SPI derived from Modern-Era Retrospective analysis for Research and Applications (MERRA-2) for drought prediction over the East Nusa Tenggara Province in Indonesia. A more recent study, Rahmati et al. (2020) compared the performance of six different ML techniques [classification and regression trees (CART), boosted regression trees (BRT), random forests, multivariate adaptive regression splines (MARS), flexible discriminant analysis (FDA), and SVM] for mapping agricultural drought hazard in the southeast region of Queensland, Australia. Similar to Park et al. (2016) and Kuswanto and Naufal (2019), they found that random forests had the best goodness-of-fit and predictive performance among the six models. Zaniolo et al. (2018) contributed to the FRIDA (FRamework for Index-based Drought Analysis) for the automatic design of basin-customized drought indexes across different types of basins by applying a ML-powered variable selection algorithm. The algorithm is based on a Wrapper for Quasi-Equally Informative Subset Selection (W-QEISS), which applies a multi-objective evolutionary algorithm to identify Pareto-efficient subsets of variables. This technique is able to maximize the wrapper accuracy, minimize the number of selected variables, and optimize relevance and redundancy of the subset. As a result, the framework is able to build an index that represents a surrogate of the drought conditions in a basin through the computation and combination of all the relevant available information regarding the water cycle in the system identified using the feature selection algorithm.

ANN ML techniques (see section 2.1.1) have also been used for drought forecasting (Mishra et al., 2007; Morid et al., 2007; Belayneh and Adamowski, 2012; Belayneh et al., 2014). Belayneh et al. (2016) coupled a wavelet transform data processing technique (see section 2.1.3), bootstrapping and boosting ensemble approaches with ANN and Support Vector Regression (SVR) (see section 2.1.1) for drought prediction in the Awash river basin of Ethiopia. Bootstrapping is a resampling technique with replacement that was used to create bootstrap ANN and SVR ensemble models to reduce model prediction uncertainty. Boosting techniques improve the performance of an algorithm by producing a series of models focusing on training cases that were not well predicted previously. The researchers found that the coupled models showed an improved performance and provided more robust SPI predictions as compared to either of ANN or SVR alone.

ANN models can be limited by model interpretability, local minima traps, and computational efficiency issues. Thus, alternatively, XGBoost has been gaining popularity due to its high execution speed and improved model performance as compared to other ML techniques such as SVM, ANN, and random forests (Fan et al., 2018; Shimoda et al., 2018; Zhang R. et al., 2019). XGBoost is an ensemble technique that implements a gradient boost decision tree algorithm to produce an ensemble of weak prediction models. Models are subsequently added to improve errors until an optimum performance is achieved. Zhang R.

et al. (2019) compared the performance of XGBoost with a traditional statistical model and an ANN model for Standardized Precipitation Evapotranspiration Index (SPEI) prediction with a lead time of 1–6 months for 32 weather stations in the Shaanxi Province of China. In their study, the XGBoost model showed the best performance for SPEI prediction, achieved highest user's and producer's accuracies and was much faster than the ANN model.

2.3. Water Quality

The deterioration of water quality in both groundwater and surface water has become a major concern causing negative impacts on human well-being, eco-systems, water supply, and infrastructure around the world (UN, 2012; Khan and See, 2016). According to United Nations (UN), more than 880 million people are living in water scarcity without adequate safe drinking water, and 2.6 billion people lack access to basic sanitation due to water shortage (UN, 2010, 2012). Effective management of water supply systems and watersheds often requires reliable and timely approaches for predicting water quality and forecasting future water quality trends (Wang et al., 2017; Bui et al., 2020). Based on established water quality standards (Nowell and Resek, 1994; EPA, 2012), water quality is often estimated using a combination of water quality parameters that reflect the physical, biological, or chemical characteristics of the air, watershed hydrology, soils, and sediment transported in the aquatic system (Hou et al., 2013; EPA, 2019). Developing accurate and timely prediction of water quality is a challenging effort. The traditional approaches utilize water quality models for analyzing and predicting water quality parameters. Most of these models consist of mathematical representations of physical mechanisms that determine (a) the fate, transport, and degradation of pollutants within a water body, and (b) the movement of pollutants from land-based sources to a water body (Refsgaard and Henriksen, 2004). Despite their usefulness for modeling specific scenarios, water quality models can only provide one line of evidence that serves as an imperfect approximation of reality (Kebede, 2009). This is because of process complexity of the water quality problems in that (1) there is a large number of interconnected multi-domain processes (e.g., physical transport, hydrological, chemical, and biological); and that (2) many underlying mechanisms that may affect water quality are still unknown. Complex water quality models often involve time-consuming and labor-intensive processes (Ahmed et al., 2019), rendering them costly and ineffective for supporting many time-critical water resources management tasks that have limited budgets. Compared with process-based (mechanistic) models, the newly emerging data-driven approaches for water quality predictions often rely on a large volume of water quality and hydrological data from various sources (Khan and See, 2016). Examples of these data sources include the United States Geological Survey (USGS) online resource—National Water Information System (NWIS) and the United States Environmental Protection Agency's (USEPA) STORET Data Warehouse (Beran and Piasecki, 2008). These analyses normally consider the combined effect of multiple water quality parameters, such as ammoniacal nitrogen (NH₃-N), suspended solid (SS), dissolved oxygen (DO), pH, and salinity. As many of these parameters are dynamic and affected by natural

watershed hydrology, their influences on water quality may vary across watersheds (EPA, 2019). In different watersheds, some parameters may have greater and more noticeable influences on water quality than others (Khan and See, 2016). In response to this challenge, the water quality index (WQI) has been proposed as a representation of several water quality variables simultaneously considered. However, calculating WQI using traditional approaches consumes time and is often filled with errors during derivations of sub-indices (Bui et al., 2020). To address these limitations and improve water quality analysis and prediction, researchers have applied many ML techniques (Khan and See, 2016; Ahmed et al., 2019; Bui et al., 2020), as well as developed a few hybrid approaches that combine various traditional methods with ML techniques (Taskaya-Temizel and Casey, 2005; Wang et al., 2017). We discuss the application of some of these approaches next.

Palani et al. (2008) and Singh et al. (2009) applied ANN models to predict river and coastal water quality in India and Singapore respectively. Each found that the ANN-computed values of water quality indicators were in close agreement with their respective measured values in the river water. García-Alba et al. (2019) developed an ANN model to estimate bathing water quality in estuaries and found that ANN models are able to estimate *Escherichia coli* concentrations comparable to those estimated by process-based models, and at much lower computational cost. In more recent studies, combinations of multiple ML and data analytic techniques applied to a problem are preferred to analysis with a single ML technique. For example, Lu and Ma (2020) proposed coupling two ML models to improve water quality prediction: XGBoost (section 2.2.1), and a random forest algorithm (section 2.2.1). They found that while the hybrid XGBoost model performed better for PH values, turbidity, and fluorescent dissolved organic matter predictions, and the random forest model performed better for temperature, dissolved oxygen, and specific conductance prediction; the combined performance of the two models was the best for optimizing the calculation of a water quality index. Barzegar et al. (2020) applied two standalone deep learning (DL) models, a convolutional neural network (CNN), an ANN with a convolutional activation function, and the long short-term memory (LSTM) model, which includes feedback in addition to feedforward networks, and a combined CNN–LSTM model to predict two water quality variables, dissolved oxygen (DO; mg/L), and chlorophyll-a (Chl-a; $\mu\text{g/L}$), in the Small Prespa Lake in Greece. Assessment of the model performance using statistical metrics, showed that LSTM outperformed the CNN model for DO prediction, but the standalone DL models yielded similar performances for Chl-a prediction. The combined CNN–LSTM model, however, outperformed the standalone models for predicting both DO and Chl-a. By coupling the LSTM and CNN models, both the low and high levels of water quality parameters were successfully captured, particularly for the DO concentrations (Barzegar et al., 2020). Similar successful approaches involving the coupling of multiple ML algorithms for the short-term prediction of water quality parameters include Li et al. (2018) and Lu and Ma (2020). Bui et al. (2020) applied four standalone algorithms [random forests and three variants: M5P (similar to Cubist, section 2.2.1),

random tree (RT), reduced error pruning tree (REPT)], and developed 12 algorithm combinations among these methods to predict water quality in northern Iran. They found fecal coliform concentrations to have the most effect and total solids to have the least effect on the predictions. Finally, Read et al. (2019) integrated theory with state-of-the-art ML techniques to improve predictions of water quality related parameters guided by physical laws. The study presented a use case for a Process-Guided Deep Learning (PGDL) hybrid modeling framework for predicting depth-specific lake water temperature, which serves as an important water quality parameter. The PGDL consisted of three primary components: a deep learning (many-layered neural network) model with temporal awareness (long short-term memory recurrence), theory-based feedback (model penalties for violating conservation of energy), and model pre-training to initialize the network with synthetic data (water temperature predictions from a process-based model) (Read et al., 2019). Through the use case the researchers demonstrated that the integration of scientific knowledge into deep learning tools shows promise for improving predictions of many important environmental variables.

2.4. Soil Erosion and Sediment Transport

Erosion and sedimentation are naturally occurring processes that include the detachment, transportation, and deposition of soil particles through the action of wind, water, and ice (NRCS, 2008). However, excessive soil erosion and sedimentation rates are results of anthropogenic activities (e.g., urbanization and agriculture) where soil surfaces are exposed and initially not revegetated (e.g., construction sites). Without proper mitigation, erosion and sedimentation in urban areas can cause a series of adverse impacts to the environment and urban areas (Guy, 1970; Hewett et al., 2018), which include water pollution, degradation of aquatic habitat, infrastructure damage (e.g., sediment blockage in urban waterways, storm sewer, and stream crossings, as well as silting of roadways, utility supply networks, and fences), increase in water-treatment costs, and stream bank instabilities (e.g., gullying and land-slides) (NRCS, 2008).

2.4.1. Machine Learning Techniques for Sediment Research

To tackle sediment-related problems, the predictions of sediment production and transport are required to inform urban planning and watershed management communities of the major source of sediment and erosion-prone areas. Conventionally, these predictions are addressed through a wide variety of erosion and sediment transport models (Merritt et al., 2003; Nearing et al., 2005). Despite the usefulness and maturity of these traditional approaches, the prediction of sediment-related parameters (e.g., soil losses, in-stream sediment load, and sediment delivery ratio) is still challenging because of the following model limitations: (a) running many physically-based erosion and sediment transport models are time- and resource-intensive, and requires the consideration of more physical processes in addition to the hydrological process making models are less applicable to sediment-related predictions in large watersheds and areas (Abaci and Papanicolaou, 2009); (b) most models are

designed to simulate a specific type of erosion (e.g., rill, gully, and stream bank erosion) and sediment transport (e.g., suspended load and bed load) (Wischmeier and Smith, 1978; Ganasri and Gowda, 2015), while sediment-related problems in urban areas and urban waterways often entail multiple types of erosions and sediment transport therefore requiring the integration of a variety of models; and (c) most erosion and sediment transport models do not cover sediment transport and deposition at man-made structures (Rowley, 2014) in urban areas. A comparative study conducted by Liang et al. (2019) showed that data-driven models can effectively inform and complement the simulations conducted with physics based models. Currently, there are many studies that utilize various ML methods to address various issues in sediment research. We summarize a list of example studies by their application areas and their applied ML methods:

1. Modeling sediment transport
 - (a) Artificial neural networks (Tayfur, 2002; Lin and Montazeri Namin, 2005; Bhattacharya et al., 2007; Yang et al., 2009),
 - (b) Adaptive-network-based fuzzy inference system: (Lin and Montazeri Namin, 2005; Bakhtyar et al., 2008; Wieprecht et al., 2013),
 - (c) M5 Model trees (Onderka, 2012; Goyal, 2014).
2. Predicting sediment load
 - (a) Random forests (Francke et al., 2008; López-Tarazón et al., 2012),
 - (b) Genetic algorithms (Altunkaynak, 2009; Yadav et al., 2019b),
 - (c) Unsupervised techniques (Ahmed et al., 2018; Xu et al., 2019a).
3. Predicting soil erosion
 - (a) Tree-based ML methods (e.g., random forest, gradient boosted regression tree, naïve Bayes tree, and tree ensemble models) (Rahmati et al., 2017; Hosseinalizadeh et al., 2019),
 - (b) Support vector machine (SVM) (Pourghasemi et al., 2017; Mustafa et al., 2018),
 - (c) Artificial neural networks (Abdollahzadeh et al., 2011; Pourghasemi et al., 2017; Rahmati et al., 2017).
4. Sediment-related impacts on urban infrastructure
 - (a) Random forest (Xu et al., 2019a),
 - (b) Adaptive-Network-based Fuzzy Inference System (ANFIS) (Azamathulla et al., 2011, 2012).

In general, erosion and sediment research is a broad subject that provides numerous opportunities for ML applications. By reviewing the above-mentioned example studies, we have summarized that (a) compared with traditional erosion and sedimentation transport models, ML methods are easier and cheaper (Cigizoglu, 2002; Tayfur and Guldal, 2006; Yadav et al., 2019a), and can be readily applied to solve complex sediment problems that entail human factors and multiple erosion and sediment transport processes (Xu et al., 2019a), (b) ML models that rely on field data generally produce better and more

reliable results than those obtained from experimental models (Kitsikoudis et al., 2014).

