TU Delft

Delft University of Technology

Integrating Qualitative Flow Observations in a Lumped Hydrologic Routing Model

Mazzoleni, M.; Amaranto, A.; Solomatine, D. P.

DOI
10.1029/2018WR023768

Publication date
2019

Document Version
Final published version

Published in
Water Resources Research

Citation (APA)
Mazzoleni, M., Amaranto, A., & Solomatine, D. P. (2019). Integrating Qualitative Flow Observations in a Lumped Hydrologic Routing Model. *Water Resources Research*, *55*(7), 6088-6108. https://doi.org/10.1029/2018WR023768

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR023768

Key Points:
- Novel approaches used in robotics and missile tracking can be used to integrate qualitative observations in hydrological applications
- The integration of qualitative observations can improve flow estimation
- Whereas state updating provides the best model results, output correction approaches are not affected by model error and observation bias

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4

Correspondence to:
M. Mazzoleni,
maurizio.mazzoleni@geo.uu.se

Citation:
Mazzoleni, M., Amaranto, A., & Solomatine, D. P. (2019). Integrating qualitative flow observations in a lumped hydrologic routing model. *Water Resources Research*, *55*, 6088–6108. https://doi.org/10.1029/2018WR023768

Received 24 JUL 2018
Accepted 24 JUN 2019
Accepted article online 9 JUL 2019
Published online 26 JUL 2019

©2019. American Geophysical Union. All Rights Reserved.

Integrating Qualitative Flow Observations in a Lumped Hydrologic Routing Model

M. Mazzoleni[1,2], A. Amaranto[3], and D.P. Solomatine[4,5,6]

[1]Department of Earth Sciences, Uppsala University, Uppsala, Sweden, [2]Centre of Natural Hazards and Disaster Science, Uppsala, Sweden, [3]Department of Electronics, Information, and Bioengineering, Piazza Leonardo da Vinci, Milano, Italy, [4]Hydroinformatics Chair Group, IHE Delft Institute for Water Education, Delft, The Netherlands, [5]Water Resources Section, Delft University of Technology, Delft, The Netherlands, [6]Water Problems Institute of RAS, Moscow, Russia

Abstract This study aims at proposing novel approaches for integrating qualitative flow observations in a lumped hydrologic routing model and assessing their usefulness for improving flood estimation. Routing is based on a three-parameter Muskingum model used to propagate streamflow in five different rivers in the United States. Qualitative flow observations, synthetically generated from observed flow, are converted into fuzzy observations using flow characteristic for defining fuzzy classes. A model states updating method and a model output correction technique are implemented. An innovative application of Interacting Multiple Models, which use was previously demonstrated on tracking in ballistic missile applications, is proposed as state updating method, together with the traditional Kalman filter. The output corrector approach is based on the fuzzy error corrector, which was previously used for robots navigation. This study demonstrates the usefulness of integrating qualitative flow observations for improving flood estimation. In particular, state updating methods outperform the output correction approach in terms of average improvement of model performances, while the latter is found to be less sensitive to biased observations and to the definition of fuzzy sets used to represent qualitative observations.

1. Introduction

Severity and frequency of extreme flood events are increasing worldwide (Willner et al., 2018). Advanced mathematical models are used to simulate and predict flood events on urbanized areas, but this is hindered by the lack of gauge data with proper spatial and temporal resolution. On the other hand, remote sensing and, more recently, citizen science initiatives provide useful data for hydrological purposes (Buytaert et al., 2014). The use of crowdsourced observations for hydrological modeling applications has rapidly increased in the last years (Assumpção et al., 2018; de Vos et al., 2017; Le Coz et al., 2016; Walker et al., 2016; Yu et al., 2016; Yang & Ng, 2017). In particular, different studies have been carried out for integrating quantitative expert-based and citizens' observations within hydrological and hydraulic models to improve model calibration and validation in ungauged basins (Seibert & McDonnell, 2002; van Meerveld et al., 2017), to better characterize the spatial and temporal dimensions of a catchment hydrological processes (Starkey et al., 2017; Weeser et al., 2018), and to improve flood predictions (Etter et al., 2018; Mazzoleni et al., 2015, Mazzoleni, Verlaan, et al., 2017; Mazzoleni, Cortes Arevalo, et al., 2018).

Besides quantitative crowdsourced observations, the growing availability of qualitative flow information, retrieved, for example, from social media, is opening more opportunities to complement standard sources of information from gauges or remote sensors (Eilander et al., 2016). In fact, information from pictures (e.g., Fohringer et al., 2015; Giuliani et al., 2016; Smith et al., 2015; Starkey et al., 2017), videos (e.g., Aronica et al., 1998; Le Boursicaud et al., 2016; Le Coz et al., 2016; Michelsen et al., 2016), or opinions shared within social networks (e.g., Flickr and Twitter) have been increasingly and successfully used in combination with traditional measurements to better identify in (near) real-time flood extent areas (Brouwer et al., 2017; Eilander et al., 2016; Fohringer et al., 2015; Li et al., 2017; Schnebele et al., 2014; Smith et al., 2015; Sun et al., 2015). Among these pioneering studies, Rosser et al. (2017) proposed a data fusion method to rapidly estimate flood inundation extent by combining observations from remote sensing, social media, and high-resolution terrain mapping. In Restrepo-Estrada et al. (2018), social media messages and rainfall measurements from authoritative sources are used to create a proxy variable for rainfall, which is then used as input to a hydrological and hydraulic model for improving flow estimation. Wang et al. (2018) used social

media and crowdsourcing to complement the traditional data sets and remote sensing data to properly assess flood extent. The authors pointed out that Twitter data seem to be weakly correlated with rainfall for flood modeling purposes.

Despite these recent studies, only few research activities assessed the usefulness of qualitative observations for hydrological and/or hydraulic modeling purposes. For example, Seibert and McDonnell (2002) developed an approach to calibrate hydrological models using both quantitative and qualitative data provided by expert users. Kovitz and Christakos (2004) assimilated fuzzy (i.e., qualitative) data sets assigning probabilities of plausible events based on general knowledge through information maximization and then applying Bayesian maximum entropy method. Ross et al. (2008) developed a fuzzy Kalman filter in which the initial estimate of the hydraulic conductivity field is based on expert judgment. Starkey et al. (2017) achieved substantial improvement in the spatial and temporal characterization of the catchment response when traditional network is complemented with community-based qualitative and quantitative observations (citizens' readings of water level gauge, pictures, and videos collected from Twitter, and other crowdsourced sources as email, website, and mobile app) for modeling purposes. Despite the encouraging results, these studies showed that expert knowledge can be used to provide a preliminary information of a given hydrological variable (e.g., hydraulic conductivity in Ross et al., 2008) to then be updated by more reliable and precise data from traditional sensors.

However, none of the previous studies jointly used qualitative observations and hydrological models for improving river flow assessment. For this reason, the aim of this research is to improve flood estimation by developing and testing different approaches for updating model states and output by integrating qualitative flow observations into a lumped hydrological model. In particular, Kalman filter (KF) and Interacting Multiple Model (IMM), which was used in missile tracking applications (Bar-Shalom et al., 2001; Cooperman, 2002), are used as model state update methods, while a fuzzy error corrector used in robot tracking (Moulton et al., 2001) is used as the model output correction method. These approaches are applied in five different river reaches in the United States, for a total of 100 flood events of various durations.

2. Methodology

2.1. From Qualitative Information to Fuzzy Flow Observations

Qualitative observations provide linguistic information about a certain physical variable. For example, during a flood event, information of *high water levels* or *intense precipitation* may be shared by people on social networks (as reported in Figure 1). However, the underpinning assumption of any traditional mathematical model implementation is its requirement of quantitative data. Therefore, for a qualitative information to be integrated into a model, it needs to be converted into a quantitative value first.

Usually, integration of quantitative measurements within mathematical models requires indication of the observations accuracy, expressed as probability distribution. However, a qualitative information can be also described in possibilistic terms rather than probabilistic, that is, by the language of fuzzy logic, and there are multiple examples of this in various areas. The methodology we are proposing here is based indeed on the use of fuzzy logic, rather than probabilistic approaches, to represent uncertainty in qualitative flow observations.

Examples of fuzzy descriptors of qualitative observations are reported, for example, by Ross et al. (2008) and Salinas et al. (2015), who converted linguistic characteristics of hydraulic conductivity from expert users (Ross et al., 2008) and historical qualitative flow (Salinas et al., 2016) into fuzzy observations by means of specific values of membership functions (similar to the ones represented in Figure 2). These membership functions may be wider or narrower depending on the vagueness of the statements.

In our study, flow is categorized based on predefined five-class fuzzy sets (dry, low, medium, high, and extreme flow) according to the statistical characteristics of the historical observed flow values. The feasibility interval of the five classes is represented by the 0.01 and 0.99 quantiles $q_{0.01}$ and $q_{0.99}$ (see Figure 2). Fuzzy sets are defined as equally spaced within the interval. The membership to each category is characterized by a trapezoidal or triangular functions with support (base of the triangular function) estimated as a linear function between the difference of two adjacent fuzzy cores and coefficient α (e.g., L^{up} in Figure 2). A higher value of the α coefficient corresponds to a bigger overlap between adjacent fuzzy sets (i.e., higher

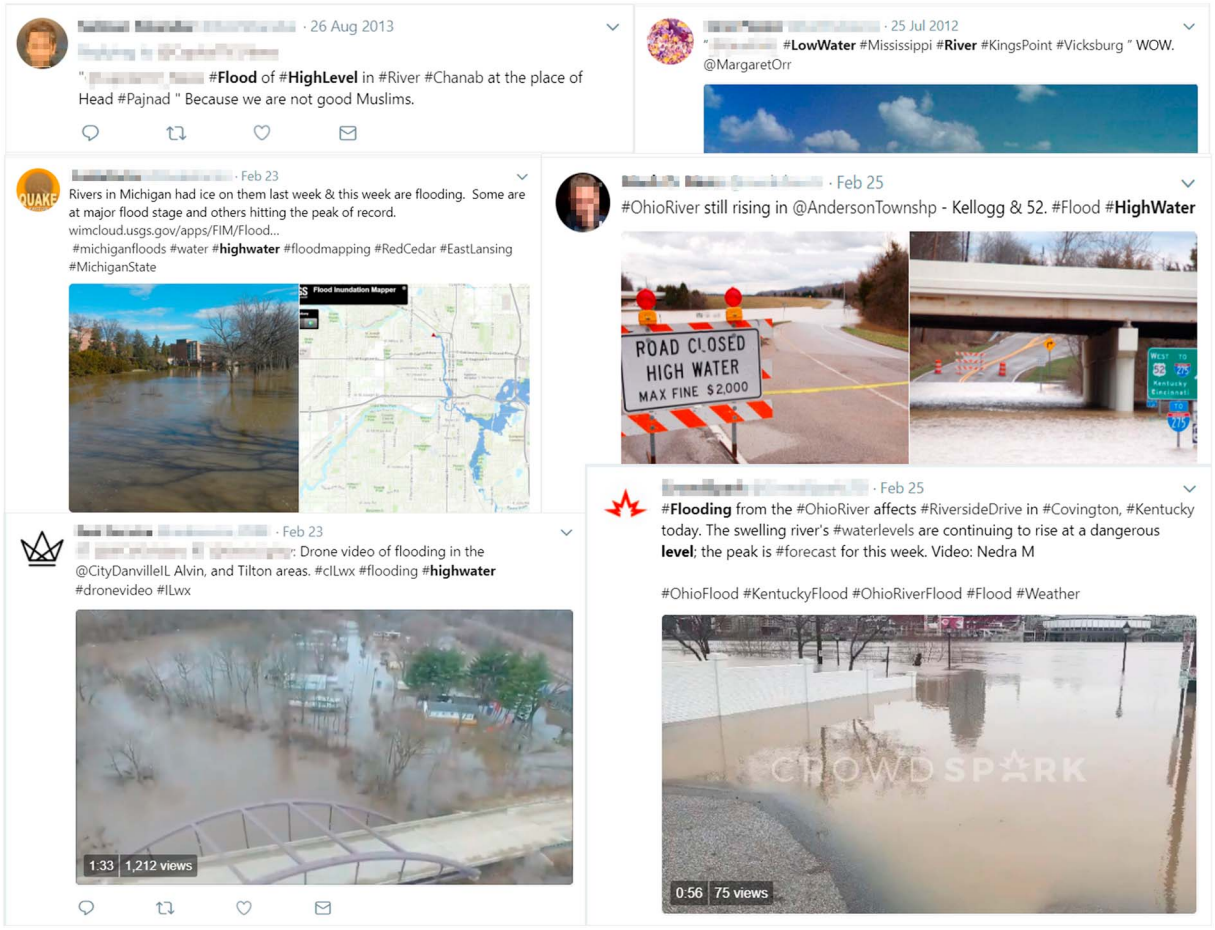

Figure 1. Example of tweets reporting qualitative flow observations from several random Twitter accounts.

uncertainty). We considered two different coefficients α, α_O and α_M used to identify the fuzzy class support to the observations and the model output. The different values of α_O and α_M are described in section 3. Moreover, D, L, M, H and E are the fuzzy cores of the Dry, Low, Medium, High, and Extreme flow fuzzy classes, respectively. The 0.01 and 0.99 quantiles are used to define the dry and extreme core values of the flow sets and indirectly also the low, medium, and high sets.