2.4.2. Hybrid Modeling Techniques for Sediment Research

In addition to its application to previously described hydrological studies, hybrid modeling has also been applied to sediment research (Merritt et al., 2003; Hajigholizadeh et al., 2018). Through the fusion of inductive data-driven models and deductive process-based models (Goldstein and Coco, 2015), hybrid models inherit the strengths of both the ML methods and physics-based models in a single model that has an increased performance in terms of speed (Babovic et al., 2001; Hall, 2004), accuracy (Krasnopolsky and Fox-Rabinovitz, 2005; Goldstein and Coco, 2015), and the capability of addressing soil-water problems with complex and multi-scale physical processes (Hajigholizadeh et al., 2018). An additional benefit of hybrid modeling is that ML models and data can be directly coupled to improve the calibration of process-based models (Knaapen and Hulscher, 2003; Ruessink, 2005; Mekonnen et al., 2012). Hajigholizadeh et al. (2018) summarized a table of hybrid modeling applications that integrate statistical models with process-based models in sediment research including:

- Modified Morgan, Morgan and Finney (MMM) (Morgan et al., 1984),
- Sediment river network model (SEDNET) (Prosser et al., 2001),
- Erosion Assessment Tool of MIKE BASIN & MILW (SEAGIS) (DHI, 2003),
- Automated Geospatial Watershed Assessment (AGWA) (Scott et al., 2002).

2.5. Application of Machine Learning to Remotely-Sensed Data for Water Hazard Prediction and Mitigation

Remotely-sensed (RS) data, due to its wide spatial coverage, provides a synoptic view of disaster affected areas. It is also frequently available during the disaster response phase providing a temporal overview of the disaster situation. Due to the recent advancements in satellite sensor technology, RS data is now available at various spatial resolutions (i.e., low, medium, and high) affording local, regional, and global coverage, and various spectral resolutions, from a few spectral bands in optical sensors to several hundreds of spectral bands in hyperspectral sensors. Additionally, advancements in the RS field have resulted in a continuous growth in Earth Observation (EO) data archives. Due to these characteristics, RS data is a potential data source for each stage during hydrological prevent planning and post-event countermeasures (Ge et al., 2020). Nevertheless, it is not always possible and is often dangerous to conduct ground surveys of disaster affected areas. Often the disaster destroys the transportation and communication facilities making ground-based survey impossible. In such time-critical situations, the proper selection of the sensor type, spatial resolution, and satellite revisit period is crucial, as pre-disaster and ancillary data can provide a wide coverage of the

disaster affected area (Ge et al., 2020). Despite these occasional limitations, various powerful approaches have been developed recently in the context of advanced ML and computer vision to exploit the wealth of information that can be found in RS data to address various urban water hazards related events (Kurte et al., 2017).

2.5.1. Flood Management

Over the last two decades, RS data have successfully contributed to various stages of flood management (Rahman and Di, 2017) such as flood risk assessment and flood emergency planning and management. Flood risk assessment requires the performance of flood hazard assessment, exposure risk assessment, and vulnerability assessment. As a part of the flood hazard assessment, RS data have been analyzed for flood forecasting and evaluation of flood inundation. As a part of flood emergency planning and management, RS data have been widely used in flood early warning systems, rescue and relief operations, post-flood damage assessment and policy making. Various recent approaches have used advanced ML techniques and RS during various stages of flood management.

Flood forecasting requires accurate estimation of rainfall. Although satellite RS has limited direct applicability to flood forecasting, it has been widely used for precipitation estimation, which is an important input for flood forecasting models. In the late 90s, Tsintikidis et al. (1997) used a shallow neural network with one hidden layer to estimate rainfall from a passive microwave radiometer SSM/I data. The network considered brightness temperature and associated polarization information as inputs and it output the rainfall rates. A random forest based ML algorithm was used to estimate the precipitation which used satellite-derived information on cloud-top height, cloud-top temperature, cloud phase, and cloud water path retrieved from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) (Kühnlein et al., 2014). Recently, Shi et al. (2015) proposed a spatio-temporal sequence forecasting approach using Convolutional Long-Short Term Memory (ConvLSTM) with RADAR echo data in 2D from a ground-based RADAR for precipitation nowcasting by forecasting the RADAR echo data. Pan et al. (2019) proposed a Convolutional Neural Network (CNN) based approach to improve the precipitation estimates from numerical weather prediction (NWP) models. The authors stated that the method outperformed reanalysis precipitation products as well as statistical downscaling (SD) products using linear regression, nearest neighbors, random forests, or fully connected deep neural networks. In an another recent work, Hayatbini et al. (2019) proposed a precipitation estimation framework using a fully convolutional neural network and the advanced baseline imager data from GOES-16, a multispectral geostationary satellite. Specifically, they proposed that the U-net CNN architecture could perform rain/no-rain classification using satellite imagery. The study was based on the earlier work of Hong et al. (2004) on precipitation estimation using remote sensing data and an ANN.

Flash flood susceptibility mapping is another important process in flood risk assessment. Recently, Costache et al. (2019) used a Digital Elevation Model (DEM) with 30 m spatial

resolution obtained from Shuttle Radar Topography Mission (SRTM), and which was developed using the technique called SAR interferometry, to derive seven flash-related conditioning factors such as slope angle, aspect, profile curvature, and other factors. In addition, the authors used aerial imagery from Google Earth to delineate the torrential areas along with the land use/cover data, CORINE, which was derived from Sentinel-2 and Landsat-8 RS images. K-nearest neighbors (kNN), K-Start (KS), and Analytical Hierarchy Process (AHP) algorithms were then applied to obtain the flash-flood susceptibility mapping. Thus, RS techniques played a crucial role in obtaining eight out of 10 flash-flood conditioning factors. In a similar work, Shahabi et al. (2020) used a ML ensemble method with four different k-nearest neighbor (kNN) algorithms for flood detection and susceptibility mapping. Authors used Sentinel-1 images to generate the flood inventory and SRTM DEM to obtain various flood-related conditioning factors. These two works show that ML ensemble methods are gaining traction in flood susceptibility mapping.

Mapping of flooded areas is important to performing damage assessment, deploying rescue and relief operations and developing policies. An example of applying RS and ML to this undertaking is Feng et al. (2015), who developed a random forest based approach to map accurately a flooded area using high-resolution (0.2 m) imagery obtained from Unmanned Aerial Vehicle (UAV) imagery. The data were obtained for Yuyao City of Zhejiang Province in Eastern China during the flooding that occurred due to the extreme rainfall event on October 7, 2013. Additionally, Jain et al. (2020) developed a hybrid approach to combine the strength of the traditional water indices from RS imagery and generalization capability of Convolutional Neural Networks (CNN). The authors proposed a new water index which minimized cloud interference in the RS image and used it with a pre-trained VGG-16 model (Simonyan and Zisserman, 2014) and a transfer learning based approach to re-train the model for a new task of flood water detection. In a similar work, Potnis et al. (2019) used an Encoder-Decoder Neural Network based on the Efficient Residual Factorized Convnet (ERFNet) architecture for multi-class segmentation of urban floods satellite imagery from WorldView-2 of floods in Srinagar, India during September 2014. Recently, Jiang et al. (2020) proposed an approach to obtain waterlogging depth from video images using CNN. The approach generated synthetic images from the set of images of reference objects and flood surface, which was further used to train the CNN model to obtain the waterlogging depth. This method can also be employed to obtain waterlogging depth from the images taken of the flooded area using recent drone-based video surveillance. Cervone et al. (2017) added to these techniques a methodology to fuse social media data with the RS data during a flood situation to improve the flood mapping capability.

Recently, a few approaches to model the semantics in RS images were proposed for flood detection and mapping. Kurte et al. (2017) proposed a semantics enabled framework to model the spatial relationships among various regions in the RS images to enable spatial-relationships-based queries such as *Retrieve all images in the ALI repository having Built Up region externally connected to the Stagnated Flood Water*. Later this work was extended to accommodate the temporal aspect to enable the

spatio-temporal semantic queries such as *Show road segments which were completely submerged during 9th September 2014 to 22nd September 2014* (Kurte et al., 2019). In a similar semantics based approach, Potnis et al. (2018) developed a flood scene ontology (FSO) which formally defines complex classes such as *Flooded_Residential_Buildings*, *Accessible_Residential_Buildings*, *Operational_Roads*. After detecting various objects in the RS imagery using any supervised classification approach, the ontology can be used to infer complex classes which are very important for flood mapping.

2.5.2. Water Quality Monitoring

RS data has been used over the past 50 years to monitor water quality. For instance, RS data can be used to measure water turbidity, or lack of transparency, which is a good measure of the water quality. Clear water shows high absorptivity in the infrared and near-infrared wavelength regions. It also shows some reflectivity in the visible regions. Reflectivity in this application can reveal variations in water quality due to salinity, temperature, and turbidity. In the past decade, much research has been published in which remote sensing and ML approaches are used to estimate additional water quality parameters. For example, Dogan et al. (2009) explored the non-linear capability of ANN to improve the accuracy of biological oxygen demand (BOD) estimation. Wu et al. (2014) compared multiple regression (MR) with ANN for total suspended solid (TSS) turbidity estimations using data measured with a hyperspectral spectroradiometer and found that the non-linear transformation function of ANN performed better than MR. Wang et al. (2011) used the support vector regression (SVR) method to retrieve various water quality estimators from SPOT-5 satellite data. SVRs showed potential in solving problems with small sample size, non-linearity, or high dimension (Vapnik, 1995). Huo et al. (2014) stated that the lakes near urban areas or inside urban areas are becoming eutrophied or even hypereutrophied due to excessive urbanization and a fast growing economy. The authors used genetic algorithms combined with support vector machines (GA-SVM) to build an inversion model for eutrophic indicators such as Chl-a from Landsat ETM imagery. They showed that the GA-SVM based method had better prediction accuracy than the traditional statistical regression methods and ANN based approaches. According to Sharaf El Din et al. (2017), modeling water quality using satellite data is a complex problem, and conventional regression-based approaches can not perform well while modeling such complex relationships between water quality and RS data. The authors claimed that the proposed Landsat8-based-BPNN—back propagation neural network—to estimate water quality (both optical and non-optical) worked better than SVM-based methods. Moreover, the authors mentioned that, compared to the BPNN-based methods, the SVM-based methods could produce very different results due to differences in parameter selections, kernel-selection, high algorithmic complexity, and extensive memory requirement. The developed model showed $R^2 > 0.9$ for the water quality indicators such turbidity, total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD), and dissolved oxygen (DO). Recently, Hafeez et al. (2019) compared

several ML techniques including artificial neural networks, random forests, cubist regression, and support vector regression for estimating the concentrations of suspended solids (SS), Chl-a, and turbidity using Landsat data. The results showed that the ANN-based model achieved the highest accuracy in estimating the above mentioned water quality indicators. In another recent study, Govedarica and Jakovljević (2019) used 4-years of time-series data of *in-situ* monitoring of surface water bodies for the calibration and validation of a water quality estimation based on SVM and ANN algorithms using Landsat 8 data. The work also compared the estimations based on Landsat 8 with the Sentinel-2 data and found that, due to higher spatial and spectral resolution, Sentinel-2 data is a better alternative for water quality monitoring. Interestingly, the results showed that SVM produced more accurate results than ANN when used with Landsat data, whereas ANN provided better estimation accuracy for turbidity and TSS than SVM, and lower accuracy for TN and TP than SVM when used with Sentinel-2 data. Finally, Wang et al. (2017) conducted a study that combined a ML algorithm and remote sensing spectral indices [difference index (DI), ratio index (RI), and normalized difference index (NDI)] through fractional derivatives methods and in turn establishes a model for estimating and assessing the water quality index (WQI) (2.3). For this study, the WQI was calculated using sensitive wave bands and a spectral index of hyperspectral data, and particle swarm optimization (Kennedy and Eberhart, 1995; Shi and Eberhart, 1998)—support vector regression models (PSO-SVR), which deploy a population of candidate solutions over the SVR search space. Through comparisons of the predictive effects of the 22 water quality index estimations determined by the PSO-SVR, Wang et al. (2017) demonstrated that the model based on RI, DI, and NDI values of the 1.6 order was better performing than the others for predicting the water quality index of the semi-arid area of central Asia [R^2 (0.92), RMSE = 58.4, RPD (2.81) and a slope of curve fitting of 0.97].

2.5.3. Impervious Surface Detection

Urban impervious surfaces such as roads, driveways, sidewalks, and parking lots prevent water from infiltrating into soil, which has impacts on urban hydrology, groundwater, and water quality. Impervious surfaces facilitate pollutant's movements to nearby water bodies during heavy rain and urban flooding (Hall and Hossain, 2020). In the context of ML, identifying impervious surfaces from RS data is fundamentally a classification approach. However, many index-based approaches for sighting impervious surfaces using RS (e.g., Weng, 2012) focus on the developments in this area that use ML algorithms. Recently, Yao et al. (2017) adopted a one-class classification approach to detect impervious surfaces using high-resolution GF-1 satellite images, and found that Presence and Background Learning (PBL) and Positive Unlabeled Learning (PUL) outperformed SVM models in detecting impervious surfaces. Miao et al. (2019) also used a one class classification technique and Landsat-8 imagery for impervious surface classification. In a similar study, Bian et al. (2019) used a random forest algorithm and time-series data from multiple satellites HJ-1A/B and GF-1/2 to estimate the changes in the impervious surface percentage over the years 2009–2017.

Lin et al. (2019) addressed the challenges in detecting impervious surfaces due to the diversity of land use and shadow effects in high-resolution satellite imagery using a dictionary sparse representation classification and data fusion approach with WV-2, GeoEye-1, TerraSAR-X, and LiDAR. Zhang H. et al. (2019) addressed similar issues by using a deep CNN approach with data fusion from optical and SAR satellites WV-3, Sentinel-2, and Radarsat-2. Similar other works, Sun et al. (2019) (used 3D CNN with WV-3 and LiDAR), McGlinchy et al. (2019) (used UNet with WV-2), show increasing trends of using deep learning based approaches with multi-satellite data fusion.