Based on the fuzzy sets illustrated in Figure 2, we considered two possible alternatives to assign a quantitative value to qualitative flow information. In the first approach, the qualitative information is converted into

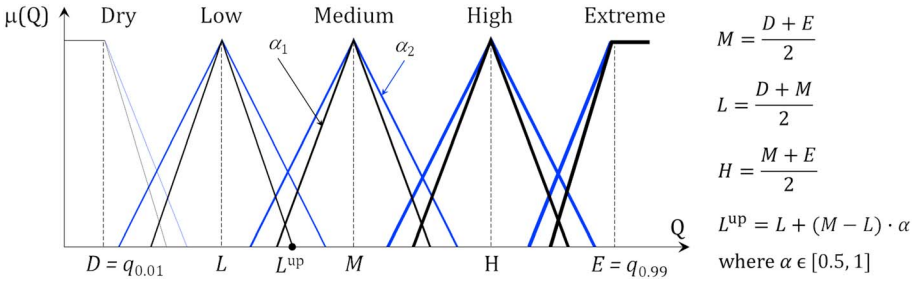

Figure 2. Classification of the five fuzzy sets used to classify qualitative river flow classes (dry, low, medium, high, and extreme) in case of two different hypothetical α values, where $\alpha_1 < \alpha_2$ represents the fuzzy support and $\mu(Q)$ is the membership function. Calculation of the fuzzy cores D, L, M, H, and E and upper extreme of the low fuzzy set class L^{up} is presented in the right part of the figure.

a deterministic value represented by the core value of the considered fuzzy set. However, in case of qualitative observations the observed true flow value might be any within the assigned fuzzy set. For this reason, in the second approach, an ensemble of flow observations is randomly generated within the range of the fuzzy set class in which the qualitative observation falls. For each ensemble member, the membership function value is actually treated as probability—to be further used in calculations by the data integration procedures.

2.2. Hydrologic Routing Model

The hydrologic routing method implemented in this study is the three-parameter Muskingum model (Musk3p) developed by O'Donnell (1985). This model is based on the same assumptions of the standard Muskingum method, with the difference that an additional model parameter is included to account for possible distributed lateral flow along the river reach. As described in O'Donnell (1985), in case of distributed lateral flow the Muskingum model can be represented as

$$Q_{t+1}^{D} = C_1 Q_t^{U} + C_2 Q_{t+1}^{U} + C_3 Q_t^{D} \tag{1}$$

where Q^{U} and Q^{D} are the upstream and downstream flows (model output), t is the time index, and C_1, C_2, and C_3 are the model coefficients, which are estimated as

$$\begin{aligned} C_1 &= (1+\varepsilon)\cdot\frac{0.5\cdot\Delta t + K\cdot X}{K\cdot(1-X) + 0.5\cdot\Delta t} \\ C_2 &= (1+\varepsilon)\cdot\frac{0.5\cdot\Delta t - K\cdot X}{K\cdot(1-X) + 0.5\cdot\Delta t} \\ C_3 &= \frac{K\cdot(1-X) - 0.5\cdot\Delta t}{K\cdot(1-X) + 0.5\cdot\Delta t} \end{aligned} \tag{2}$$

where Δt is the model time step (hours), X is a weighting factor, K is the parameter function of the storage constant (hours), and ε is a parameter representing the distributed lateral inflow. If no lateral inflow is present, ε equals zero and the coefficients in equation 2 are equivalent to the ones proposed in the standard formulation of the Muskingum model (Todini, 2007).

The state-space form of Musk3p can be expressed by means of states and measurements equations. The model states at time $t+1$, that is, the streamflow Q, are expressed as a function of the model states at time t based on the following standard system model formulations typically used in data assimilation (Georgakakos et al., 1990):

$$\boldsymbol{x}_{t+1} = \boldsymbol{\Phi}\boldsymbol{x}_t + \boldsymbol{\Gamma}\boldsymbol{I}_t + \boldsymbol{w}_t \tag{3}$$

where $\mathbf{x}$ is the $n_{\text{state}} \times 1$ model state matrix representing the calculated streamflow at each river reach, n_{state} is the number of reaches into which the river is divided, $\mathbf{I}_t$ is the upstream boundary conditions vector (defined as $[Q_t^u, Q_{t+1}^u]$), and $\mathbf{w}_t$ is the white noise vector (a diagonal matrix $n_{\text{state}} \times n_{\text{state}}$) with zero mean and covariance $\mathbf{S}$. The matrix $\boldsymbol{\Phi}$ ($n_{\text{state}} \times n_{\text{state}}$) is the state-transition matrix, and $\boldsymbol{\Gamma}$ ($n_{\text{state}} \times 2$) is the input-transition matrix. Both $\boldsymbol{\Phi}$ and $\boldsymbol{\Gamma}$ are functions of the three model coefficients C_1, C_2, and C_3, derived based on equation 1, following Georgakakos et al. (1990) in the general case of $n_{\text{state}} \geq 1$.

The corresponding measurement equation representing the flow along the river channel, is expressed as

$$\boldsymbol{z}_{t+1} = \boldsymbol{H}_t\boldsymbol{x}_{t+1} + \boldsymbol{v}_t \tag{4}$$

where $\mathbf{z}_t$ is the $n_{\text{obs}} \times 1$ observation matrix representing the streamflow at the observation location, $\mathbf{H}$ is the $n_{\text{obs}} \times n_{\text{states}}$ output matrix, and $\mathbf{v}_t$ is the measurement noise represented by normal distribution with zero mean and covariance $\mathbf{R}$.

In this study, a lumped version of the Musk3p model is applied on all river reaches. Therefore, n_{states} is set equal to 1 and consequently the model states matrix corresponds to the scalar streamflow value at the outlet of the river reach. For this reason, the matrix $\mathbf{H}$ is set equal to 1. It is worth noting that a distributed model structure may allow to provide more information regarding flow propagation for long river reaches. However, a comparison between integration of flow observations in lumped and distributed model

structures, using different model updating methods, has been already carried out in Mazzoleni, Alfonso, et al. (2017) and Mazzoleni, Chacon-Hurtado, et al. (2018). In those studies, the authors also accounted for other sources of uncertainty, related to estimation of model error covariance matrix and to location of the assimilation point. Because of this, we decided to test our new approaches using the lumped model structure and thus avoid other sources of uncertainty (e.g., related to model calibration and to choice of model spatial and temporal time steps), which may additionally affect assimilation performances. In addition, we believe that the use of a lumped model is legitimate due the fact that the flow observations are assimilated only at the river outlet and not at internal points.

2.3. Model Updating

Model updating methods are used for integrating the real-time observations within mathematical models and thus to correct model input, states, parameters, or outputs to reduce model uncertainty (Lahoz et al., 2010; Liu et al., 2012; Refsgaard, 1997; WMO, 1992). In this study, states updating and output correction methods are used to improve flow estimation. This is done by integrating hourly (model updating frequency) fuzzy flow observations derived from qualitative information. In case of hydrologic routing methods, regardless of the lumped or distributed structure, river flow (which is the observed and assimilated variable) corresponds to the state of the model. For this reason, the integration of flow observations would update model states using both output correction and model states techniques. Therefore, an updated estimate of the initial state can be used to initialize flow prediction. However, in this study no flow prediction is provided. It is worth noting that output correction methods applied to a lumped hydrological model (e.g., HBV model) would only provide correction of the output at the current time step while no update will be performed to the model states.

2.3.1. States Update

One method we used to update model states is the KF (Kalman, 1960). KF is a mathematical tool that allows for estimating the model states of a linear system, in an efficient optimal recursive way, as the response to new available hydrological noisy observations (Heemink & Segers, 2002; Kalman, 1960; Liu & Gupta, 2007; McLaughlin, 2002; Walker & Houser, 2005). The KF method can be divided into two parts: the update (or analysis) and the forecast (or background). The update equations are generally expressed as

$$\boldsymbol{K}_t = \frac{\boldsymbol{P}_t^- \boldsymbol{H}^T}{\boldsymbol{H}\boldsymbol{P}_t^- \boldsymbol{H}^T + \boldsymbol{R}_t} \tag{5}$$

$$\boldsymbol{x}_t^+ = \boldsymbol{x}_t^- + \boldsymbol{K}_t \cdot \left(\boldsymbol{Q}_t^O - \boldsymbol{H}\boldsymbol{x}_t^-\right) \tag{6}$$

$$\boldsymbol{P}_t^+ = (\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{H})\boldsymbol{P}_t^- \tag{7}$$

where $\mathbf{K}$ is the $n_{\text{states}} \times n_{\text{obs}}$ Kalman gain matrix, $\mathbf{P}$ is the $n_{\text{states}} \times n_{\text{states}}$ error covariance matrix, and $\mathbf{Q}^{\text{O}}$ is the real-time vector of flow observations. Because of the lumped nature of the Musk3p model, $\mathbf{Q}^{\text{O}}$, $\mathbf{K}$, and $\mathbf{P}$ are reduced to scalars.

The forecast equations use the system model formulation in equation 3 and are as follows:

$$\boldsymbol{P}_{t+1}^- = \boldsymbol{\Phi}\boldsymbol{P}_t^+\boldsymbol{\Phi}^T + \boldsymbol{\Gamma}\boldsymbol{M}_b\boldsymbol{\Gamma}^T + \boldsymbol{S}_t \tag{8}$$

where $\mathbf{M}_{\text{b}}$ is the 2×2 covariance matrix of the boundary conditions matrix $\mathbf{I}$ of equation 3, the superscript $^+$ indicates the updated states values, while the superscript $^-$ represents the background estimates. An important aspect of the correct implementation of KF is the proper estimation of model covariance matrix $\mathbf{S}$. An overview on the methods used to calculate model error is given by Maybeck (1982). The value of the $n_{\text{states}} \times n_{\text{states}}$ model covariance matrix $\mathbf{S}$ is estimated as

$$\boldsymbol{S}_t = (\gamma \cdot \boldsymbol{x}_t)^2 = \gamma^2 \boldsymbol{x}_t \boldsymbol{x}_t^T \tag{9}$$

where γ is the coefficient that may take different values depending on the assumptions about the model error; they are reported in section 2.4. Similarly, the $n_{\text{obs}} \times n_{\text{obs}}$ observational error matrix can be estimated as

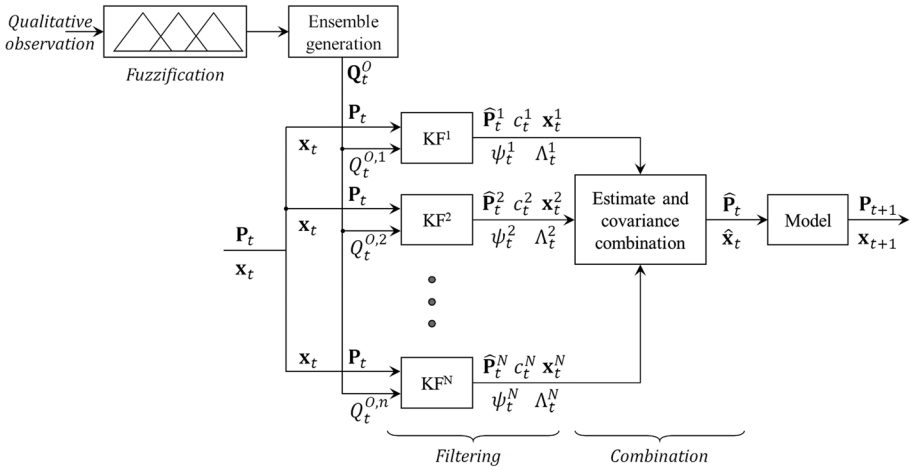

Figure 3. Schematic representation of the modified IMM procedure with both filtering and combination steps reported. In particular, P indicates the error covariance matrix, **x** is the model state, $\mathbf{Q}^O$ the ensemble of fuzzy observations, **c** is a normalizing factor, **Λ** the likelihood of the observations, **ψ** is the probability of each filter KF, where the superscript number indicates the ensemble member. IMM = Interacting Multiple Model; KF = Kalman filter.

$$\boldsymbol{R}_t = \left(\beta \cdot \boldsymbol{Q}_t^O\right)^2 = \beta^2 \boldsymbol{Q}_t^O \boldsymbol{Q}_t^{O,T} \tag{10}$$

where β is the coefficient related to observation error ranging between 0 and 1, as defined later in section 3, and T indicates the transposed matrix. Because of the lumped structure of the model, the matrices **S** and **R** are considered to be scalar values.