3. IDENTIFICATION AND ASSESSMENT OF MULTI-HAZARD RISK

Multi-hazard identification and compound risk assessment inform effective planning activities and strategies (FEMA, 2015), and help water managers prioritize attention, investment, and recourse (Dickson-Anderson et al., 2016) to target the most urgent and the highest impact risks. Risk is defined as a combination of hazard, exposure, and vulnerability (Garrick and Hall, 2014). Because exposure in urban areas is relatively high due to the high density of population and man-made structures (Hoekstra et al., 2018), cities without proper preparedness and adaptation strategies are vulnerable to a wide variety of urban water hazards (Shaw et al., 2016; Eldho et al., 2018; Hoekstra et al., 2018; Gangrade et al., 2019; Rahmasary et al., 2019) that are often causally linked to further hazards. Additionally, coincidental hazards may occur, resulting in a compounding effect overwhelming the ability of local or national governments to respond (Liu and Huang, 2014). For example, a specific urban water hazard such as flooding can lead to multiple risks (Dai et al., 2017; Cook et al., 2019) that include inundation of building structures, damage to infrastructure, and/or the spread of water-borne diseases (Gangrade et al., 2018; Pereira, 2018). Consequently, multi-hazard risk assessment techniques must be conducted in the urban water management sector in a manner that considers the combined effects and interactive reactions of multiple urban water hazards in urban areas (Garcia-Aristizabal and Marzocchi, 2013; Gruber and Mergili, 2013; FEMA, 2015; Karlsson et al., 2017).

Despite its usefulness for hazard mitigation planning, multi-hazard risk assessment has been under-emphasized in natural disaster management and planning (Rahmati et al., 2019) due to the difficulty of analyzing the risk for more than one hazard in the same area, and of analyzing their interaction. In the past, studies have focused primarily on forecasting and controlling hazards, and their physical processes (Kalantari et al., 2019) in natural areas, without considering the social and economic impacts of these hazards in urban areas (e.g., hazard effects on buildings, infrastructures, and agriculture). Previous studies, which intended to analyze hazard risk and social vulnerabilities, only analyzed the risks of single hazards separately (Bühler et al., 2013; Statham et al., 2017) using physical or statistical models [e.g., flood impact using the HEC-FIA model (Lehman and Light, 2016) or economic damage to

fisheries caused by surface water pollution using AQUATOX model (Park et al., 2008)]. In general, most past studies do not consider the multi-hazard chain (hazard interaction) and the combined risk of coupled hazard events (Garcia-Aristizabal and Marzocchi, 2013; Rahmati et al., 2019). Although a few studies (Freeman and Warner, 2001; Newman et al., 2017) analyze the components of different types of vulnerability and risk by evaluating physical, social, and economic consequences of a chain of urban hazards, developing a systematic approach for multi-hazard risk assessment using conventional modeling methods faces multiple challenges. These challenges are primarily associated with (a) integrating multiple physical or statistical models and domain-data that only target single hazards to simulate a multi-hazard chain and predict the combined effect of multiple urban water hazards, and (b) in-depth understanding of hazards, including interconnections between different hazards, and dynamics behind multiple hazards. In the presence of hydro-complexities, many underlying mechanisms of urban water hazards remain unknown. Therefore, conventional methods based on physical modeling alone may not be the best way to assess multi-hazard risk in urban water systems.

In recent years, advanced ML methods have been used to develop innovative multi-hazard risk assessment frameworks and workflows, which are able to address the challenges associated with conventional risk assessment techniques. The feasibility of applying ML to multi-hazard risk assessment is shown by the following: (a) ML is a subfield of artificial intelligence and data-driven analysis where ML models can easily identify trends, patterns, and empirical relationships in a large volume of data without considering detailed physical processes behind a phenomenon, such as the interactive reactions between multiple water hazards (Dibike and Solomatine, 2000; Rahmati et al., 2019), and (b) ML models are capable of handling data that are multi-dimensional and multi-domain (Anzai, 2012). In this section, we review several ML workflows and applications that are designed to support the analysis of multi-hazard risk for mitigating water-related hazards.

For example, Rahmati et al. (2019) investigated and mapped multi-hazard exposure using several ML models including BRT (Boosted Regression Trees), GAM (Generalized Additive Model, a regression which can include linear or non-linear predictor variables and predicted values potentially following any of a variety of probability distribution functions), and SVM (Support Vector Machines), and they evaluated the performance of these ML models using threshold-dependent and threshold-independent methods. The study consists of several steps: (1) selection of predictive factors for modeling multiple hazards (e.g., flood, landslide, soil erosion, and debris flow), (2) creation of Multi-Hazard Inventory using records from road organization and the regional water company (RWC) to document the occurrence of various hazards, (3) application of ML models to predict and map the exposure of multiple hazards, and (4) evaluation of the accuracy of these models. The results of this study indicate that (a) different ML models differed in their accuracy of predicting the different hazards (Rahmati et al., 2019), and (b) the applied ML models are useful and generalizable for multi-risk mapping around the world.

Another example of a multi-hazard multi-model approach is Chen et al. (2019), in which the researchers evaluate the risk of regional flood disaster in the Yangtze River Delta (YRD) region. Based on the driving force, pressure, state, impact, and response (DPSIR) conceptual framework, the study first applies a random forest algorithm to screen important indices of flood risk. They then construct a radial basis function (RBF) neural network to evaluate the flood risk level. In this study, the radial basis function is the activation function for the ANN. The study approaches the urban flood risk assessment as a multi-classification problem using ML methods and indicates that only a few of the previous studies use ML theory to assess the urban flood disaster risks that are complex and associated with multiple sources and contributing factors. The study concludes that the level of urban flood disaster is closely related to rainfall, topography, economic development, land use, soil erosion, urban flood control investment, and disaster emergency response capability, shedding light on effective regulation measures for improving flood prevention in urban environments.

3.1. Exploration of Complex and Interconnected Hazards and Risks

To explore complex and interconnected hazards and risks, Xu et al. (2019a) present a visual analytics framework that combines various types of ML applications (e.g., feature selection, classification, and multivariate clustering analysis) with different geo-visualization techniques to analyze multi-hazard risk at culverts due to flooding and sedimentation. ML models applied in this study include the classification schemes, random forests and Self Organizing Maps (SOM), and are used for exploratory data analysis, aiming to improve the understanding of the factors and interconnected hazards (e.g., flooding, excessive erosion, and sediment transport in rivers) that contribute to the sedimentation and flood over-topping of culverts (transportation infrastructure). The results of the study show that ML application can be used not only for multi-risk assessment and hazard prediction but also for exploring the complex and interconnected processes behind multiple hazards. Additionally, the same framework can be readily extended to analyze multiple hazards at other hydraulic structures, such as bridges and weirs. Pourghasemi et al. (2020) presented a ML workflow, debuted as the Sendai framework, for assessing and mapping multi-hazard risk susceptibility, with an overall objective of reducing hazard risk and increasing sustainable development in urban areas. The workflow entails three main steps: (1) data preparation for obtaining the location of various hazards (floods, forest fires, and landslides), (2) recognition of the most important factors contributing to the occurrence of different hazards using the Boruta algorithm (a wrapper around random forest classification that iteratively removes irrelevant features from the data), and (3) construction of multi-hazard susceptibility maps along with validation processes using the random forest model and the preparation of a Multi-hazard Probability Index (MHPI) for the study area. The significance of the Sendai framework is that it (a) creates a reasonable understanding of the factors controlling flood and forest fire through ML-powered variable ranking

and landslide occurrence, and (b) produces a multi-hazard probability map for facilitating integrated and comprehensive watershed management and land use planning.

3.2. Hybrid Modeling for Multi-Hazard Risk Assessment

A few researchers have applied hybrid models to water-related multi-hazard risk assessment. For example, Yang T. et al. (2019) used long short-term memory units (LSTM) to improve the timing component of the amplitude of peak discharge for flood simulations produced with global hydrological models over different climate zones. Hajjigholizadeh et al. (2018) used hybrid models for predicting and assessing water erosion vulnerability and risks, as well as for the optimization of management strategies for agricultural or soil and water conservation practices. Application of hybrid models to these multi-hazard hydrological risks is still emerging within the domain, but the utility of this approach continues to be demonstrated across a variety of hydrological applications.

4. SELECTION OF BEST MANAGEMENT PRACTICES

The proper selection and placement of Best Management Practices (BMPs) is a critical planning process that helps many watershed and urban planning communities effectively mitigate water-related hazards and manage urban water resources (e.g., stormwater management, water pollution reduction, and erosion controls) (Cheng et al., 2006; NRCS, 2011; USEPA, 2018). These BMPs are carefully selected from a pool of planning and mitigation alternatives that exists in various forms. Based on their spatial scales, these alternatives can be categorized as either localized alternatives, which are city-scale practices for protecting the municipal water supply and infrastructure through structural actions and non-structural actions, and watershed alternatives, which represent the management of land cover and land-use at the watershed scale (Carson et al., 2018). The selection of BMPs is a complex multi-objective optimization problem that requires the consideration of multiple planning objectives and criteria, which aim to maximize the environmental and social benefits for multiple urban communities, while minimizing the economic cost for the implementation of these management practices (Maringanti et al., 2008; Rodriguez et al., 2011). The development and advancement in GA (section 2.4.1) have provided watershed management communities with a method for solving complicated optimization problems that are associated with the selection of BMPs. GA are capable of handling complex and irregular solution spaces when searching for a global optimum (Chambers, 2000; Rodriguez et al., 2011) in a multiobjective optimization. Multiobjective optimization has been defined as “vector optimization” (Cohon and Marks, 1975) for which the objective function is a vector containing scalar objectives subject to a set of constraints, and for which Pareto optimal solutions show the best performance. Reed et al. (2013) evaluated a variety of multiobjective optimization GA as applied to rainfall-runoff calibration, long-term groundwater

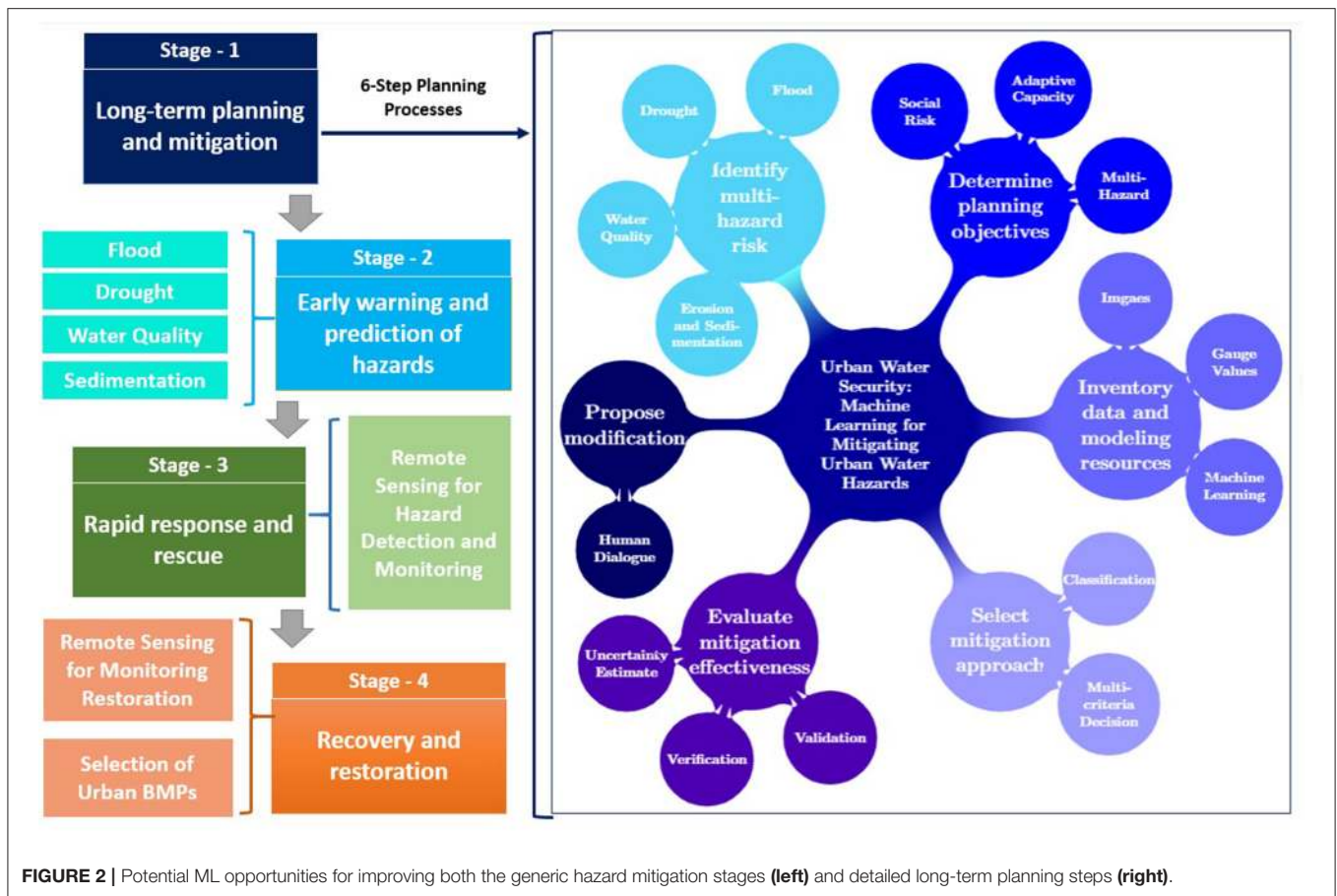
monitoring, and risk-based water supply portfolio planning. They found five best performing algorithms, of which their high-performance adaptive search Borg algorithm (Hadka and Reed, 2013) was the most scalable and the best performing, and has shown particular stakeholder usefulness in its incorporation into a visual and interactive decision support framework (Reed and Kollat, 2013).

In the water quality management sector, several studies applied GA-based optimization models to find optimal solutions to water quality problems for several watersheds in the United States by connecting non-point pollution reduction models with economic components (Srivastava et al., 2002; Chen et al., 2015). In the stormwater management sector, Limbrunner et al. (2013) applied classic optimization techniques to stormwater and non-point source pollution management at the watershed scale, and compared their effectiveness for finding optimal solutions to that of genetic algorithms, and linear and dynamic programming. Dynamic programming proved to find the most efficient solution to the sediment-management-optimization problem.

In addition to the optimization of planning alternatives, ML methods can enable selection for optimal management practices (Savic, 2019). AI-driven applications are envisioned to learn from the human decision-making process, during which best management practices are selected by planners and watershed managers based on their past experiences.

5. VISION: NEW APPLICATIONS OF MACHINE LEARNING TO URBAN WATER SECURITY

In order to ensure high-quality and timely water availability in the right quantities for urban areas, water resources must be managed well. In order for water resources to be managed well, a planning system leading to actions that promote sustainability and urban water security must be in place at the municipal level. We have shown that ML can help with this system as it applies to every stage of disaster management and planning, as outlined sequentially on the left hand side of **Figure 2** and shown as an interconnected and cyclical process on the right side. That is, we have outlined a variety of ML applications for facilitating the individual disaster management stages and planning processes. For long-term planning and mitigation, we have presented studies that use ML methods to identify and assesses multi-hazard risks and vulnerability in urban water systems, taking into account socio-economic factors and the multi-hazard chain. We have also discussed how ML can help optimize the selection of urban best management practices for reducing water pollution and supporting storm water management. For early warning and hazards prediction, we have examined a range of ML applications for supporting the prediction of various water-hazard related parameters. We included studies that combine ML methods with process-based models (e.g., conceptual and physics-based hydrological and sediment transport models) into hybrid models to increase the accuracy and speed of the predictions for water hazard-related parameters. We have also discussed how



innovative combinations of ML and remote sensing technologies can improve the discovery and extraction of useful hazard information and features that are critical to early-warning, rapid response and rescue, and recovery and restoration.

Our vision is that these methods can be combined into ML water management workflows that build on those already in use for characterizing and predicting multi-hazard hydrological events. By weaving together the ML methods we have described, long-term management processes including the six steps shown on the right hand side of **Figure 2** and outlined in the introduction can be captured. For example, risks associated with flood, drought and water quality can be identified using genetic algorithms, artificial neural networks, support vector machines, random forests, and other types of regression and hybrid models. Then planning objectives can be determined by weighing social risk and adaptive capacity using agent-based models, boosted regression trees, generalized additive models, and support vector machines. To inventory data, ground-based and satellite-based data can be reckoned, cataloged, and formatted for use in spatial-relationships-based queries, k-nearest neighbors, analytical hierarchy processes, and convolutional neural networks. To select mitigation approaches, classification schemes can be used along with multi-criteria decision methods. Uncertainty estimates can be used to evaluate the mitigation approaches selected. Finally, the insight gained from the ML results may

be discussed by the planners to modify and implement the approaches determined.

ML is often not the first choice of analytical tools for planners for a variety of reasons. The first is that reasonably robust methods with known uncertainty for analyzing water risks are well established and accepted in the water management community. ML methods are less proven even if they often can perform better on data than the traditional methods. To address the uncertainty in ML methods, some researchers (e.g., Morrison et al., 2003; Duncan, 2014) use metrics such as Receiver Operating Characteristic Curves for scoring the diagnostic ability of a binary (or higher dimensional) classifier system, or alternative goodness-of-fit measures for evaluating the reliability of ML output. Others (e.g., Munafò and Smith, 2018) suggest a method of investigation called *triangulation*, in which multiple approaches (at least 3) are used to address one question. The uncertainty associated with a complete model chain is large, especially at the required level of decision-making under climate change, urbanization (Dessai et al., 2009), and the accumulation of uncertainty at each level of the assessment (Merz et al., 2010). However, while each ML method may have its own strengths, weaknesses, and unrelated assumptions, uncertainty quantification can help assign some degree of confidence to results obtained.

We observe that many aspects of urban water security and hazard modeling are still underrepresented as ML problems, in particular, those pertaining to the prediction of indirect effects of water-related hazards and their associated risks. Additionally, the use of ML techniques often requires additional mathematical and computational training (and often large high performance compute resources) beyond traditional statistical methods, and time constraints of working water managers may not allow for this additional training. Nevertheless, understanding the development of sustainable urban water management planning, we can draw lessons from history and devise sensible approaches for the future that include ML. If we view hydrological systems as “structurally co-constituted of natural, engineered, and social elements,” (Brelsford et al., 2020), we may more readily employ ML to integrate disparate data and discover new perspectives on management practices based on the new patterns these methods reveal. In the near future, We also envision an increase in the applications of the hybrid modeling approaches (i.e., theory-guided ML) (Mekonnen et al., 2012; Karpatne et al., 2017; Frame, 2019) in the urban water management sector through the integration of data-driven ML methods and conventional process-based domain models.

REFERENCES

- Abaci, O., and Papanicolaou, T. (2009). Long-term effects of management practices on water-driven soil erosion in an intense agricultural sub-watershed: monitoring and modelling. *Hydrol. Process.* 23, 2818–2837. doi: 10.1002/hyp.7380
- Abdollahzadeh, A., Mukhlisin, M., and Shafie, A. (2011). Predict soil erosion with artificial neural network in Tanakami (Japan). *WSEAS Trans. Comput.* 10, 51–60.
- Abdulkadir, T. S., Muhammad, R. U. M., Yusof, K. W., Ahmad, M. H., Aremu, S. A., Gohari, A., et al. (2019). Quantitative analysis of soil erosion causative factors for susceptibility assessment in a complex watershed. *Cogent Eng.* 6:1594506. doi: 10.1080/23311916.2019.1594506
- Abolnaga, H. T., Ribbe, L., Frechen, F.-B., and Saghir, J. (2019). Urban water security: Definition and assessment framework. *Resources* 8:178. doi: 10.3390/resources8040178
- Adamala, S. (2017). An overview of big data applications in water resources engineering. *Mach. Learn. Res.* 2, 10–18. doi: 10.11648/j.ml.20170201.12
- AghaKouchak, A. (2015). A multivariate approach for persistence-based drought prediction: application to the 2010–2011 east Africa drought. *J. Hydrol.* 526, 127–135. doi: 10.1016/j.jhydrol.2014.09.063
- Ahmed, A. N., Othman, F. B., Afan, H. A., Ibrahim, R. K., Fai, C. M., Hossain, M. S., et al. (2019). Machine learning methods for better water quality prediction. *J. Hydrol.* 578:124084. doi: 10.1016/j.jhydrol.2019.124084
- Ahmed, N., Mahmud, S., Elahi, M., Ahmed, S., and Sujauddin, M. (2018). Forecasting river sediment deposition through satellite image driven unsupervised machine learning techniques. *Remote Sens. Appl. Soc. Environ.* 133, 435–444. doi: 10.1016/j.rsase.2018.12.011
- Ait Kadi, M., and Arriens, W. L. (2012). “Increasing water security: a development imperative,” in *Perspectives Paper of the Global Water Partnership Technical Committee*. Stockholm: Global Water Partnership.
- Allen, M. R., Zaidi, S. M. A., Chandola, V., Morton, A. M., Brelsford, C. M., McManamay, R. A., et al. (2018). A survey of analytical methods for inclusion in a new energy-water nexus knowledge discovery framework. *Big Earth Data* 2, 197–227. doi: 10.1080/20964471.2018.1524344

AUTHOR CONTRIBUTIONS

MA-D suggested the focus on urban hydrology, co-wrote the introduction and vision sections, and wrote the text on machine learning for flooding. HX developed the outline and organization of the paper based on urban water management practices, co-wrote the introduction and vision, and wrote the text for multi-hazard risk, soil erosion and sediment transport, and selection of best management practices. KK co-wrote the introduction, provided the research trend analysis on the historical application of machine learning to water management and hazard, and wrote the text on applications of machine learning to disaster management using remote sensing. DR wrote the sections on machine learning application to drought and water quality characterization and prediction. All authors edited and revised the document throughout.

FUNDING

Funding was provided by the Oak Ridge National Laboratory Computational Science and Engineering Division.

- Altunkaynak, A. (2009). Sediment load prediction by genetic algorithms. *Adv. Eng. Softw.* 40, 928–934. doi: 10.1016/j.advengsoft.2008.12.009
- Anzai, Y. (2012). *Pattern Recognition and Machine Learning*. Elsevier: Amsterdam.
- Auerbach, D., Buchanan, B., Alexiades, A., Anderson, E., Encalada, A., Larson, E., et al. (2015). Towards catchment classification in data-scarce regions. *Ecohydrology* 9, 1235–1247. doi: 10.1002/eco.1721
- Azamathulla, H., Ab Ghani, A., and Fei, S. (2012). Anfis-based approach for predicting sediment transport in clean sewer. *Appl. Soft Comput.* 12, 1227–1230. doi: 10.1016/j.asoc.2011.12.003
- Azamathulla, H., Asce, M., and Ab Ghani, A. (2011). Anfis-based approach for predicting the scour depth at culvert outlets. *J. Pipeline Syst. Eng. Practice* 3, 1227–1230. doi: 10.1061/(ASCE)PS.1949-1204.0000066
- Babovic, V., Canizares, R., Jensen, H., and Kliting, A. (2001). Neural networks as routine for error updating of numerical models. *J. Hydraul. Eng. ASCE* 127:3. doi: 10.1061/(ASCE)0733-9429(2001)127:3(181)
- Badrzadeh, H., Sarukkalige, R., and Jayawardena, A. (2013). Impact of multi-resolution analysis of artificial intelligence models inputs on multi-step ahead river flow forecasting. *J. Hydrol.* 507, 75–85. doi: 10.1016/j.jhydrol.2013.10.017
- Bakhtyar, R., Ghaheeri, A., Yeganeh-Bakhtiary, A., and Baldock, T. (2008). Longshore sediment transport estimation using fuzzy inference system. *Appl. Ocean Res.* 30, 273–286. doi: 10.1016/j.apor.2008.12.001
- Barzegar, R., Aalami, M. T., and Adamowski, J. (2020). Short-term water quality variable prediction using a hybrid CNN-LSTM deep learning model. *Stochast. Environ. Res. Risk Assess.* 34, 1–19. doi: 10.1007/s00477-020-01776-2
- Belayneh, A., and Adamowski, J. (2012). Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Appl. Comput. Intell. Soft Comput.* 2012:794061. doi: 10.1155/2012/794061
- Belayneh, A., Adamowski, J., Khalil, B., and Ozga-Zielinski, B. (2014). Long-term SPI drought forecasting in the awash river basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *J. Hydrol.* 508, 418–429. doi: 10.1016/j.jhydrol.2013.10.052
- Belayneh, A., Adamowski, J., Khalil, B., and Quilty, J. (2016). Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmos. Res.* 172, 37–47. doi: 10.1016/j.atmosres.2015.12.017