The presented procedure is applied for the integration of one single fuzzy observation (corresponding to the core value of the flow fuzzy class) each hourly model time step within the Musk3p. On the other hand, in case of the ensemble of fuzzy flow observations it would be natural to use multiple updating filters for each member of the fuzzy observations ensemble. We decided to modify the IMMs approach (originally proposed used by Bar-Shalom et al., 2001 and Cooperman, 2002) to integrate each observation within the Musk3p model. In his study, Cooperman (2002) developed a tactical ballistic missile tracker within an IMM framework in order to fuse measurements from multiple sensors in the three different models representing the boost, the exo-atmospheric, and the endo-atmospheric phases of a flight. An IMM framework can be divided into the three main parts: (a) interaction, (b) filtering, and (c) combination. The initial condition for each model is obtained from the state estimates of the filter at previous time steps. Then, multiple model updating methods (e.g., KF) are applied to the models to update them as a response to the observed data. Finally, a weighted combination of the updated states and models covariance matrices is performed yielding the combined estimation of **x** and **P**.

However, we did not apply IMM in case of multiple models but in case of single model and multiple observations. For this reason, in this modified version of the IMM the updating part of the assimilation process (equations 5–7 for the KF) consists only in the filtering and combination phases (Figure 3), while the interaction phase is not considered.

The first step in our approach is the implementation of the observations ensemble. The ensemble of fuzzy flow observations $\mathbf{Q}^O$ is randomly generated at each time step, by sampling from the minimum and maximum value of the membership function range in which an observation falls. During the updating phase, we used the filtering equations 5–7 to assimilate each flow ensemble member. Then, in addition to the updated values of **x** and **P**, we also estimated the likelihood of the observations for each filter as

$$\Lambda_t^i = N\left(v_t^i, 0, S_t^i\right) \tag{11}$$

where v is the observation error. Then, the probability of each filter i and time step t is estimated as

$$\psi_t^i = \frac{1}{c}\Lambda_t^i \cdot \mu(Q)_t^i \tag{12}$$

where c a normalizing factor calculated as

$$c = \sum_{i=1}^{N_{ens}} \Lambda_t^i \cdot \mu(Q)_t^i \tag{13}$$

where N_{ens} is the ensemble number and $\mu(Q)$ is the membership function value for the giver ensemble member i.

In the combination part, the combined values of the states and covariance matrices are estimated as

$$\widehat{\boldsymbol{x}}_t = \sum_{i=1}^{N_{ens}} \psi_t^i \cdot \boldsymbol{x}_t^i \tag{14}$$

$$\widehat{\boldsymbol{P}}_t = \sum_{i=1}^{N_{ens}} \psi_t^i \times \left\{ \boldsymbol{P}_t^i \left[\boldsymbol{x}_t^i - \widehat{\boldsymbol{x}}_t\right]\left[\boldsymbol{x}_t^i - \widehat{\boldsymbol{x}}_t\right]^T \right\} \tag{15}$$

The updated values of **x** and **P** are then used for the next model time step where the procedure is repeated for the new observations ensemble (see Figure 3). It is worth mentioning that ensemble Kalman filter (EnKF) could also be considered as a possible state updating method instead of IMM as both can be applied in case of an ensemble of observations. However, due to (a) the type of probabilistic distribution of the flow ensemble, (b) the linear nature of the Musk3p model, and (c) applicability of IMM for multiple independent flow observations, we decided to implement the IMM method instead of EnKF.

The first reason is related to the type of probabilistic distribution used in this study to convert the fuzzy information in an ensemble of quantitative flow observations. We adopted a uniform distribution for the observations ensemble. However, in case of qualitative observations the ensemble could be associated with any type of probabilistic distribution, which will result in a nonoptimal implementation of the EnKF. Second, EnKF is mainly used in case of highly nonlinear models while the Musk3p model is linear. Last but not least, IMM can be applied in case of multiple flow observations that are independently taken and are not generated as ensemble of a single flow information as for EnKF.

2.3.2. Output Correction

One disadvantage of the state updating methods is the need to quantify the model covariance matrix **S** (Liu et al., 2012) which in many practical situation is hardly possible. Therefore, the second approach that we implemented is the output correction, in which no information regarding model covariance matrix is required. The approach we proposed to assimilate qualitative streamflow observations is based on the ideas presented in the study of Moulton et al. (2001), where a Fuzzy Error Correction (FEC) control system is used to navigate a robot along a path of magnetic disks.

In Moulton et al. (2001), the purpose of the robot is to follow a path defined by a sequence of ferromagnetic disks located on the floor. It is suggested that the robot can better follow smooth path, while the path of ferromagnetic disks can assume any shape. At the bottom of the robot the array of Hall sensors is located, positioned between the two driving wheels, used to detect the metal disks. The direction in which the robot moves defines the azimuth of the robot, while the Hall arrays defines the x axis. Each time that the robot passes over a ferromagnetic disk, one or two sensors of the Hall arrays produce an electric signal that identifies the position of the disk on the x axis of the robot. The FEC is used to correct the trajectory of the robot based on the position of the magnetic disk with respect to the array of Hall sensors. The FEC control system has two inputs: (a) the signal produced by the Hall arrays to identify the location of the disk, and (b) the azimuth of the robot indicating its current direction. The output of the FEC control system is the angle of rotation, that is, the error correction of the robot to be added to the azimuth to get the updated direction of the robot. Ideally, the robot would move directly to the disk in case of steady path. However, because of the random location of the ferromagnetic disks the robot will have to adjust its position each moment.

It is suggested that the described approach by Moulton et al. (2001) for robot navigation can be reformulated for the task of simulated flow correction. In fact, similar to the robot (which has to follow the path made by ferromagnetic disks), the simulated hydrograph (calculated with the hydrologic model) needs to move

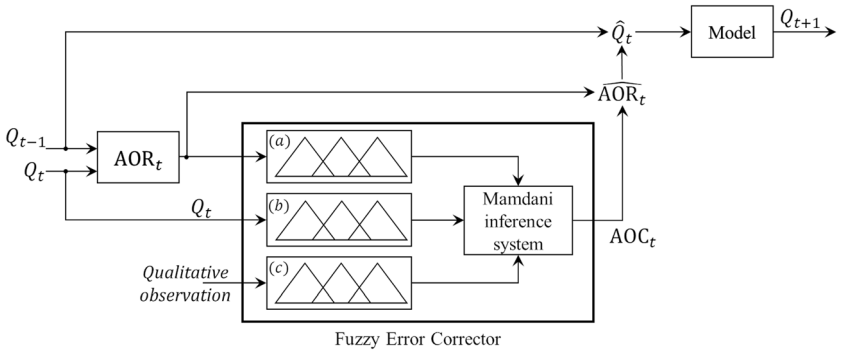

Figure 4. Schematic representation of the output updating method, where Q is the simulated flow, AOR is the angle of rotation, AOC is the angle of correction, and $\widehat{\mathrm{AOR}}$ is the updated angle of rotation.

forward the qualitative flow observations as well. The position of the robot can be seen as the hydrologic model at a particular time step, the robot trajectory as the simulated flow hydrograph, while the magnetic disk as the observed flow value.

The idea behind our approach is to correct the direction of the simulated flow hydrograph at the time step t as a response to the observed qualitative observations using the correction factor in a similar fashion to the one proposed by Moulton et al. (2001). The only difference is that, while the robot does not have an original trajectory, the simulated flow has a trajectory defined by the lumped hydrologic routing model. In addition, the robot can potentially move in any physical direction, while the simulated flow hydrograph can only move forward positive time directions.

The first step for the implementation of the output updating method is the estimation of the angle of rotation (AOR) of the simulated flow hydrograph (see Figures 4 and 5) calculated as

$$\mathrm{AOR} = \tan\left(\frac{Q_t^D - Q_{t-1}^D}{dt}\right) \tag{16}$$

The AOR corresponds to the azimuth direction in the robot application proposed by Moulton et al. (2001), which can assume values only between the −90° and + 90° range. The preliminary step in the FEC method is the fuzzification of the AOR together with the model output at time t, indicated with a and b in Figure 4. It is worth noting that the fuzzification of the qualitative observation (named c in Figure 4) is already performed in the previous steps. In particular, the fuzzification of the model output and qualitative observations at time step t is performed using the fuzzy sets defined in section 2.2 for different values of the α coefficient reported in section 3. Regarding the AOR, seven fuzzy sets, Large Negative, Medium Negative, Small Negative, Zero, Small Positive, Medium Positive, and Large Positive are identified and shown in Figure S1c of the supporting information. Similarly, we defined the AOC to be between −179 and +179° and 13 fuzzy sets are introduced. In addition to the previous seven fuzzy sets, the other six sets, Large Negative Extreme, Medium Negative Extreme, Small Negative Extreme, Small Positive Extreme, Medium Positive Extreme, and Large Positive Extreme, are considered for the AOC. These extra sets are introduced to account for situations in which a model, observations, and AOR are in disagreement with each other. An example of this situation is described below and reported in Figure 5.

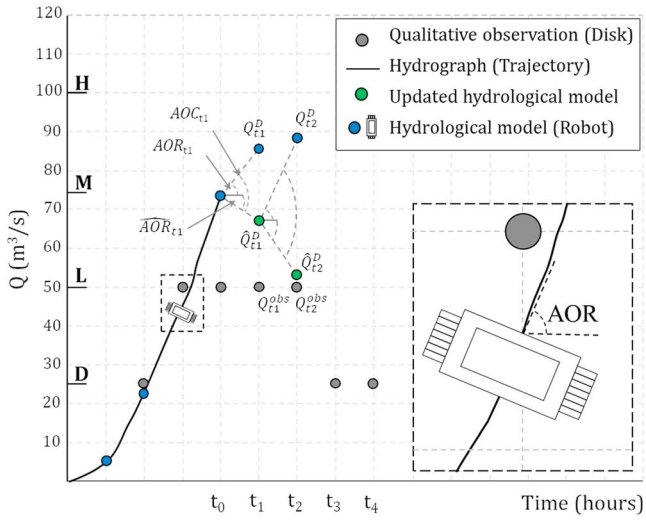

Figure 5. Left side: graphical representation of the analogies between flood prediction problem and robot tracking problem with indication of the main variables used in the explanatory example of the error correction method and reported in Table 1.

The three fuzzified variables are then fed into a Mamdani's fuzzy inference system (Mamdani & Assilian, 1975), with the centroid-based defuzzification, to calculate the angle of correction (AOC) as reported in Figure 4.

Table 1
Calculation Results Obtained When Applying Error Correction in the Four Time Steps Example Showed in Figure 5, Where Model Input, Parameters, and Error Coefficients Are $Q_{t0}^{U} = 45\ m^3/s$, $Q_{t0}^{D} = 75\ m^3/s$, $K = 2.59$ hr, $X = 0.14$, $\varepsilon = 0.02$, $\alpha_M = 0.55$, and $\alpha_O = 0.75$

	t_1	t_2	t_3	t_4
Q_t^{obs} (m^3/s)	50.0	50.0	25.0	25.0
Q_t^{U} (m^3/s)	85.0	75.0	60.0	55.0
Q_t^{D} (m^3/s)	85.8	89.3	78.1	68.5
AOR_t (degrees)	47.3	64.3	69.2	71.9
AOC_t (degrees)	−79.9	−123.5	−119.9	−119.9
$\widehat{AOR}_t$ (degrees)	−32.7	−59.3	−54.4	−47.9
$\widehat{Q}_t^{D}$ (m^3/s)	68.6	51.8	37.8	26.7
Error (m^3/s)	18.6	1.8	12.3	1.7

Note. AOR = angle of rotation; AOC = angle of correction; $\widehat{AOR}$ = updated angle of rotation.

As expected, a crucial part of the proposed method is the proper assessment of the fuzzy rule in the Mamdani inference method. The 175 rules used in the Mamdani inference method are reported in Table S1 as supporting information. For each combination of inputs, there is an associated AOC output value, based on the amount of correction necessary to keep the flow hydrograph trajectory close to observations. It can be noted that there is a certain amount of symmetry present in the fuzzy rules used to assess AOC. An important assumption made is that when both the model and observations are in the same fuzzy set, no correction is made. Moreover, the farther the observation is from the model class, depending on the model gradient, the larger the AOC will be.

Once the AOC is calculated from the FEC, it is then used to correct the AOR at time step t (and as a consequence also the flow hydrograph trajectory) and estimate its updated value as

$$\widehat{AOR} = AOR + AOC \tag{17}$$

Based on the updated AOR value, the updated model output $\widehat{Q}_t$ is assessed applying basic trigonometric computations using the model output Q at time $t-1$

$$\widehat{Q}_t^{D} = Q_{t-1}^{D} + \mathrm{atan}\left(\widehat{\mathrm{AOR}}\right)\cdot \mathrm{d}t \tag{18}$$

Finally, the model output at $t+1$ is calculated using $\widehat{Q}_t$ within equations 3 and 4.

In order to strengthen the explanation of the output correction approach, a small example is presented in Figure 5 and Table 1.

In this example, we assumed a model time step of 10 hr, while in the other experiments we considered hourly time step. Hypothetical values of model inputs, parameters, and error coefficients are reported in Table 1.

In addition, we assumed four different qualitative flow observation values belonging to the low and dry flow class (see Figure 5). The Musk3P model of equation 3 is used to calculate the value of Q^D at time step t_1, forcing the model with the information of upstream and downstream simulated flow at t_0. Then, the AOR at t_1 is calculated using equation 16. Knowing the value of AOR, simulated flow, and qualitative flow observations, it is possible to use them as input in the FEC. Their corresponding fuzzy classes are reported in Figure S1 and derived from the values α_O, α_M, and fuzzy set cores.

From Table 1 it can be observed that simulated flow Q_{t1}^{D} tends to overestimate the observed fuzzy value Q_{t1}^{obs} and has a high value of AOR in the range of Medium Positive (angle). In order to correct the model trajectory toward the lower values of flow, the resulting AOC provides a negative correction value to force the model toward the low flow. The derived AOC value is then used to correct the value of AOR, using equation 17, and the consequent corrected value of $\widehat{Q}_{t1}^{D}$ (see equation 18). Finally, $\widehat{Q}_{t1}^{D}$ is used in equation 3 to calculate the corrected value of $\widehat{Q}_{t2}^{D}$. This sequence is repeat for each time step in which a flow observation is available.

The aforementioned output correction procedure refers to the case of qualitative observation being equal to the core of the corresponding fuzzy set. When an ensemble of observations is considered, the FEC procedure is applied to each ensemble member of the observation vector $\mathbf{Q}^{O}$. The corresponding model output is then calculated as the arithmetic average of the ensemble model output. In the following, the ensemble variant of the FEC method is called EnsFEC.

One of the limitations of the output correction methods is the lack of a spatial error covariance for a distributed structure of the Musk3p model. In fact, once a given flow observation is integrated at a specific location of the river, the updating benefits come from two aspects: (a) from the downstream propagation of the updated river flow following the state-space equation of the Musk3p model (see equation 3); and (b) from a distributed update along the river reach due to the spatial characteristic of the covariance error (e.g., the

Table 2
Summary of the Main Features of the State Update and Error Correction Methods Used to Integrate Qualitative Flow Observations

Feature	KF	IMM	FEC	EnsFEC
Updated variables	Model states	Model states	Model output	Model output
Applicable to	Linear models	Linear models	Any model	Any model
Type of errors	Observation (β) and model (γ) errors	Observation (β) and model (γ) errors	Observation (α_O) and model (α_M) errors	Observation (α_O) and model (α_M) errors
Type of update	Recursive	Recursive	Recursive (in hydrologic routing models)	Recursive (in hydrologic routing models)
Type of observation error	Normal error	Normal error	Fuzzy sets	Fuzzy sets
Ensemble based	No	Yes	No	Yes

Note. KF = Kalman filter; IMM = Interacting Multiple Model; FEC = Fuzzy Error Correction; EnsFEC = ensemble variant of the FEC method.

Kalman gain matrix in case of IMM and KF). For FEC, only the state update due to the downstream flow propagation of equation 3 is carried out. This will result in a lower model update at the downstream or upstream location of the assimilation point. Another limitation would emerge if the output correction method would be used to more complex hydrological models, in which the output updating methods would only provide a correction of the output while no update will be made to the model states (as in case of hydrologic routing models). A summary of the features of the state updating and output correction methods is reported in Table 2.

2.4. Performance Measures

Two metrics are used to assess the improvement in the Musk3p model performance after integrating the quantitative and qualitative observations. First, Nash-Sutcliffe efficiency (NSE)

$$NSE = 1 - \frac{\sum_{t=1}^{T}\left(Q_t^m - Q_t^o\right)^2}{\sum_{t=1}^{T}\left(Q_t^m - \overline{Q}^o\right)^2} \tag{19}$$

and, second, bias ratio (BR)

$$BR = \frac{\sum_{t=1}^{T} Q_t^m}{\sum_{t=1}^{T} Q_t^o} \tag{20}$$

where Q_t^m is the simulated streamflow at time step t; Q_t^o is the observed streamflow value; $\overline{Q}^o$ is the average observed streamflow.

3. Experimental Setup

Four numerical experiments are carried out to explore the benefit of assimilating qualitative observations in hydrological modeling. Different scenarios of observation bias, core (D and E), and support values of the fuzzy sets to generate fuzzy model output and flow observations (α_M and α_O) from qualitative indications are proposed. Model results integrating quantitative flow observations using KF are used as benchmark to compare the integration of qualitative observations. To assess the sensitivity of states updates to model error, two different values of model error γ are considered in all experiments: low model error of $\gamma = 0.01$ (γ_{low}) and high model error of $\gamma = 0.1$ (γ_{high}).

In the first experiment, the comparison between the methods used to integrate qualitative observations into the Musk3p model is carried out. In this case, fixed values of α_M, α_O, and β are considered (Table 3). The model fuzzy sets are narrower, that is, more accurate, than the observation fuzzy sets ($\alpha_{M} < \alpha_O$). Both deterministic (KF, FEC) and ensemble (IMM, EnsFEC) methods are used. In this experiment, an ensemble of flow observations with 100 members is generated.

The second experiment focuses on understanding the influence of the core value of each fuzzy set on the assimilation of qualitative observations. For this reason, the 0.01 and 0.99 quantile values are perturbed

Table 3
Summary of the Experiments Performed in This Study Where α_O and α_M Are Coefficients of the Fuzzy Set Support for the Estimation of Qualitative Observations and Modeled Flow Values, γ and β Are the Model and Observation Errors Used in KF and IMM, λ Is a Perturbation Coefficient of the Fuzzy Sets Core, and E and D Are the Core Values of the Dry and Low Fuzzy Sets, While $q_{0.01}$ and $q_{0.99}$ Are the 0.01 and 0.99 Quantiles

	Experiment	Fuzzy sets	α_O	α_M	γ	β	Biased
1	Initial comparison	$D = q_{0.01}$ $E = q_{0.99}$	0.75	0.55	0.01, 0.1	0.4	No
2	Effect of fuzzy sets core choice	$D = q_{0.01} + U(0,\lambda)$ $E = q_{0.99} - U(0,\lambda)$	0.75	0.55	0.01, 0.1	0.4	No
3	Effect of biased observations	$D = q_{0.01}$ $E = q_{0.99}$	0.75	0.55	0.01, 0.1	0.4	Yes
4	Effect of fuzzy sets support choice	$D = q_{0.01}$ $E = q_{0.99}$	[0.5, 1.5]	0.5, 1.0, 1.5	0.01, 0.1	0.2, 0.4, 0.6	No

adding a uniform noise that varies from zero to λ, a value proportional to the distance between E and D as (E-D)/8. In particular, in case of D this noise is added to the 0.01 quantile, while for E it is subtracted from the 0.99 quantile (see Table 3). Because of the approach use to estimate the other core value sets, a change in E and D generates a resulting proportional variation in L, M, and H.

In the third experiment, the influence of biased qualitative information on the updating performances are studied. Four different scenarios of biases are considered: (1) no bias, (2) negative bias, (3) positive bias, and (4) random bias at each time step during the considered flood event. Regardless the type of bias, the observation classes is biased only by one positive or negative fuzzy class. For example, in case of negative bias, if the qualitative observation at time t belongs to high flow the corresponding biased flow value is medium flow and not low or dry flow classes.

In the fourth and the last experiments, the effect of the fuzzy sets' support coefficient α_O on the estimation of fuzzy observations is studied. The value of α equal to 0.5 indicates no overlap between sets, while the higher values represent higher overlap and hence a higher uncertainty. This experiment is carried out only for ensemble methods (IMM and EnsFEC). In fact, varying the membership functions' support in deterministic approaches (KF and FEC) would not induce any change in the assimilation performance since the observed flow value is chosen equal to the core of the fuzzy set, so independent of the set support value. Different values of observational error (β, model state updating) and fuzzy support coefficient of mode output (α_M, output correction).

4. Material

4.1. Case Study and Data Set

Five rivers in the United States (Figure 6) are included in this study: the McKenzie, Missouri, Allegheny, Sabine, and Trinity Rivers. These rivers cover a wide range of hydrometeorological conditions, and for each of these rivers data from an upstream and a downstream gauging station were available.

The McKenzie river extends for about 150 km in the state of Oregon (north-west United States), draining an area of about 3400 km^2. The Missouri River flows for 3,768 km from Montana to Missouri (where it joins the Mississippi River), and it is the longest in the contiguous United States. The drainage area is 1,371,000 km^2, and it is largely semiarid. The Allegheny flows through the Allegheny Plateau of Pennsylvania and New York (north-west USA) for about 507 km. The Sabine River is a transboundary river between the states of Texas and Louisiana and has a drainage area of about 25,270 km^2, while the Trinity River flows for 1,140 km across the state of Texas, draining an area of 46,100 km^2. For all the river reaches hourly discharge data (i.e., the same time step of the model time step) are available from monitoring stations managed by the U.S. Geological Survey (USGS, 2016). Table 4 summarizes the considered upstream and downstream stations for each river reach, the length of the reaches, and the period of the flow time series (more details are provided in the next section).