- Beran, B., and Piasecki, M. (2008). Availability and coverage of hydrologic data in the US geological survey national water information system (NWIS) and us environmental protection agency storage and retrieval system (storet). *Earth Sci. Inform.* 1, 119–129. doi: 10.1007/s12145-008-0015-2
- Bertolotto, M., Martino, S. D., Ferrucci, F., and Kechadi, T. (2007). Towards a framework for mining and analysing spatio-temporal datasets. *Int. J. Geogr. Inform. Sci.* 21, 895–906. doi: 10.1080/13658810701349052
- Bhattacharya, B., Price, R., and Solomatine, D. (2007). A machine learning approach to modeling sediment transport. *J. Hydraul. Eng. ASCE* 133, 440–450. doi: 10.1061/(ASCE)0733-9429(2007)133:4(440)
- Bian, J., Li, A., Zuo, J., Lei, G., Zhang, Z., and Nan, X. (2019). Estimating 2009–2017 impervious surface change in Gwadar, Pakistan using the HJ-1a/b constellation, GF-1/2 data, and the random forest algorithm. *ISPRS Int. J. Geo-Inf.* 8:443. doi: 10.3390/ijgi8100443
- Brelsford, C., Dumas, M., Schlager, E., Dermody, B. J., Aiuvalasit, M., Allen-Dumas, M. R., et al. (2020). Developing a sustainability science approach for water systems. *Ecol. Soc.* doi: 10.5751/ES-11515-250223
- Bühler, Y., Kumar, S., Veitinger, J., Christen, M., Stoffel, A., and Snehmani (2013). Automated identification of potential snow avalanche release areas based on digital elevation models. *Nat. Hazards Earth Syst. Sci.* 13, 1321–1335. doi: 10.5194/nhess-13-1321-2013
- Bui, D. T., Khosravi, K., Tiefenbacher, J., Nguyen, H., and Kazakis, N. (2020). Improving prediction of water quality indices using novel hybrid machine-learning algorithms. *Sci. Tot. Environ.* 2020:137612. doi: 10.1016/j.scitotenv.2020.137612
- Cao, L., and Gu, Q. (2002). Dynamic support vector machines for non-stationary time series forecasting. *Intell. Data Anal.* 6, 67–83. doi: 10.3233/IDA-2002-6105
- Cardwell, H. E., Langsdale, S., and Stephenson, K. (2009). “Developing best practices for computer aided dispute resolution,” in *World Environmental and Water Resources Congress May 17* (Kansas City, MI: American Society of Civil Engineers (ASCE) Library). doi: 10.1061/41036(342)484
- Carlisle, D., Falcone, J., and Meador, M. (2008). Predicting the biological condition of streams: use of geospatial indicators of natural and anthropogenic characteristics of watersheds. *Environ. Monitor. Assess.* 151, 143–160. doi: 10.1007/s10661-008-0256-z
- Carson, A., Windsor, M., Hill, H., Haigh, T., Wall, N., Smith, J., et al. (2018). Serious gaming for participatory planning of multi-hazard mitigation. *Int. J. River Basin Manage.* 16, 1–47. doi: 10.1080/15715124.2018.1481079
- Cervone, G., Schnebele, E., Waters, N., Moccaldi, M., and Sicignano, R. (2017). “Using social media and satellite data for damage assessment in urban areas during emergencies,” in *Seeing Cities Through Big Data*, eds P. Thakuriah, N. Tilahun, and M. Zellner (Cham: Springer), 443–457. doi: 10.1007/978-3-319-40902-3_24
- Chambers, L. (2000). *The Practical Handbook of Genetic Algorithms: Applications, 2nd Edn* Boca Raton, FL: CRC Press. doi: 10.1201/9781420035568
- Chandola, V., Cheboli, D., and Kumar, V. (2009a). Anomaly detection: a survey. *ACM Comput. Surveys* 41, 1–58. doi: 10.1145/1541880.1541882
- Chandola, V., Cheboli, D., and Kumar, V. (2009b). *Detecting Anomalies in a Time Series Database*. Computer Science Department, University of Minnesota, Tech. Rep.
- Chang, F.-J., Hsu, K., and Chang, L.-C. (2019). *Flood Forecasting Using Machine Learning Methods*. MDPI.
- Chen, C.-Y., Li, Wang, and Deng (2019). A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: a case study of the Yangtze river delta, China. *Int. J. Environ. Res. Public Health* 17:49. doi: 10.3390/ijerph17010049
- Chen, L., Wei, G., and Shen, Z. (2015). An auto-adaptive optimization approach for targeting nonpoint source pollution control practices. *Sci. Rep.* 5:15393. doi: 10.1038/srep15393
- Chen, S., Dong, S., Cao, Z., and Guo, J. (2020). A compound approach for monthly runoff forecasting based on multiscale analysis and deep network with sequential structure. *Water* 12:2274. doi: 10.3390/w12082274
- Cheng, M.-S., Zhen, J., and Shoemaker, L. (2006). BMP decision support system for evaluating stormwater management alternatives. *Front. Environ. Sci. Eng. China* 3, 453–463. doi: 10.1007/s11783-009-0153-x
- Cigizoglu, H. (2002). Suspended sediment estimation for rivers using artificial neural networks and sediment rating curves. *Turkish J. Eng. Environ. Sci.* 26, 27–36.
- Cohon, J. L., and Marks, D. H. (1975). A review and evaluation of multiobjective programming techniques. *Water Resour. Res.* 11, 208–220. doi: 10.1029/WR011i002p00208
- Cook, E. R., Seager, R., Cane, M. A., and Stahle, D. W. (2007). North American drought: reconstructions, causes, and consequences. *Earth Sci. Rev.* 81, 93–134. doi: 10.1016/j.earscirev.2006.12.002
- Cook, S., [van Roon], M., Ehrenfried, L., LaGro, J., and Yu, Q. (2019). “Chapter 27 - WSUD 'best in class' Case Studies from Australia, New Zealand, United States, Europe, and Asia,” in *Approaches to Water Sensitive Urban Design*, eds A. K. Sharma, T. Gardner, and D. Begbie (Amsterdam: Elsevier; Woodhead Publishing), 561–585. doi: 10.1016/B978-0-12-812843-5.00027-7
- Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T. T., Abba, S., Vojtek, M., et al. (2019). Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and GIS techniques. *Remote Sens.* 12, 1–26. doi: 10.3390/rs12010106
- Costello, C., Ensor, C., and Chadwick, S. (2019). *Indirect Losses From Flood Disasters*. Available online at: <https://firststreet.org/flood-lab/research/indirect-losses-from-flood-disasters> (accessed May 02, 2020).
- Crossman, J., Futter, M., Oni, S., Whitehead, P., Jin, L., Butterfield, D., et al. (2013). Impacts of climate change on hydrology and water quality: future proofing management strategies in the Lake Simcoe Watershed, Canada. *J. Great Lakes Res.* 39, 19–32. doi: 10.1016/j.jglr.2012.11.003
- Dai, A. (2011). Drought under global warming: a review. *Wiley Interdisc. Rev.* 2, 45–65. doi: 10.1002/wcc.81
- Dai, L., Rijswijk, H., Driessen, P., and Keessen, A. (2017). Governance of the sponge city programme in china with Wuhan as a case study. *Int. J. Water Resour. Dev.* 34, 1–19. doi: 10.1080/07900627.2017.1373637
- Das, M., and Parthasarathy, S. (2009). “Anomaly detection and spatio-temporal analysis of global climate system,” in *Proceedings of the Third International Workshop on Knowledge Discovery From Sensor Data* (Paris), 142–150. doi: 10.1145/1601966.1601989
- de Goyet, C. D. V., Marti, R. Z., and Osorio, C. (2006). *Natural Disaster Mitigation and Relief. In Disease Control Priorities in Developing Countries*. 2nd edition. The International Bank for Reconstruction and Development/The World Bank.
- Dessai, S., Hulme, M., Lempert, R., and Pielke R. Jr. (2009). Do we need better predictions to adapt to a changing climate? *EOS Trans. Am. Geophys. Union* 90, 111–112. doi: 10.1029/2009EO130003
- DHI (2003). *A Versatile Decision Support Tool for Integrated Water Resources Management Planning*. Available online at: <https://www.slideshare.net/indiaWRM/14-dhi-integrated-decision-support-tools-for-water-resources-managementsep17> (accessed April 14, 2020).
- Dibike, Y., and Solomatine, D. (2000). River flow forecasting using artificial neural networks. *Phys. Chem. Earth Part B* 26, 1–7. doi: 10.1016/S1464-1909(01)85005-X
- Dickson-Anderson, S., Wallace, C., and Newton, J. (2016). Water security assessment indicators: the rural context. *Water Resour. Manage.* 30, 1567–1604. doi: 10.1007/s11269-016-1254-5
- Dogan, E., Sengorur, B., and Koklu, R. (2009). Modeling biological oxygen demand of the Melen river in turkey using an artificial neural network technique. *J. Environ. Manage.* 90, 1229–1235. doi: 10.1016/j.jenvman.2008.06.004
- Dörendahl, E. I. (2013). “Boundary work and water resources: towards improved management and research practice?,” in *ZEF Working Paper Series*, No. 122; Zentrum für Entwicklungsforschung (Bonn).
- Duncan, A. P. (2014). *The analysis and application of artificial neural networks for early warning systems in hydrology and the environment* (Ph.D. thesis). University of Exeter, Exeter, United Kingdom. doi: 10.13140/RG.2.1.1602.4806
- Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., et al. (2014). Global meteorological drought-part 2: seasonal forecasts. *Hydrol. Earth Syst. Sci.* 18, 2669–2678. doi: 10.5194/hess-18-2669-2014
- Eldho, T., Zope, P., and Kulkarni, A. (2018). “Urban flood management in coastal regions using numerical simulation and geographic information system,” in *Integrating Disaster Science and Management* (Elsevier), 205–219. doi: 10.1016/B978-0-12-812056-9.00012-9
- Emerson, K., Nabatchi, T., and Balogh, S. (2012). An integrated framework for collaborative governance. *J. Publ. Administr. Res. Theory* 22:1. doi: 10.1093/jopart/mur011

- EPA (2012). *Water Quality Standards 101*. Technical report, U.S. Environmental Protection Agency.
- EPA (2019). *Examples of Water Quality Assessments for Watershed Health*. Available online at: <https://www.epa.gov/hwp/examples-water-quality-assessments-watershed-health> (accessed October 12, 2020).
- Eriksson, M., Nutter, J., Day, S., Guttman, H., James, R., and Quibell, G. (2015). Challenges and commonalities in basin-wide water management. *Aquat. Proc.* 5, 44–57. doi: 10.1016/j.aqpro.2015.10.007
- Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., et al. (2018). Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: a case study in china. *Energy Convers. Manage.* 164, 102–111. doi: 10.1016/j.enconman.2018.02.087
- FEMA (2015). *Integrating Disaster Data into Hazard Mitigation Planning*. Federal Emergency Management Agency.
- Feng, Q., Liu, J., and Gong, J. (2015). Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—a case of Yuyao, China. *Water* 7, 1437–1455. doi: 10.3390/w7041437
- Fernández, C., Vega, J. A., Fonturbel, T., and Jiménez, E. (2009). Streamflow drought time series forecasting: a case study in a small watershed in North West Spain. *Stochast. Environ. Res. Risk Assess.* 23:1063. doi: 10.1007/s00477-008-0277-8
- FHWA (2000). *Stormwater Best Management Practices in an Ultra-Urban Setting: Selection and Monitoring*. Technical report, Federal Highway Administration.
- Frame, J. M. (2019). “Toward global terrestrial hydrology with theory guided machine learning,” in *Proceedings of the AGU Annual Meeting* (San Francisco, CA: NASA Goddard Space Flight Center).
- Francke, T., López-Tarazón, J., and Schröder, B. (2008). Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrol. Process.* 22, 4892–4904. doi: 10.1002/hyp.7110
- Freeman, P., and Warner, K. (2001). *Vulnerability of Infrastructure to Climate Variability: How Does This Affect Infrastructure Lending Policies?* The World Bank, ProVention Consortium.
- Ganasri, B., and Gowda, R. (2015). Assessment of soil erosion by Rusle model using remote sensing and GIS - a case study of Nethravathi basin. *Geosci. Front.* 7, 953–961. doi: 10.1016/j.gsf.2015.10.007
- Gangrade, S., Kao, S.-C., Dullo, T. T., Kalyanapu, A. J., and Preston, B. L. (2019). Ensemble-based flood vulnerability assessment for probable maximum flood in a changing environment. *J. Hydrol.* 576, 342–355. doi: 10.1016/j.jhydrol.2019.06.027
- Gangrade, S., Kao, S.-C., Naz, B. S., Rastogi, D., Ashfaq, M., Singh, N., et al. (2018). Sensitivity of probable maximum flood in a changing environment. *Water Resour. Res.* 54, 3913–3936. doi: 10.1029/2017WR021987
- García-Alba, J., Bárcena, J. F., Ugarteburu, C., and García, A. (2019). Artificial neural networks as emulators of process-based models to analyse bathing water quality in estuaries. *Water Res.* 150, 283–295. doi: 10.1016/j.watres.2018.11.063
- García-Aristizabal, A., and Marzocchi, W. (2013). New methodologies for multi-hazard and multi-risk assessment methods for Europe. EU Matrix project.
- Garrick, D., and Hall, J. (2014). Water security and society: Risks, metrics, and pathways. *Annu. Rev. Environ. Resour.* 39, 611–639. doi: 10.1146/annurev-environ-013012-093817
- Ge, P., Gokon, H., and Meguro, K. (2020). A review on synthetic aperture radar-based building damage assessment in disasters. *Remote Sens. Environ.* 240:111693. doi: 10.1016/j.rse.2020.111693
- Goldstein, E., and Coco, G. (2015). Machine learning components in deterministic models: hybrid synergy in the age of data. *Front. Environ. Sci.* 3:33. doi: 10.3389/fenvs.2015.00033
- Govedarica, M., and Jakovljević, G. (2019). “Monitoring spatial and temporal variation of water quality parameters using time series of open multispectral data,” in *Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019)*, Vol. 11174 (Paphos: International Society for Optics and Photonics). doi: 10.1117/12.2533708
- Goyal, M. (2014). Modeling of sediment yield prediction using m5 model tree algorithm and wavelet regression. *Water Resour. Manage.* 28, 1991–2003. doi: 10.1007/s11269-014-0590-6
- Gruber, F., and Mergili, M. (2013). Regional-scale analysis of high-mountain multi-hazard and risk indicators in the Pamir (Tajikistan) with grass GIS. *Nat. Hazards Earth Syst. Sci.* 13, 2779–2796. doi: 10.5194/nhess-13-2779-2013
- Guidolin, M., Chen, A. S., Ghimire, B., Keedwell, E. C., Djordjević, S., and Savić, D. A. (2016). A weighted cellular automata 2d inundation model for rapid flood analysis. *Environ. Model. Softw.* 84, 378–394. doi: 10.1016/j.envsoft.2016.07.008
- Guy, H. P. (1970). *Sediment Problems in Urban Areas*. United States Geological Survey. doi: 10.3133/cir601E
- Hadka, D., and Reed, P. (2013). Borg: an auto-adaptive many-objective evolutionary computing framework. *Evol. Comput.* 21, 231–259. doi: 10.1162/EVCO_a_00075
- Hafeez, S., Wong, M. S., Ho, H. C., Nazeer, M., Nichol, J., Abbas, S., et al. (2019). Comparison of machine learning algorithms for retrieval of water quality indicators in case-II waters: a case study of Hong Kong. *Remote Sens.* 11:617. doi: 10.3390/rs11060617
- Haghiabi, A., Nasrolahi, A., and Parsaie, A. (2018). Water quality prediction using machine learning methods. *Water Qual. Res. J.* 53:wqrj2018025. doi: 10.2166/wqrj.2018.025
- Hajigholizadeh, M., Melesse, A., and Fuentes, H. (2018). Erosion and sediment transport modelling in shallow waters: a review on approaches, models and applications. *Int. J. Environ. Res. Publ. Health* 15:518. doi: 10.3390/ijerph15030518
- Hall, J., and Hossain, A. (2020). Mapping urbanization and evaluating its possible impacts on stream water quality in Chattanooga, Tennessee, using GIS and remote sensing. *Sustainability* 12:1980. doi: 10.3390/su12051980
- Hall, J. W. (2004). Comment on ‘of data and models’. *J. Hydroinform.* 6, 75–77. doi: 10.2166/hydro.2004.0006
- Hao, Z., Hao, F., Singh, V. P., Ouyang, W., and Cheng, H. (2017). An integrated package for drought monitoring, prediction and analysis to aid drought modeling and assessment. *Environ. Model. Softw.* 91, 199–209. doi: 10.1016/j.envsoft.2017.02.008
- Hao, Z., Singh, V. P., and Xia, Y. (2018). Seasonal drought prediction: advances, challenges, and future prospects. *Rev. Geophys.* 56, 108–141. doi: 10.1002/2016RG000549
- Hayatbini, N., Kong, B., Hsu, K.-I., Nguyen, P., Sorooshian, S., Stephens, G., et al. (2019). Conditional generative adversarial networks (cGANs) for near real-time precipitation estimation from multispectral GOES-16 satellite imageries-PERSIANN-cGAN. *Remote Sens.* 11:2193. doi: 10.3390/rs11192193
- Hewett, C., Simpson, C., Wainwright, J., and Hudson, S. (2018). Communicating risks to infrastructure due to soil erosion: a bottom-up approach. *Land Degrad. Dev.* 29, 1282–1294. doi: 10.1002/ldr.2900
- Hodgson, M., Davis, B., and Kotelenska, J. (2010). “Remote sensing and GIS data/information in the emergency response/recovery phase,” in *Geospatial Techniques in Urban Hazard and Disaster Analysis*, eds P. Showalter and Y. Lu (Dordrecht: Springer), 327–354. doi: 10.1007/978-90-481-2238-7_16
- Hoekstra, A. Y., Buurman, J., and van Ginkel, K. C. H. (2018). Urban water security: a review. *Environ. Res. Lett.* 13:053002. doi: 10.1088/1748-9326/aaba52
- Holzbecher, E., Hadidi, A., Barghash, H., and Balushi, K. (2019). “Application of big data and technologies for integrated water resources management - a survey,” in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (Grenada), 309–315. doi: 10.1109/SNAMS.2019.8931722
- Hong, Y., Hsu, K.-L., Sorooshian, S., and Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteorol.* 43, 1834–1853. doi: 10.1175/JAM2173.1
- Hosseinalizadeh, M., Kariminejad, N., Chen, W., Pourghasemi, H. R., Alinejad, M., Mohammadian Behbahani, A., et al. (2019). Gully headcut susceptibility modeling using functional trees, naive Bayes tree, and random forest models. *Geoderma* 342, 1–11. doi: 10.1016/j.geoderma.2019.01.050
- Hou, D., Song, X., Zhang, G., Zhang, H., and Loaiciga, H. (2013). An early warning and control system for urban, drinking water quality protection: China’s experience. *Environ. Sci. Pollut. Res.* 20, 4496–4508. doi: 10.1007/s11356-012-1406-y
- Huang, J., Svoboda, M., Wood, A., Schubert, S. D., Peters-Lidard, C. D., Wood, E. F., et al. (2016). NOAA drought task force 2016: research to advance national drought monitoring and prediction capabilities.