4.2. Model Setup

Model calibration is performed using the least squares minimization technique by means of the Broyden-Fletcher-Goldfarb-Shanno variant of the Davidon-Fletcher-Powell minimization algorithm (Press et al.,

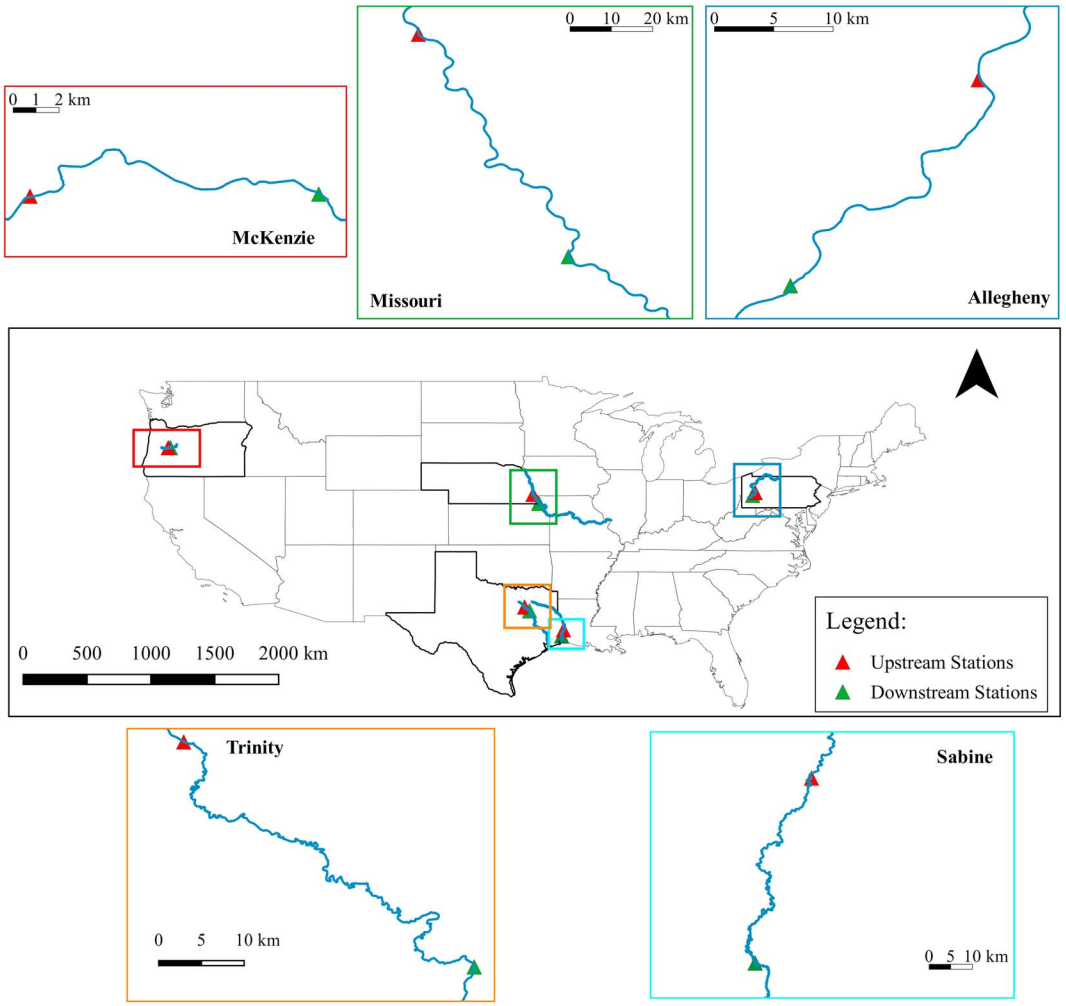

Figure 6. Location of the five river reaches and corresponding upstream and downstream flow stations.

1992). The use of gradient-based technique in this case is possible due to the analytical representation of model equations and differentiability of the error function. The optimal model parameters are reported in Table 5. The flow data period is divided in two sets for model calibration and validation. Because the Musk3p model for the Trinity and Sabine Rivers was already calibrated by Lee et al. (2011), the same optimal parameter sets are used in this study. On the other hand, the model for the remaining three river reaches is calibrated using the aforementioned approach for the periods from 1 January 1991 to 7 March 2001 (McKenzie River), 1 January 1991 to 18 February 1996 (McKenzie River), and 15 December 1991 to 6 August 1999 (Allegheny River). The resulting calibrated hydrographs for the three river reaches are

Table 4
Characteristics of the Selected River Reaches in Terms of River Length, Upstream and Downstream Station and Duration of the Flow Data Period Used for Model Calibration and Validation

River	Upstream station	Downstream station	Reach length	Flow data period
McKenzie	Leaburg dam	Vida	12 km	October 1989 to November 2017
Missouri	Nebraska city	Rulo	80 km	December 1991 to November 2017
Allegheny	Kittanning	Natrona	25 km	December 1991 to November 2017
Trinity	Rosser	Trinidad	78 km	January 2007 to January 2015
Sabine	Bon Wier	Deweyville	88 km	January 2007 to January 2015

Table 5
The Optimal Sets of Parameter Values for the Musk3p Model Implemented Along the Five River Reaches

Parameters	Trinity	Sabine	McKenzie	Missouri	Allegheny
K (hr)	47.28	78.97	2.59	10.75	15.62
X (−)	0.47	0.35	0.024	0.021	0.011
ε (−)	0.109	0.105	0.233	0.132	0.190
NSE			0.67	0.90	0.95

Note. The last row indicates the NSE values achieved in calibration. (No NSE value is reported for the Trinity and Sabine rivers as they have been calibrated in previous studies.) NSE = Nash-Sutcliffe efficiency.

reported in the supporting information. The NSE values obtained by using the calibrated sets of parameters for the three river reaches are reported in Table 5. During the validation period, 100 major flood events (20 events per reach) occurred during the flow data period are subjectively selected to assess the proposed methodology.

4.3. Synthetic Qualitative Observations

To explore the benefits of the proposed approaches under controlled experimental conditions, different assumption of fuzzy classes features, and due to the lack of real data, synthetic qualitative flow observations are used. As mentioned in section 1, one potential application of our approach could be the integration of qualitative flow observations provided by people using social media in hydrologic routing models. For this reason, the observed flow value used as input in our experiments is considered to belong to only flow class, for example, low flow.

In order to generate synthetic observations of flow classes we used observed quantitative flow observations. A number of fuzzy sets are created (using $\alpha_O = 0.5$ in order to avoid any overlapping between sets) to identify the flow classes for each river reach. Then, the observed flows are combined with the fuzzy set information to identify in which flow class each flow observation belongs and provide an indication of the qualitative flow observation class.

Figure 7 show the comparison between quantitative and qualitative observations for five randomly selected flood events occurred in the river reaches of this study. The ensemble of qualitative observations is generated by sampling from a uniform distribution with the ranges coinciding with the support of a fuzzy set.

5. Results and Discussions

5.1. Experiment 1: Initial Comparison

The aim of this experiment is to compare states and output correction methods in case of base assumption on model (γ) and observation errors (β). Fixed values of α_O and β are used (see Table 3), and *Nens* is set equal to 100 to properly represent all the streamflow values belonging to a specific fuzzy set. Experiment 1 is then

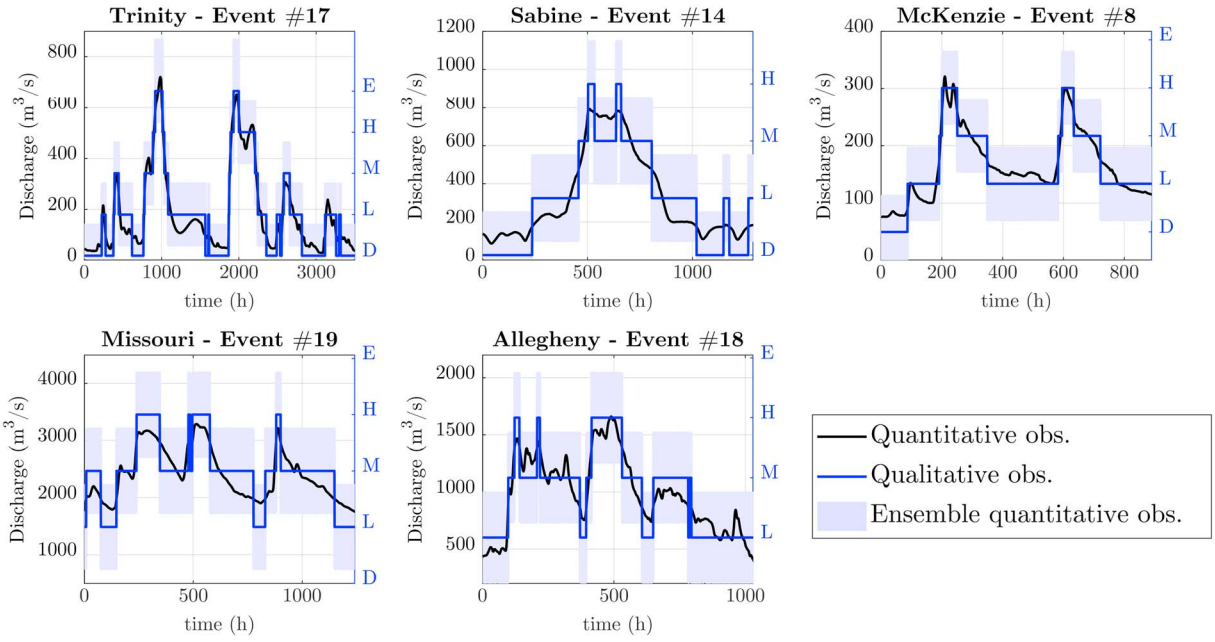

Figure 7. Comparison between quantitative and qualitative flow observations calculated with $\alpha_O = 0.5$ for the five selected flood events occurred in the river reaches. The ensemble of qualitative flow observations is represented by the light blue shaded area, while D, L, M, H and E represents the fuzzy cores for the dry, low, medium, high, and extreme flow classes.

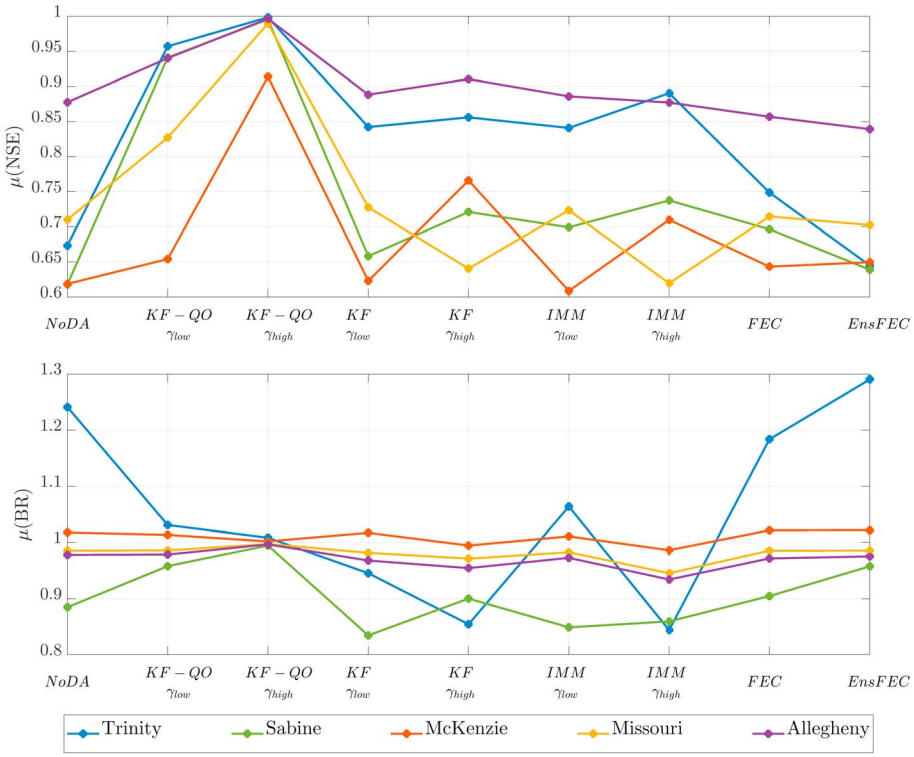

Figure 8. Average values of NSE (μ (NSE), first row) and BR (μ (BR), second row) calculated considering the 20 flood events occurred in each of the 5 case studies (different color lines) for all updating methods (with low and high values of model error γ) and model without update (NoDA) KF = Kalman filter; IMM = Interacting Multiple Model; FEC = Fuzzy Error Correction; NSE = Nash-Sutcliffe efficiency; BR = bias ratio.

performed using all the 100 flood events occurred in the five case studies. KF assimilating quantitative flow observations is used as the benchmark to compare the other techniques. The outcomes of Experiment 1 are summarized in Figure 8, where the mean of the NSE and BR indices of each updating method (estimated for each flood event) are compared.

Overall, integration of qualitative observations improves model performances if compared to the model results without update. Output correction approaches tend to provide lower results than model states updating. It can be observed that the results obtained with KF and IMM with γ_{low} in most of the river reaches are comparable to the ones achieved without model update (NoDA) in both NSE and BR. On the other hand, assimilation of qualitative observations using KF and IMM for γ_{high} values gives satisfactory improvement. However, opposite outcomes are achieved for the Missouri River, where μ (NSE) values are higher for KF and IMM with γ_{low} than with γ_{high}. This can be related to the combined effect of qualitative observations that tend to misrepresent streamflow observation and to the γ_{high} value that is driving the model closer to the qualitative observations. In fact, for γ_{low}, less trust is given to observations, and the model results are actually better than for the γ_{high} values. This brings us to a conclusion that assimilation of uncertain qualitative observations is not always beneficial to improve the model output if high model errors are assumed.

Overall, underestimation of the river flow is observed on all the river reaches, with the exemption of the McKenzie and Trinity Rivers. In particular, on the Trinity River such overestimation is significant for NoDA, FEC, and EnsFec methods. This can be due to the slow response of the output correction methods to improve overestimated simulated flow integrating qualitative observations in steep rising hydrograph. A possible solution could be to change the fuzzy set rules of the FEC in order to provide a *more negative* value of AOC when simulated flows are in high flood classes (e.g., H and E) and observed flows are in low classes.

One can see that on average, the output correction methods typically slightly improve NSE values if compared with the ones obtained without update. For the Missouri and Allegheny Rivers, output correction shows similar and lower NSE values than the ones without model update. However, the FEC and

EnsFEC methods give higher μ (NSE) and lower σ (NSE) values than KF with γ_{low}. On average, FEC and EnsFEC show comparable BR values.

Last but not least, performances of ensemble methods like IMM and EnsFEC are slightly higher than those obtained using their corresponding deterministic approaches (KF and FEC). In particular, IMM gives similar μ (NSE) and σ (NSE) values to KF in case of γ_{low} (less assimilation effect), while results of IMM are better than KF with γ_{high}. On the other hand, outcomes of EnsFEC are comparable to the ones achieved with FEC for the Missouri and McKenzie Rivers, while they are worse for the other three river reaches.

Integration of qualitative flow observations are highly dependent on the characteristic of the fuzzy sets used to calculate qualitative observations. The following experiments are specifically designed to address this issue.

5.2. Experiment 2: Effect of Fuzzy Sets Core Choice

In Experiment 2, several scenarios of core values for dry and extreme flow sets (D and E) are introduced for representing the uncertainty in the proper characterization of fuzzy sets and the effect on the assimilation performances. In this way, different values of L, M, and H are also generated. The outcomes of Experiment 2 are summarized in Figure 9. Only KF and FEC methods are considered at this stage because of the similar model performance achieved using the ensemble methods. As for Experiment 1, a better μ (NSE) is found for the model error γ_{high} in all the three river reaches displayed in Figure 9. KF with γ_{low} and FEC show similar performances in the Sabine and Allegheny Rivers. In addition to these outcomes, Experiment 2 provides the other two important outcomes.

The first one is that model results obtained with KF for γ_{low} are not influenced by the different characterizations of fuzzy sets, and, in general, by the assimilation of fuzzy observations. Similar weak dependency can be also seen for FEC. In fact, also for μ (NSE) there is no a clear correlation between μ (NSE) and the variation of E and D. Hence, the FEC method can be considered a useful approach when it is difficult to define a priori the characteristics of the fuzzy sets.

On the other hand, the results of KF with γ_{high} show higher μ (NSE) variability than the other two approaches. In particular, model output is more sensitive to the value of the dry fuzzy set rather than the extreme flow set in the Trinity and Sabine Rivers, while a mixed effect can be observed in the Allegheny River. Figure 9 shows the results achieved only on three river reaches as small variations of NSE are obtained with the McKenzie River (which does not allow to draw any general conclusions about influence of D and E values), while similar results are obtained on the Missouri River. The results for the McKenzie and Missouri Rivers are included as the supporting information.

5.3. Experiment 3: Effect of Biased Observations

In order to evaluate the effect of biased qualitative (fuzzy) observations on the model output, four diverse scenarios are considered: (1) no bias, (2) negative bias, (3) positive bias, and (4) random bias at each time step. All updating methods are considered. In case of random bias, 100 simulations are run to randomly generate negative and positive biases for each simulation time step. The outcomes are summarized in Figures 10 and 11. As in the case of Experiment 1, μ (NSE) represents the average values of the single NSE obtained in the 20 flood events occurred on the river reaches. As expected, integrating qualitative observations without bias outperforms all the remaining three bias scenarios both in terms of NSE and BR.

Positive bias leads to the lowest μ (NSE) in the Trinity and Sabine Rivers, while negative bias leads to worse performance for McKenzie, Missouri, and Allegheny. Random observation bias tends to provide higher μ (NSE) value, while μ (BR) values are found always lower than 1 and similar to those obtained integrating qualitative observations with negative bias. It is worth noting that assimilation of positively biased observations shows a μ (BR) value higher than 1, as opposed to the negatively biased observations. As expected, low variations of μ (NSE) and μ (BR) are obtained for γ_{low} as the updated model tends to *mistrust* the qualitative observations and no update is performed.

An interesting result is that output correction methods are only marginally affected by different bias scenarios. In fact, small variations of μ (NSE) and μ (BR) are achieved with FEC and EnsFEC (in particular for the McKenzie, Missouri, and Allegheny Rivers). On the other hand, KF implemented for γ_{high} shows to be very sensitive to biased observations.

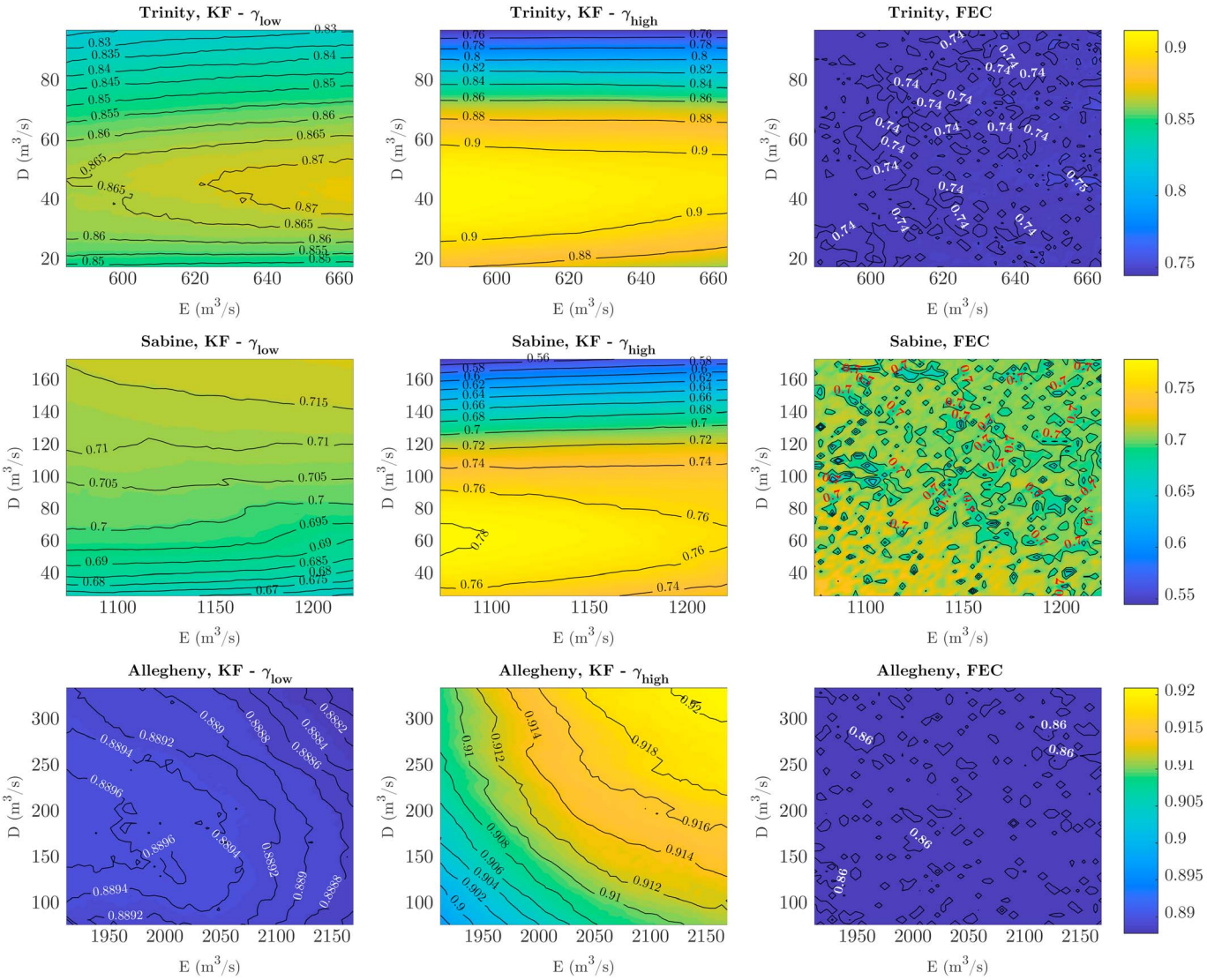

Figure 9. Contour graphs representing the μ (NSE) values obtained from Experiment 2 using different D and E values on the Trinity, Sabine, and Allegheny Rivers for KF for model error γ having low value (first column), high value (second column), and for the FEC method (third column). Nash-Sutcliffe efficiency; KF = Kalman filter; FEC = Fuzzy Error Correction

As mentioned, the previous outcomes are obtained considering the support of the fuzzy observations sets (α_O) equal to 0.75, without accounting for the uncertain nature of the fuzzy sets definition.

5.4. Experiment 4: Effect of Fuzzy Sets Support Choice

Experiment 4 aims at understanding the effect of different supports of fuzzy sets (α_O) for the qualitative flow observations. In addition, different values of model error γ, observational error β, and model fuzzy sets' support α_M are considered as well. Only the ensemble methods IMM and EnsFEC are analyzed as KF and FEC are not affected by different α_O. The range of variation of the coefficient α_O is from 0.5 to 1.5 in order to cover scenarios with no overlap between fuzzy sets ($\alpha_O = 0.5$) and scenarios with extreme overlap between sets ($\alpha_O = 1.5$). The outcomes of Experiment 4 are summarized in Figure 12. The main finding is that EnsFEC provides similar model outputs for varying values of α_O, as opposed to IMM. In addition, EnsFEC is not very sensitive to big variation of model fuzzy sets' support α_M.