- doi: 10.1201/9781315265551-10 Available online at: <https://repository.library.noaa.gov/view/noaa/11194>
- Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., and Yang, H. (2019). Machine learning and artificial intelligence to aid climate change research and preparedness. *Environ. Res. Lett.* 14:124007. doi: 10.1088/1748-9326/ab4e55
- Huo, A., Zhang, J., Qiao, C., Li, C., Xie, J., Wang, J., et al. (2014). Multispectral remote sensing inversion for city landscape water eutrophication based on genetic algorithm-support vector machine. *Water Qual. Res. J. Can.* 49, 285–293. doi: 10.2166/wqrjc.2014.040
- Jain, P., Schoen-Phelan, B., and Ross, R. (2020). “Automatic flood detection in sentinel-2 images using deep convolutional neural networks,” in *Proceedings of the 35th Annual ACM Symposium on Applied Computing* (Brno), 617–623. doi: 10.1145/3341105.3374023
- Jetten, V. G., Alkema, D., van Westen, C. J., and Brussel, M. J. G. (2014). “Development of the Caribbean handbook on disaster risk information management,” in *International Conference on Analysis and Management of Changing Risks for Natural Hazards 2014* (Padua).
- Jiang, J., Qin, C.-Z., Yu, J., Cheng, C., Liu, J., and Huang, J. (2020). Obtaining urban waterlogging depths from video images using synthetic image data. *Remote Sens.* 12:1014. doi: 10.3390/rs12061014
- Joslin, P. (2016). *Data-driven analyses of watersheds as coupled human-nature systems* (Ph.D. thesis). University of Illinois at Urbana-Champaign, Champaign, IL, United States.
- Kalantari, Z., Ferreira, C., Koutsouris, A., Ahlmer, A.-K., Cerdà, A., and Destouni, G. (2019). Assessing flood probability for transportation infrastructure based on catchment characteristics, sediment connectivity and remotely sensed soil moisture. *Sci. Tot. Environ.* 661, 393–406. doi: 10.1016/j.scitotenv.2019.01.009
- Kam, J., Sheffield, J., and Wood, E. F. (2014). A multiscale analysis of drought and pluvial mechanisms for the southeastern United States. *J. Geophys. Res.* 119, 7348–7367. doi: 10.1002/2014JD021453
- Kao, S.-C., and Ganguly, A. R. (2011). Intensity, duration, and frequency of precipitation extremes under 21st-century warming scenarios. *J. Geophys. Res.* 116, 1–14. doi: 10.1029/2010JD015529
- Kappes, M., Keiler, M., v. Elverfeldt, K., and Glade, T. (2012). Challenges of dealing with multi-hazard risk: a review. *Nat. Hazards* 64, 1925–1958. doi: 10.1007/s11069-012-0294-2
- Karlsson, C., Kalantari, Z., Mörtberg, U., Olofsson, B., and Lyon, S. (2017). Natural hazard susceptibility assessment for road planning using spatial multi-criteria analysis. *Environ. Manage.* 60, 823–851. doi: 10.1007/s00267-017-0912-6
- Karpatne, A., Atluri, G., Faghmous, J., Steinbach, M., Banerjee, A., Ganguly, A., et al. (2016). Theory-guided data science: a new paradigm for scientific discovery. *arXiv[Preprint].arXiv:1612.08544*. doi: 10.1109/TKDE.2017.2720168
- Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., et al. (2017). Theory-guided data science: a new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Eng.* 29, 2318–2331.
- Kebede, A. (2009). *Water Quality Modeling An Overview*. Available online at: https://files.nc.gov/ncdeq/WaterQuality/Planning/TMDL/Modeling/Modeling_101_for_FON_stakeholder_May09.pdf (accessed October 10, 2020).
- Kennedy, J., and Eberhart, R. (1995). “Particle swarm optimization,” in *Proceedings of ICNN’95-International Conference on Neural Networks, Vol. 4* (Perth, WA), 1942–1948. doi: 10.1109/ICNN.1995.488968
- Khan, Y., and See, C. S. (2016). “Predicting and analyzing water quality using machine learning: a comprehensive model,” in *2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)* (Farmingdale, NY), 1–6. doi: 10.1109/LISAT.2016.7494106
- Kitsikoudis, V., Sidiropoulos, E., and Hrisanthou, V. (2014). Assessment of sediment transport approaches for sand-bed rivers by means of machine learning. *Hydrol. Sci. J.* 60, 1566–1586. doi: 10.1080/02626667.2014.909599
- Knaepen, M., and Hulscher, S. (2003). Use of a genetic algorithm to improve predictions of alternate bar dynamics. *Water Resour. Res.* 39, 1–11. doi: 10.1029/2002WR001793
- Komendantova, N., Mrzyglocki, R., Mignan, A., Khazai, B., Wenzel, F., Patt, A., et al. (2014). Multi-hazard and multi-risk decision-support tools as a part of participatory risk governance: feedback from civil protection stakeholders. *Int. J. Disast. Risk Reduct.* 8, 50–67. doi: 10.1016/j.ijdr.2013.12.006
- Konapala, G., and Mishra, A. (2020). Quantifying climate and catchment control on hydrological drought in continental United States. *Water Resour. Res.* 56:e2018WR024620. doi: 10.1029/2018WR024620
- Krajewski, W., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., et al. (2016). Real-time flood forecasting and information system for the state of Iowa. *Bull. Am. Meteorol. Soc.* 98, 539–554. doi: 10.1175/BAMS-D-15-00243.1
- Krasnopolsky, V., and Fox-Rabinovitz, M. (2005). “Complex hybrid models combining deterministic and machine learning components as a new synergetic paradigm in numerical climate modeling and weather prediction, Vol. 3,” in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005* (Montreal, QC), 1615–1620.
- Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T. (2014). Precipitation estimates from msg seviri daytime, nighttime, and twilight data with random forests. *J. Appl. Meteorol. Climatol.* 53, 2457–2480. doi: 10.1175/JAMC-D-14-0082.1
- Kurte, K., Potnis, A., and Durbha, S. (2019). “Semantics-enabled spatio-temporal modeling of earth observation data: an application to flood monitoring,” in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Advances on Resilient and Intelligent Cities* (Chicago, IL), 41–50. doi: 10.1145/3356395.3365545
- Kurte, K. R., Durbha, S. S., King, R. L., Younan, N. H., and Vatsavai, R. (2017). Semantics-enabled framework for spatial image information mining of linked earth observation data. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 10, 29–44. doi: 10.1109/JSTARS.2016.2547992
- Kuswanto, H., and Naufal, A. (2019). Evaluation of performance of drought prediction in Indonesia based on TRMM and merra-2 using machine learning methods. *MethodsX* 6, 1238–1251. doi: 10.1016/j.mex.2019.05.029
- Lambert, P. (2014). *Caribbean Handbook on Risk Management*. Technical report, ACP-EU NDRR Program.
- Lehman, W., and Light, M. (2016). Using hec-fia to identify indirect economic losses. *E3S Web Confer.* 7:05008. doi: 10.1051/e3sconf/20160705008
- Li, C., Li, Z., Wu, J., Zhu, L., and Yue, J. (2018). A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features. *Inform. Process. Agric.* 5, 11–20. doi: 10.1016/j.inpa.2017.11.002
- Liang, J., Li, W., Bradford, S., and Simunek, Jiri, J. (2019). Physics-informed data-driven models to predict surface runoff water quantity and quality in agricultural fields. *Water* 11:200. doi: 10.3390/w11020200
- Limbrunner, J., Chapra, S., and Kirshen, P. (2013). Classic optimization techniques applied to stormwater and nonpoint source pollution management at the watershed scale. *Journal of Water Resources Plann. Manage.* 139, 486–491. doi: 10.1061/(ASCE)WR.1943-5452.0000361
- Lin, B., and Montazeri Namin, M. (2005). Modelling suspended sediment transport using an integrated numerical and ANNs model. *J. Hydraul. Res.* 43, 302–310. doi: 10.1080/00221680509500124
- Lin, Y., Zhang, H., Li, G., Wang, T., Wan, L., and Lin, H. (2019). Improving impervious surface extraction with shadow-based sparse representation from optical, SAR, and LIDAR data. *IEEE Journal of Selected Top. Appl. Earth Observ. Remote Sens.* 12, 2417–2428. doi: 10.1109/JSTARS.2019.2907744
- Liu, M., and Huang, M. C. (2014). *COMPOUND DISASTERS AND COMPOUNDING PROCESSES: Implications for Disaster Risk Management*. Prepared for the Global Assessment Report on Disaster Risk Reduction 2015, United Nations Office for Disaster Risk Reduction. Asian Development Bank Institute / National Graduate Institute for Policy Studies. Available online at: <https://www.preventionweb.net/english/hyogo/gar/2015/en/bgdocs/inputs/Liu%20and%20Huang,%202014.%20Compound%20disasters%20and%20compounding%20processes%20-%20Implications%20for%20Disaster%20Risk%20Management.pdf>
- López-Tarazón, J., Batalla, R. J., Vericat, D., and Francke, T. (2012). The sediment budget of a highly dynamic mesoscale catchment: the river isbena. *Geomorphology* 138, 15–28. doi: 10.1016/j.geomorph.2011.08.020
- Lu, H., and Ma, X. (2020). Hybrid decision tree-based machine learning models for short-term water quality prediction. *Chemosphere* 249:126169. doi: 10.1016/j.chemosphere.2020.126169
- Manning, R., Griffith, J. P., Pigot, T., and Vernon-Harcourt, L. F. (1890). On the flow of water in open channels and pipes. *Inst. Civil Eng. Trans.* 20, 161–207.
- Maringanti, C., Chaubey, I., Arabi, M., and B. E. (2008). A multi-objective optimization tool for the selection and placement of bmps for pesticide control. *Hydrol. Earth Syst. Sci. Discuss.* 5, 1821–1862. doi: 10.5194/hessd-5-1821-2008