Overall, IMM with γ_{high} gives higher NSE values than IMM with γ_{low}. In addition, more stable and higher μ (NSE) values are achieved by IMM with respect to FEC for σ_O lower than 1. In fact, as in Experiment 1, no

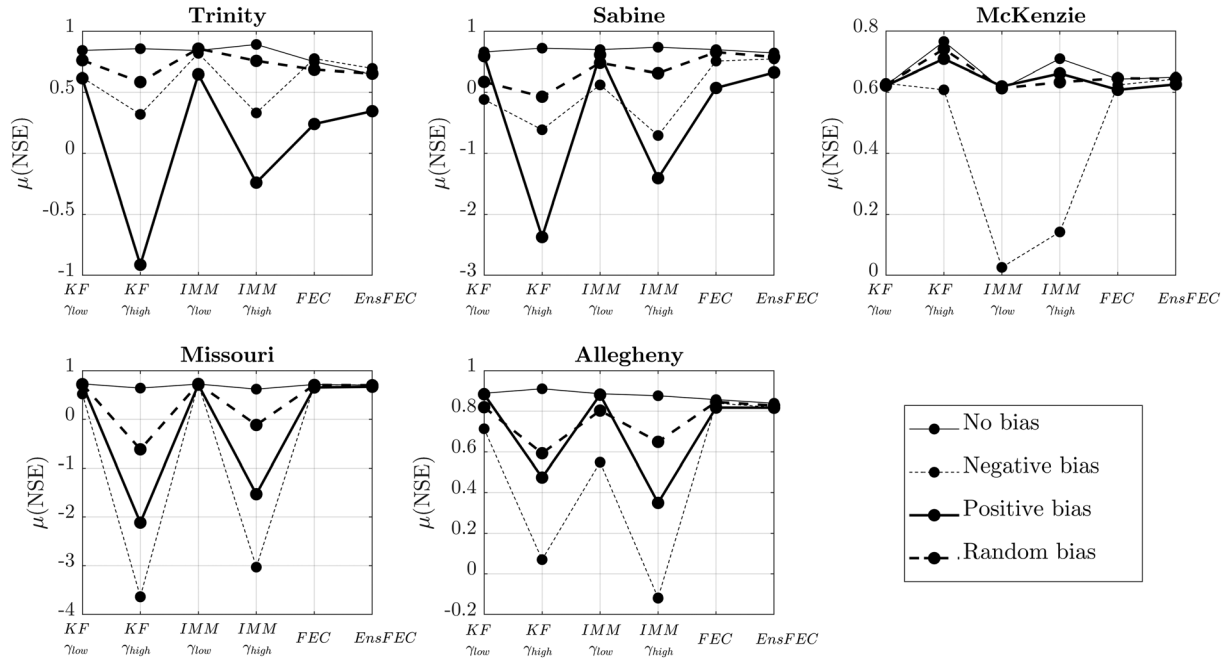

Figure 10. μ (NSE) values obtained in case of different assumptions for the bias of the qualitative fuzzy observations on the five river reaches for all types of updating techniques with low and high values of model error γ.

assimilation effect is observed for low model error so that also small influence of α_O can be observed in the model results. This means that for low value of observational error, the state updating methods outperform output correction techniques. However, IMM shows that a significant reduction in the μ (NSE) values is obtained for α_O values higher than 1. Similar trend in μ (NSE) are achieved in the Trinity, Sabine, and

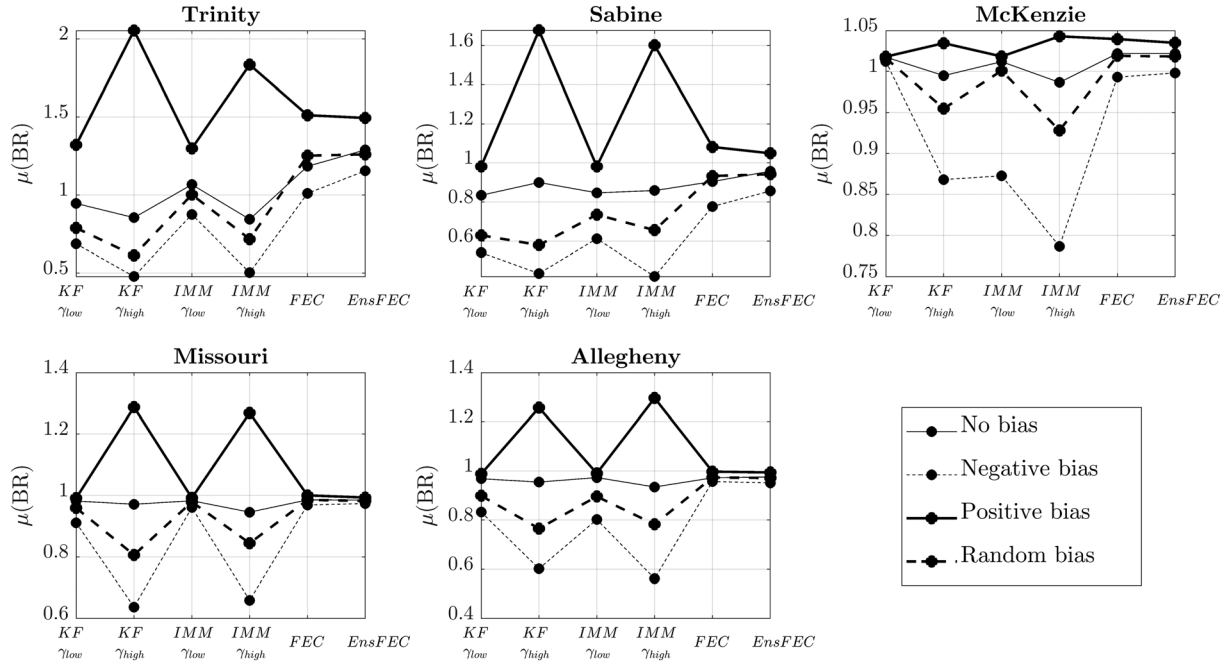

Figure 11. μ (Bias) values obtained in case of different assumptions for the bias of the qualitative fuzzy observations on the five river reaches for all types of updating techniques with low and high values of model error γ.

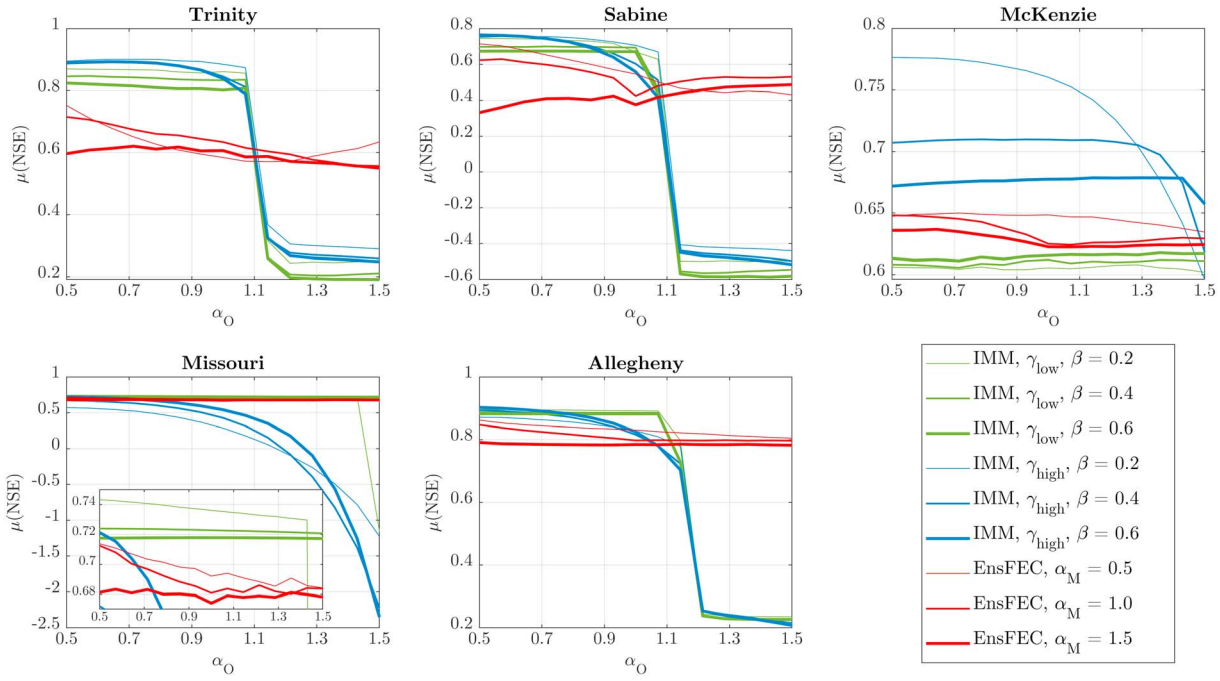

Figure 12. μ (NSE) values obtained for different values of the coefficient α_O on the five river reaches in case of low and high values of model error γ, various assumption on observation error β for IMM method, and three values of coefficient α_M for the FEC approach. IMM = Interacting Multiple Model; FEC = Fuzzy Error Correction; NSE = Nash-Sutcliffe efficiency.

Allegheny Rivers. On the McKenzie River the μ (NSE) achieved with IMM is stable and higher than FEC only for γ_{high}. A higher variability of model's results is achieved with small values of β, that is, small observational error and more influence on the state update technique. An interesting behavior can be observed in the Missouri River, where a drastic reduction of μ (NSE) occurs with IMM for high model error. This is not due to the low model performances without update but to the misrepresentation of the observed flow value by the synthetic qualitative observations; as demonstrated in Experiment 1 and Figure 8. In this way, the γ_{high} value drives the model closer to the uncertain qualitative observations.

As for the previous experiments, also Experiment 4 shows the importance of an accurate estimation of the model error γ in state update approaches. For this purpose, data-driven models estimating model errors (e.g., Abebe & Price, 2003; Pathiraja et al., 2018) or the model error distribution (Solomatine & Shrestha, 2009) may be useful to apply. In addition, for uncertain estimation of the set support for the fuzzy observations, output update methods tend to provide more stable results than the state updates techniques.

Experiment 4 strengthens the outcomes of Experiments 2 and 3 and demonstrates that FEC method can be a powerful tool when the characteristics of the flow fuzzy sets (core and support values) are not a priori known. However, the low sensitivity of the FEC method to different cores and supports fuzzy sets' values does not results in sizeable improvement of the NSE, if compared to data integration is applied as previously described in Experiment 1 and Figure 8.

6. Conclusions

This study proposes two innovative approaches to integrate qualitative flow observations, expressed as fuzzy numbers, within a hydrologic routing model. A lumped three-parameter Muskingum model is used to propagate river flow. Deterministic and ensemble synthetic qualitative observations are generated based on observed flow measurements. Two model update methods are implemented: (a) state updating (Kalman Filter, KF, and IMM) and (b) output correction method (Fuzzy Error Corrector, FEC). The IMM and FEC methods implemented in this study are based on techniques employed for the problems of missile tracking and robotics. The methodology is applied to 100 flood events occurred in five river reaches in the United

States. Different numerical experiments are performed assuming various cores and supports of the fuzzy sets for demonstrating the pros and cons of assimilating qualitative observations into a hydrologic routing model.

This study demonstrates that qualitative flow observations can provide valuable information for improving streamflow estimation if integrated within a hydrological model using states or output updating techniques. We prove that state updating methods are strongly affected by the estimation of the model error, which is in accordance with the earlier studies on the topic. The model performance after applying KF, in case of low model error, is comparable to the one without model update, so the assimilation effects are negligible in this case. When the model error is assumed to be high, KF-based approaches, unsurprisingly, provide the best model improvement. However, if streamflow observations are misrepresented by observers, and their estimates are inaccurate, model updating method can bring model closer to uncertain observation and lead to lower performances than the ones without model update. The comparison between deterministic and ensemble methods shows a slightly higher performance by the latter.

Different experiments are carried out to investigate the sensitivity of the updating approaches to the characteristics (cores and supports) of the flow fuzzy sets used to estimate the fuzzy flow observations. An important result is that the output correction techniques provide robust results when the characteristics of the flow fuzzy sets are not a priori known and in case of biased fuzzy observations. In fact, output correction methods show steady performance for any core and support value of the fuzzy sets, although NSE values are lower than the ones obtained with state updating methods. For high model errors, the experiments show that the sensitivity to the fuzzy set core values is higher for the state updating methods than that for the output correction methods. In particular, model performances depend more on the assumed core value for the *dry* fuzzy set, rather than of the *extreme flow* set. Moreover, high NSE values are obtained with IMM for low uncertainty of the fuzzy sets, while a fast reduction of model performance occurs when such spread increases.