- McGlinchy, J., Johnson, B., Muller, B., Joseph, M., and Diaz, J. (2019). "Application of UNET fully convolutional neural network to impervious surface segmentation in urban environment from high resolution satellite imagery," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium* (Yokohama), 3915–3918. doi: 10.1109/IGARSS.2019.8900453
- Medema, W., Wals, A., and Adamowski, J. (2014). Multi-loop social learning for sustainable land and water governance: towards a research agenda on the potential of virtual learning platforms. *Wageningen J. Life Sci.* 69, 23–38. doi: 10.1016/j.njas.2014.03.003
- Mekanik, F., Imteaz, M., Gato-Trinidad, S., and Elmahdi, A. (2013). Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes. *J. Hydrol.* 503, 11–21. doi: 10.1016/j.jhydrol.2013.08.035
- Mekonnen, B., Nazemi, A., Elshorbagy, A., Mazurek, K., and Putz, G. (2012). Hybrid modelling approach to prairie hydrology: fusing data-driven and process-based hydrological models. *Hydrol. Sci. J.* 60:6562. doi: 10.1080/02626667.2014.935778
- Merritt, W., Letcher, R., and Jakeman, A. (2003). A review of erosion and sediment transport models. *Environ. Model. Softw.* 18, 761–799. doi: 10.1016/S1364-8152(03)00078-1
- Merz, B., Hall, J., Disse, M., and Schumann, A. (2010). Fluvial flood risk management in a changing world. *Nat. Hazards Earth Syst. Sci.* 3, 509–527. doi: 10.5194/nhess-10-509-2010
- Miao, Z., Xiao, Y., Shi, W., He, Y., Gamba, P., Li, Z., et al. (2019). Integration of satellite images and open data for impervious surface classification. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 12, 1120–1133. doi: 10.1109/JSTARS.2019.2903585
- Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., et al. (2008). Stationarity is dead: whither water management? *Science* 319, 573–574. doi: 10.1126/science.1151915
- Mishra, A., Desai, V., and Singh, V. (2007). Drought forecasting using a hybrid stochastic and neural network model. *J. Hydrol. Eng.* 12, 626–638. doi: 10.1061/(ASCE)1084-0699(2007)12:6(626)
- Mo, K. C., and Lyon, B. (2015). Global meteorological drought prediction using the North American multi-model ensemble. *J. Hydrometeorol.* 16, 1409–1424. doi: 10.1175/JHM-D-14-0192.1
- Mohanty, J. R., and Mohapatra, M. R. (2018). Rainfall prediction using support vector machine (SVM). *IOSR J. Comput. Eng.* 20, 6–13. doi: 10.9790/0661-2003020613
- Morgan, R., Morgan, D., and Finney, H. (1984). A predictive model for assessment of erosion risk. *J. Agric. Eng. Res.* 30, 245–253. doi: 10.1016/S0021-8634(84)80025-6
- Morid, S., Smakhtin, V., and Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices. *Int. J. Climatol.* 27, 2103–2111. doi: 10.1002/joc.1498
- Morrison, A. M., Coughlin, K., Shine, J. P., Coull, B. A., and Rex, A. C. (2003). Receiver operating characteristic curve analysis of beach water quality indicator variables. *Appl. Environ. Microbiol.* 69, 6405–6411. doi: 10.1128/AEM.69.11.6405-6411.2003
- Mosavi, A., Ozturk, P., and Chau, K.-W. (2018). Flood prediction using machine learning models: literature review. *Water* 10:1536. doi: 10.3390/w10111536
- Munafó, M. R., and Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature* 553, 399–401. doi: 10.1038/d41586-018-01023-3
- Mustafa, M., Abdulkadir, T. S., Wan Yusof, K., Hashim, A., Waris, M., and Shahbaz, M. (2018). SVM-based geospatial prediction of soil erosion under static and dynamic conditioning factors. *MATEC Web Conf.* 203:04004. doi: 10.1051/mateconf/201820304004
- Nearing, M., Jetten, V., Baffaut, C., Cerdan, O., Couturier, A., Hernandez, M., et al. (2005). Modeling response of soil erosion and runoff to changes in precipitation and cover. *Catena* 61, 131–154. doi: 10.1016/j.catena.2005.03.007
- Newman, J., Maier, H., Riddell, G., Zecchin, A., Daniell, J., Schfer, A., et al. (2017). Review of literature on decision support systems for natural hazard risk reduction: current status and future research directions. *Environ. Model. Softw.* 96, 378–409. doi: 10.1016/j.envsoft.2017.06.042
- Nowell, L. H., and Resek, E. A. (1994). "National standards and guidelines for pesticides in water, sediment, and aquatic organisms: application to water-quality assessments," in *Reviews of Environmental Contamination and Toxicology*, ed G. W. Ware (New York, NY: Springer), 1–154. doi: 10.3133/ofr9444
- Noymanee, J., Nikitin, N. O., and Kalyuzhnaya, A. V. (2017). Urban pluvial flood forecasting using open data with machine learning techniques in Pattani basin. *Proc. Comput. Sci.* 119, 288–297. doi: 10.1016/j.procs.2017.11.187
- NRCS (2003). *National Planning Procedures Handbook (NPPH), Amendment 4*. Available online at: https://nutrientmanagement.tamu.edu/content/resources/nrcs_handbook.pdf (accessed February 5, 2020).
- NRCS (2008). *Urban Soil Erosion and Sediment Control*. USDA.
- NRCS (2011). *Conservation Planning and Regulatory Compliance Handbook*. Technical report, Pennsylvania NRCS.
- Olden, J., Kennard, M., and Pusey, B. (2012). A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology* 5, 503–518. doi: 10.1002/eco.251
- Onderka, M. (2012). Dynamics of storm-driven suspended sediments in a headwater catchment described by multivariable modeling. *J. Soils Sedim.* 12, 620–635. doi: 10.1007/s11368-012-0480-6
- Pagán, B. R., Ashfaq, M., Rastogi, D., Kendall, D. R., Kao, S.-C., Naz, B. S., et al. (2016). Extreme hydrological changes in the southwestern us drive reductions in water supply to southern California by mid century. *Environ. Res. Lett.* 11:094026. doi: 10.1088/1748-9326/11/9/094026
- Palani, S., Liong, S.-Y., and Tkalic, P. (2008). An ANN application for water quality forecasting. *Mar. Pollut. Bull.* 56, 1586–1597. doi: 10.1016/j.marpolbul.2008.05.021
- Palmer, R., Cardwell, H., Lorie, M., and Werick, W. (2013). Disciplined planning, structured participation, and collaborative modeling-applying shared vision planning to water resources. *JAWRA J. Am. Water Resour. Assoc.* 49, 614–628. doi: 10.1111/jawr.12067
- Pan, B., Hsu, K., AghaKouchak, A., and Sorooshian, S. (2019). Improving precipitation estimation using convolutional neural network. *Water Resour. Res.* 55, 2301–2321. doi: 10.1029/2018WR024090
- Park, R., Clough, J., and Wellman, M. (2008). Aquatox: Modeling environmental fate and ecological effects in aquatic ecosystems. *Ecol. Model.* 213, 1–15. doi: 10.1016/j.ecolmodel.2008.01.015
- Park, S., Im, J., Jang, E., and Rhee, J. (2016). Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agric. For. Meteorol.* 216, 157–169. doi: 10.1016/j.agrformet.2015.10.011
- Pendergrass, A. G., Meehl, G. A., Pulwarty, R., Hobbins, M., Hoell, A., AghaKouchak, A., et al. (2020). Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nat. Clim. Change* 10, 191–199. doi: 10.1038/s41558-020-0709-0
- Pereira, J. (2018). Science and technology to enhance disaster resilience in a changing climate. *Sci. Technol. Disaster Risk Reduc. Asia* 2018, 31–38. doi: 10.1016/B978-0-12-812711-7.00003-1
- Pham, B., Tien Bui, D., Prakash, I., and Dholakia, M. (2016). Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena* 149, 52–63. doi: 10.1016/j.catena.2016.09.007
- Philbrick, C. R. Jr., and Kitanidis, P. K. (1999). Limitations of deterministic optimization applied to reservoir operations. *J. Water Resour. Plann. Manage.* 125, 135–142. doi: 10.1061/(ASCE)0733-9496(1999)125:3(135)
- Pielke, R. Jr. (2000). Flood impacts on society: damaging floods as a framework for assessment. *Floods* 1, 133–156.
- Poff, N., and Allan, J. D. (1995). Functional organization of stream fish assemblages in relation to hydrologic variability. *Ecology* 76, 606–627. doi: 10.2307/1941217
- Potnis, A. V., Durbha, S. S., and Kurte, K. R. (2018). "A geospatial ontological model for remote sensing scene semantic knowledge mining for the flood disaster," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (Valencia), 5274–5277. doi: 10.1109/IGARSS.2018.8517680
- Potnis, A. V., Shinde, R. C., Durbha, S. S., and Kurte, K. R. (2019). "Multi-class segmentation of urban floods from multispectral imagery using deep learning," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (Yokohama), 9741–9744. doi: 10.1109/IGARSS.2019.8900250
- Pourghasemi, H. R., Kariminejad, N., Amiri, M., Zarafshar, M., thomasBlaschke, and Edalat, M. (2020). Assessing and mapping multi-hazard risk

- susceptibility using a machine learning technique. *Sci. Rep.* 10:3203. doi: 10.1038/s41598-020-60191-3
- Pourghasemi, H. R., Yousefi, S., Kornejady, A., and Cerdà, A. (2017). Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci. Tot. Environ.* 609, 764–775. doi: 10.1016/j.scitotenv.2017.07.198
- Prosser, I., Young, W., Rustomji, P., Hughes, A., and Moran, C. (2001). “A model of river sediment budgets as an element of river health assessment,” in *Proceedings of the International Congress on Modelling and Simulation MODSIM2001*, 861–866.
- 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
- Rahmasary, A. N., Robert, S., Chang, I.-S., Jing, W., Park, J., Bluemling, B., Koop, S., and Van Leeuwen, K. V. L. (2019). Overcoming the challenges of water, waste and climate change in asian cities. *Environ. Manage.* 63, 520–535. doi: 10.1007/s00267-019-01137-y
- Rahmati, O., Falah, F., Dayal, K. S., Deo, R. C., Mohammadi, F., Biggs, T., et al. (2020). Machine learning approaches for spatial modeling of agricultural droughts in the south-east region of Queensland Australia. *Sci. Tot. Environ.* 699:134230. doi: 10.1016/j.scitotenv.2019.134230
- Rahmati, O., Tahmasebipour, N., Haghizadeh, A., Pourghasemi, H. R., and Feizizadeh, B. (2017). Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology* 298, 118–137. doi: 10.1016/j.geomorph.2017.09.006
- Rahmati, O., Yousefi, S., Kalantari, Z., Uemaa, E., Teimurian, T., Keesstra, S., et al. (2019). Multi-hazard exposure mapping using machine learning techniques: a case study from Iran. *Remote Sens.* 11:1943. doi: 10.3390/rs11161943
- Read, J. S., Jia, X., Willard, J., Appling, A. P., Zwart, J. A., Oliver, S. K., et al. (2019). Process-guided deep learning predictions of lake water temperature. *Water Resour. Res.* 55, 9173–9190. doi: 10.1029/2019WR024922
- Reed, P. M., Hadka, D., Herman, J. D., Kasprzyk, J. R., and Kollat, J. B. (2013). Evolutionary multiobjective optimization in water resources: the past, present, and future. *Adv. Water Resour.* 51, 438–456. doi: 10.1016/j.advwatres.2012.01.005
- Reed, P. M., and Kollat, J. B. (2013). Visual analytics clarify the scalability and effectiveness of massively parallel many-objective optimization: a groundwater monitoring design example. *Adv. Water Resour.* 56, 1–13. doi: 10.1016/j.advwatres.2013.01.011
- Refsgaard, J. C., and Henriksen, H. J. (2004). Modelling guidelines—terminology and guiding principles. *Adv. Water Resour.* 27, 71–82. doi: 10.1016/j.advwatres.2003.08.006
- Reidy Liermann, C., Nilsson, C., Robertson, J., and Ng, R. (2012). Implications of dam obstruction for global freshwater fish diversity. *BioScience* 62, 539–548. doi: 10.1525/bio.2012.62.6.5
- Rice, J., Emanuel, R., Vose, J., and Nelson, S. (2015). Continental U.S. streamflow trends from 1940 to 2009 and their relationships with watershed spatial characteristics. *Water Resour. Res.* 51, 6262–6275. doi: 10.1002/2014WR016367
- Rodriguez, F., Andrieu, H., and Creutin, J.-D. (2003). Surface runoff in urban catchments: morphological identification of unit hydrographs from urban databanks. *J. Hydrol.* 283, 146–168. doi: 10.1016/S0022-1694(03)00246-4
- Rodriguez, H., Popp, J., Maringanti, C., and Chaubey, I. (2011). Selection and placement of best management practices used to reduce water quality degradation in Lincoln lake watershed. *Water Resour. Res.* 47, 1–13. doi: 10.1029/2009WR008549
- Rowley, K. J. (2014). *Sediment transport conditions near culverts* (Master’s thesis). Brigham Young University, Provo, UT. doi: 10.1061/9780784413548.141
- Rozos, E. (2019). Machine learning, urban water resources management and operating policy. *Resources* 8:173. doi: 10.3390/resources8040173
- Ruessink, G. (2005). Calibration of nearshore process models - application of a hybrid genetic algorithm. *J. Hydroinform.* 7, 135–149. doi: 10.2166/hydro.2005.0012
- Sahoo, S., Russo, T. A., Elliott, J., and Foster, I. (2017). Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. *Water Resour. Res.* 53, 3878–3895. doi: 10.1002/2016WR019933
- Santelmann, M., Hulse, D., Wright, M., Enright, C., Branscomb, A., Tchintcharauli-Harrison, M., et al. (2019). Designing and modeling innovation across scales for urban water systems. *Urban Ecosyst.* 22, 1149–1164. doi: 10.1007/s11252-019-00882-6
- Savic, D. (2019). “What is artificial intelligence and how can water planning and management benefit from it?,” in *International Association for Hydro-Environment Engineering and research (IAHR)*, ed S. Wiecepcht (Universit’t Stuttgart). Available online at: https://www.researchgate.net/profile/Dragan_Savic3/publication/335126234_What_is_Artificial_Intelligence_and_how_can_water_planning_and_management_benefit_from_it/links/5d51795a299bf1995b78d9e4/What-is-Artificial-Intelligence-and-how-can-water-planning-and-management-benefit-from-it.pdf
- Scott, S., Burns, I., Levick, L., Hernandez, M., and Goodrich, D. (2002). Automated geospatial watershed assessment (AGWA) - a GIS-based hydrologic modeling tool: documentation and user manual. Available online at: https://www.researchgate.net/publication/253514869_Automated_Geospatial_Watershed_Assessment_AGWA_-_A_GIS-Based_Hydrologic_Modeling_Tool_Documentation_and_User_Manual
- Selin, S., and Chevez, D. (1995). Developing a collaborative model for environmental planning and management. *Environ. Manage.* 19, 189–195. doi: 10.1007/BF02471990
- Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., et al. (2020). Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. *Remote Sens.* 12:266. doi: 10.3390/rs12020266
- Shamseldin, A. Y. (2010). Artificial neural network model for river flow forecasting in a developing country. *J. Hydroinform.* 12, 22–35. doi: 10.2166/hydro.2010.027
- Sharaf El Din, E., Zhang, Y., and Suliman, A. (2017). Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. *Int. J. Remote Sens.* 38, 1023–1042. doi: 10.1080/10413116.2016.1275056
- Shaw, R., Rahman, A.-U., Surjan, A., and Parvin, G. (2016). Urban Disasters and Resilience in Asia. Elsevier: Butterworth-Heinemann. 368.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and WOO, W.-C. (2015). “Convolutional LSTM network: a machine learning approach for precipitation nowcasting,” in *Advances in Neural Information Processing Systems* (Montreal, QC), eds C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett (Curran Associates, Inc.), 802–810.
- Shi, Y., and Eberhart, R. (1998). “A modified particle swarm optimizer,” in *1998 IEEE International Conference on Evolutionary Computation Proceedings* (Anchorage, AK), 69–73. doi: 10.1109/ICEC.1998.699146
- Shimoda, A., Ichikawa, D., and Oyama, H. (2018). Using machine-learning approaches to predict non-participation in a nationwide general health check-up scheme. *Comput. Methods Prog. Biomed.* 163, 39–46. doi: 10.1016/j.cmpb.2018.05.032
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv[Preprint].arXiv:1409.1556*.
- Singh, K. P., Basant, A., Malik, A., and Jain, G. (2009). Artificial neural network modeling of the river water quality-a case study. *Ecol. Model.* 220, 888–895. doi: 10.1016/j.ecolmodel.2009.01.004
- Snelder, T., and Biggs, B. (2007). Multi-scale river environment classification for water resources management. *JAWRA J. Am. Water Resour. Assoc.* 38, 1225–1239. doi: 10.1111/j.1752-1688.2002.tb04344.x
- Souchère, V., Millair, L., Echeverria, J., Bousquet, F., Le Page, C., and Etienne, M. (2010). Co-constructing with stakeholders a role-playing game to initiate collective management of erosive runoff risks at the watershed scale. *Environ. Model. Softw.* 25, 1359–1370. doi: 10.1016/j.envsoft.2009.03.002
- Srivastava, P., Hamlett, J., Robillard, P., and Day, R. (2002). Watershed optimization of best management practices using annaGNPS and a genetic algorithm. *Water Resour. Res.* 38, 3-1-3-14. doi: 10.1029/2001WR000365
- Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., et al. (2017). A conceptual model of avalanche hazard. *Nat. Hazards.* 90, 663–691. doi: 10.1007/s11069-017-3070-5
- Sun, A., and Scanlon, B. (2019). How can big data and machine learning benefit environment and water management: a survey of methods, applications,

- and future directions. *Environ. Res. Lett.* 14:073001. doi: 10.1088/1748-9326/ab1b7d
- Sun, J., Lou, Y., and Ye, F. (2017). "Research on anomaly pattern detection in hydrological time series," in *2017 14th Web Information Systems and Applications Conference (WISA)* (Liu Zhou), 38–43. doi: 10.1109/WISA.2017.73
- Sun, Z., Zhao, X., Wu, M., and Wang, C. (2019). Extracting urban impervious surface from worldview-2 and airborne LIDAR data using 3d convolutional neural networks. *J. Indian Soc. Remote Sens.* 47, 401–412. doi: 10.1007/s12524-018-0917-5
- Taskaya-Temizel, T. and Casey, M. C. (2005). A comparative study of autoregressive neural network hybrids. *Neural Netw.* 18, 781–789. doi: 10.1016/j.neunet.2005.06.003
- Tay, F. E., and Cao, L. (2002). Modified support vector machines in financial time series forecasting. *Neurocomputing* 48, 847–861. doi: 10.1016/S0925-2312(01)00676-2
- Tayfur, G. (2002). Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.* 47, 879–892. doi: 10.1080/02626660209492997
- Tayfur, G., and Guldal, V. (2006). Artificial neural networks for estimating daily total suspended sediment in natural streams. *Nordic Hydrol.* 37, 69–79. doi: 10.2166/nh.2006.0006
- Tayyab, M., Zhou, J., Zeng, X., and Adnan, R. (2016). Discharge forecasting by applying artificial neural networks at the Jinsha river basin, china. *Eur. Sci. J.* 12, 108–127. doi: 10.19044/esj.2016.v12n9p108
- Tsintikidis, D., Haferman, J. L., Anagnostou, E. N., Krajewski, W. F., and Smith, T. F. (1997). A neural network approach to estimating rainfall from spaceborne microwave data. *IEEE Trans. Geosci. Remote Sens.* 35, 1079–1093. doi: 10.1109/36.628775
- U.N. (2018). *World Urbanization Prospects: The 2018 Edition Highlights*. Department of Economic and Social Affairs.
- UN (2010). *Clean Water for a Healthy World*. Technical report, United Nations.
- UN (2012). *Un Water Decade Program on Advocacy and Communication*. Technical report, United Nations.
- USEPA (2012). *Integrated Planning for Municipal Stormwater and Wastewater*. Available online at: <https://www.epa.gov/npdcs/integrated-planning-municipal-stormwater-and-wastewater> (accessed February 5, 2020).
- USEPA (2018). *Best Management Practices (BMPs) Siting Tool*. Available online at: <https://www.epa.gov/water-research/best-management-practices-bmps-siting-tool> (accessed February 5, 2020).
- Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdisc. Rev.* 2, 359–392. doi: 10.1002/wat2.1085
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer Verlag. doi: 10.1007/978-1-4757-2440-0
- Wagener, T., Sivapalan, M., Troch, P., McGlynn, B., Harman, C., Gupta, H., et al. (2010). The future of hydrology: an evolving science for a changing world. *Water Resour. Res.* 46, 1–10. doi: 10.1029/2009WR008906
- Wagener, T., Sivapalan, M., Troch, P., and Woods, R. (2007). Catchment classification and hydrologic similarity. *Geogr. Compass* 1, 901–931. doi: 10.1111/j.1749-8198.2007.00039.x
- Wang, X., Fu, L., and He, C. (2011). Applying support vector regression to water quality modelling by remote sensing data. *Int. J. Remote Sens.* 32, 8615–8627. doi: 10.1080/01431161.2010.543183
- Wang, X., Zhang, F., and Ding, J. (2017). Evaluation of water quality based on a machine learning algorithm and water quality index for the Ebinur Lake Watershed, China. *Sci. Rep.* 7, 1–18. doi: 10.1038/s41598-017-12853-y
- Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remote Sens. Environ.* 117, 34–49. doi: 10.1016/j.rse.2011.02.030
- Whigham, P., and Crapper, P. (2001). Modelling rainfall-runoff using genetic programming. *Math. Comput. Model.* 33, 707–721. doi: 10.1016/S0895-7177(00)00274-0
- Wieprecht, S., Habtamu, G., and Yang, C. (2013). A neuro-fuzzy-based modelling approach for sediment transport computation. *Hydrol. Sci. J.* 58, 587–599. doi: 10.1080/02626667.2012.755264
- Wischmeier, W. H., and Smith, D. D. (1978). *Predicting-Rainfall Erosion Losses: A Guide to Conservation Planning*. USDA.
- Wood, E. F., Schubert, S. D., Wood, A. W., Peters-Lidard, C. D., Mo, K. C., Mariotti, A., et al. (2015). Prospects for advancing drought understanding, monitoring, and prediction. *J. Hydrometeorol.* 16, 1636–1657. doi: 10.1175/JHM-D-14-0164.1
- Wu, J.-L., Ho, C.-R., Huang, C.-C., Srivastav, A. L., Tzeng, J.-H., and Lin, Y.-T. (2014). Hyperspectral sensing for turbid water quality monitoring in freshwater rivers: empirical relationship between reflectance and turbidity and total solids. *Sensors* 14, 22670–22688. doi: 10.3390/s141222670
- Xu, H., Demir, I., Koylu, C., and Muste, M. (2019a). A web-based geovisual analytics platform for identifying potential contributors to culvert sedimentation. *Sci. Tot. Environ.* 692, 806–817. doi: 10.1016/j.scitotenv.2019.07.157
- Xu, H., Windsor, M., Muste, M., and Demir, I. (2019b). A web-based decision support system for collaborative mitigation of multiple water-related hazards using serious gaming. *J. Environ. Manage.* 255:109887. doi: 10.1016/j.jenvman.2019.109887
- Yadav, A., Chatterjee, S., and Equeenuddin, S. (2019a). Prediction of suspended sediment yield by artificial neural network and traditional mathematical model in Mahanadi River Basin, India. *Sustain. Water Resour. Manage.* 4, 745–759. doi: 10.1007/s40899-017-0160-1
- Yadav, A., Chatterjee, S., and Equeenuddin, S. (2019b). Suspended sediment yield estimation using genetic algorithm-based artificial intelligence models: case study of Mahanadi River, India. *Hydrol. Sci. J.* 63, 1162–1182. doi: 10.1080/02626667.2018.1483581
- Yang, C., Marsooli, R., and AALAMI, M. (2009). Evaluation of total load sediment transport formulas using ann. *Int. J. Sedim. Res.* 24, 274–286. doi: 10.1016/S1001-6279(10)60003-0
- Yang, S., Ogawa, Y., Ikeuchi, K., Akiyama, Y., and Shibasaki, R. (2019). "Firm-level behavior control after large-scale urban flooding using multi-agent deep reinforcement learning," in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation* (Chicago, IL), 24–27. doi: 10.1145/3356470.3365529
- Yang, T., Sun, F., Gentine, P., Liu, W., Wang, H., Yin, J., et al. (2019). Evaluation and machine learning improvement of global hydrological model-based flood simulations. *Environ. Res. Lett.* 14:114027. doi: 10.1088/1748-9326/ab4d5e
- Yao, Y., He, J., Zhang, J., and Zhang, Y. (2017). Extracting urban impervious surface from GF-1 imagery using one-class classifiers. *arXiv[Preprint].arXiv:1705.04824*.
- Yoe, C., and Orth, K. (1996). *Planning Manual*. U.S. Army Corps of Engineering. Water Resources Support Center.
- Yu, M., Yang, C., and Li, Y. (2018). Big data in natural disaster management: a review. *Geosciences* 8:165. doi: 10.3390/geosciences8050165
- Zaidi, S. M. A., Chandola, V., Allen, M. R., Sanyal, J., Stewart, R. N., Bhaduri, B. L., et al. (2018). Machine learning for energy-water nexus: challenges and opportunities. *Big Earth Data* 2, 228–267. doi: 10.1080/20964471.2018.1526057
- Zaniolo, M., Giuliani, M., Castelletti, A. F., and Pulido-Velazquez, M. (2018). Automatic design of basin-specific drought indexes for highly regulated water systems. *Hydrol. Earth Syst. Sci.* 22, 2409–2424. doi: 10.5194/hess-22-2409-2018
- Zhang, H., Wan, L., Wang, T., Lin, Y., Lin, H., and Zheng, Z. (2019). Impervious surface estimation from optical and polarimetric sar data using small-patched deep convolutional networks: a comparative study. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 12, 2374–2387. doi: 10.1109/JSTARS.2019.2915277
- Zhang, R., Chen, Z.-Y., Xu, L.-J., and Ou, C.-Q. (2019). Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, china. *Sci. Tot. Environ.* 665, 338–346. doi: 10.1016/j.scitotenv.2019.01.431
- Zhang, Z., Xu, L., and Liu, Q. (2019). "The identification of impervious area from sentinel-2 imagery using a novel spectral spatial residual convolution neural network," in *Proceedings of the 2019 3rd International Conference on Advances in Image Processing* (New York, NY, USA: Association for Computing Machinery).
- Zounemat-Kermani, M., Matta, E., Cominola, A., Xia, X., Zhang, Q., Liang, Q., et al. (2020). Neurocomputing in surface water hydrology and hydraulics: a

review of two decades retrospective, current status and future prospects. *J. Hydrol.* 2020:125085. doi: 10.1016/j.jhydrol.2020.125085

Disclaimer: This manuscript has been authored by UT-Battelle LLC under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

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

Copyright © 2021 Allen-Dumas, Xu, Kurte and Rastogi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.