One of main limitations of the output correction approach implemented in this study is the lack of any spatial error covariance structure for a distributed structure of the Musk3p model. The recommendation for future study is to include such spatial covariance error using a time invariant gain function to distribute model errors (based on the error at the measurement point) across the entire state of the river system, in a fashion similar to the one presented by Madsen and Skotner (2005). Moreover, in case of additional river reaches, more analysis would be needed to attribute the different performances and sensitivities of the proposed model updating depending on the locations and characteristics of these river reaches (e.g., presence of hydraulic structures, human footprint, climatic region, etc.). Finally, future research activities should focus on the integration of qualitative flow observations within a distributed structure of the Musk3p model, on employment of more complex hydrological models, and on the integration of qualitative flow observations to complement an existing network of quantitative flow data when the latter is not fully available, either temporally or spatially.

Overall, this study demonstrates, on a number of examples, the usefulness of using the qualitative flow observations (represented as fuzzy numbers) in model-based streamflow estimation, and presents the novel model updating techniques that could be used to integrate qualitative flow observations posted by people on social networks.

Acknowledgments
This research was partly funded by the European H2020 Project GroundTruth 2.0, grant agreement 689744, which also supported the PhD study of the first author at IHE Delft Institute for Water Education. Some research ideas and components were developed in the framework of the grant 17-77-30006 of the Russian Science Foundation. Historical daily streamflow data used were supplied by the United States Geological Survey (USGS) National Water Information System (https://waterdata.usgs.gov/nwis/uv/?referred_module=sw). Postprocessed input data and experiments results can be found in (https://github.com/alessandroamaranto/AQDHM). The authors would like to thank Jan Seibert for his constructive suggestion on our methodology and his comments on the earlier version of the manuscript. We also would like to thank Lucia Ciancia for preparing Figure 1 of this manuscript.

References

Abebe, A. J., & Price, R. K. (2003). Managing uncertainty in hydrological models using complementary models. *Hydrological Sciences Journal*, *48*(5), 679–692. https://doi.org/10.1623/hysj.48.5.679.51450

Aronica, G., Hankin, B., & Beven, K. (1998). Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources*, *22*(4), 349–365. https://doi.org/10.1016/S0309-1708(98)00017-7

Assumpção, T. H., Popescu, I., Jonoski, A., & Solomatine, D. P. (2018). Citizen observations contributing to flood modelling: Opportunities and challenges. *Hydrology and Earth System Sciences*, *22*, 1473–1489. https://doi.org/10.5194/hess-22-1473-2018

Bar-Shalom, Y., Li, X. R., & Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation*. New York: John Wiley & Sons, Wiley Interscience. https://doi.org/10.1002/0471221279. ISBNs: 0-471-41655-X (Hardback) 0-471-22127-9 (Electronic)

Brouwer, T., Eilander, D., van Loenen, A., Booij, M. J., Wijnberg, K. M., Verkade, J. S., & Wagemaker, J. (2017). Probabilistic flood extent estimates from social media flood observations. *Natural Hazards and Earth System Sciences*, *17*, 735–747. https://doi.org/10.5194/nhess-17-735-2017

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

Cooperman, R.L. (2002). Tactical ballistic missile tracking using the interacting multiple model algorithm, Information Fusion, Proceedings of the Fifth International Conference, 2, pp.824-831
de Vos, L., Leijnse, H., Overeem, A., & Uijlenhoet, R. (2017). The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, *21*, 765–777. https://doi.org/10.5194/hess-21-765-2017
Eilander, D., Trambauer, P., Wagemaker, J., & van Loenen, A. (2016). Harvesting social media for generation of near real-time flood maps. *Procedia Engineering*, *154*, 176–183. https://doi.org/10.1016/j.proeng.2016.07.441
Etter, S., Strobl, B., Seibert, J., & van Meerveld, H. J. I. (2018). Value of uncertain streamflow observations for hydrological modelling. *Hydrology and Earth System Sciences*, *22*, 5243–5257. https://doi.org/10.5194/hess-22-5243-2018
Fohringer, J., Dransch, D., Kreibich, H., & Schröter, K. (2015). Social media as an information source for rapid flood inundation mapping. *Natural Hazards and Earth System Sciences*, *15*(12), 2725–2738. https://doi.org/10.5194/nhess-15-2725-2015
Georgakakos, A. P., Georgakakos, K. P., & Baltas, E. A. (1990). A state-space model for hydrologic river routing. *Water Resources Research*, *26*(5), 827–838. https://doi.org/10.1029/WR026i005p00827
Giuliani, M., Castelletti, A., Fedorov, R., & Fraternali, P. (2016). Using crowdsourced web content for informing water systems operations in snow-dominated catchments. *Hydrology and Earth System Sciences*, *20*, 5049–5062. https://doi.org/10.5194/hess-20-5049-2016
Heemink, A. W., & Segers, A. J. (2002). Modeling and prediction of environmental data in space and time using Kalman filtering. *Stochastic Environmental Research and Risk Assessment*, *16*(3), 225–240. https://doi.org/10.1007/s00477-002-0097-1
Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552
Kovitz, J. L., & Christakos, G. (2004). Assimilation of fuzzy data by the BME method. *Stochastic Environmental Research and Risk Assessment*, *18*(2), 79–90. https://doi.org/10.1007/s00477-003-0128-6
Lahoz, W., Khattatov, B., & Menard, R. (2010). *Data assimilation: Making sense of observations*. Berlin, Heidelberg: Springer-Verlag, Springer Science & Business Media. https://doi.org/10.1007/978-3-540-74703-1
Le Boursicaud, R., Pénard, L., Hauet, A., Thollet, F., & le Coz, J. (2016). Gauging extreme floods on YouTube: Application of LSPIV to home movies for the post-event determination of stream discharges. *Hydrological Processes*, *30*, 90–105. https://doi.org/10.1002/hyp.10532
Le Coz, J., Patalano, A., Collins, D., Guillén, N. F., García, C. M., Smart, G. M., et al. (2016). Crowdsourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand. *Journal of Hydrology*, *541*, 766–777. https://doi.org/10.1016/j.jhydrol.2016.07.036
Lee, H., Liu, Y., He, M., Demargne, J., & Seo, D.-J. (2011). Variational assimilation of streamflowinto three-parameter Muskingum routing model forimproved operational river flow forecasting. EGU General Assembly 2011. *Geophysical Research Abstracts*, *13*, EGU2011–13073.
Li, Z., Wang, C., Emrich, C. T., & Guo, D. (2017). A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, *45*, 97–110. https://doi.org/10.1080/15230406.2016.1271356
Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*, W07401. https://doi.org/10.1029/2006WR005756
Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H. J., Kumar, S., Moradkhani, H., et al. (2012). Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, *16*(10), 3863–3887. https://doi.org/10.5194/hess-16-3863-2012
Madsen, H., & Skotner, C. (2005). Adaptive state updating in real-time river flow forecasting—A combined filtering and error forecasting procedure. *Journal of Hydrology*, *308*(1–4), 302–312. https://doi.org/10.1016/j.jhydrol.2004.10.030
Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, *7*(1), 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2
Maybeck, P. S. (1982). *Stochastic models, estimation, and control*. New York: Academic Press.
Mazzoleni, M., Alfonso, L., Chacon-Hurtado, J., & Solomatine, D. (2015). Assimilating uncertain, dynamic and intermittent streamflow observations in hydrological models. *Advances in Water Resources*, *83*, 323–339. https://doi.org/10.1016/j.advwatres.2015.07.004
Mazzoleni, M., Alfonso, L., & Solomatine, D. P. (2017). Influence of spatial distribution of sensors and observation accuracy on the assimilation of distributed streamflow data in hydrological modelling. *Hydrological Sciences Journal*, *62*, 389–407. https://doi.org/10.1080/02626667.2016.1247211
Mazzoleni, M., Chacon-Hurtado, J., Noh, S. J., Seo, D. J., Alfonso, L., & Solomatine, D. (2018). Data assimilation in hydrologic routing: Impact of model error and sensor placement on flood forecasting. *Journal of Hydrologic Engineering*, *23*(6), 04018018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001656
Mazzoleni, M., Cortes Arevalo, V. J., Wehn, U., Alfonso, L., Norbiato, D., Monego, M., et al. (2018). Exploring the influence of citizen involvement on the assimilation of crowdsourced observations: A modelling study based on the 2013 flood event in the Bacchiglione catchment (Italy). *Hydrology and Earth System Sciences*, *22*, 391–416. https://doi.org/10.5194/hess-22-391-2018
Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., & Solomatine, D. P. (2017). Can assimilation of crowdsourced data in hydrological modelling improve flood prediction? *Hydrology and Earth System Sciences*, *21*, 839–861. https://doi.org/10.5194/hess-21-839-2017
McLaughlin, D. (2002). An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering. *Advances in Water Resources*, *25*(8–12), 1275–1286. https://doi.org/10.1016/S0309-1708(02)00055-6
Michelsen, N., Dirks, H., Schulz, S., Kempe, S., AlSaud, M., & Schüth, C. (2016). YouTube as a crowd-generated water level archive. *Science of the Total Environment*, *568*, 189–195. https://doi.org/10.1016/j.scitotenv.2016.05.211
Moulton, K. M., Cornell, A., & Petriu, E. (2001). A fuzzy error correction control system. *IEEE Transactions on Instrumentation and Measurement*, *50*(5), 1456–1463. https://doi.org/10.1109/19.963224
O'Donnell, T. (1985). A direct three-parameter Muskingum procedure incorporating lateral inflow. *Hydrological Sciences Journal*, *30*(4), 479–496. https://doi.org/10.1080/02626668509491013
Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., & Geenens, G. (2018). Data-driven model uncertainty estimation in hydrologic data assimilation. *Water Resources Research*, *54*, 1252–1280. https://doi.org/10.1002/2018WR022627
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in Fortran*, (2nd ed.).
Refsgaard, J. C. (1997). Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrology*, *28*(2), 65–84. https://doi.org/10.2166/nh.1997.005
Restrepo-Estrada, C., Camargo de Andrade, S., Abe, N., Fava, M. C., Mendiondo, E. M., & de Albuquerque, J. P. (2018). Geo-social media as a proxy for hydrometeorological data for streamflow estimation and to improve flood monitoring. *Computers & Geosciences*, *111*.

Ross, J., Ozbek, M., & Pinder, G. F. (2008). Kalman filter updating of possibilistic hydraulic conductivity. *Journal of Hydrology*, *354*(1 – 4), 149–159. https://doi.org/10.1016/j.jhydrol.2008.03.005

Rosser, J. F., Leibovici, D. G., & Jackson, M. J. (2017). Rapid flood inundation mapping using social media, remote sensing and topographic data. *Natural Hazards*, *87*(1), 103–120. https://doi.org/10.1007/s11069-017-2755-0

Salinas, J. L., Kiss, A., Viglione, A., Viertl, R., & Blöschl, G. (2016). A fuzzy Bayesian approach to flood frequency estimation with imprecise historical information. *Water Resources Research*, *52*, 6730–6750. https://doi.org/10.1002/2016WR019177

Schnebele, E., Cervone, G., Kumar, S., & Waters, N. (2014). Real time estimation of the Calgary floods using limited remote sensing data. *Water*, *6*(2), 381–398. https://doi.org/10.3390/w6020381

Seibert, J., & McDonnell, J. J. (2002). On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research*, *38*(11), 1241. https://doi.org/10.1029/2001WR000978

Smith, L., Liang, Q., James, P., & Lin, W. (2015). Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management*, *10*(3), 370–380. https://doi.org/10.1111/jfr3.12154

Solomatine, D. P., & Shrestha, D. L. (2009). A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research*, *45*, W00B11. https://doi.org/10.1029/2008WR006839

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., & Gibson, C. (2017). Demonstrating the value of community based ("citizen science") observations for catchment modelling and characterisation. *Journal of Hydrology*, *548*, 801–817. https://doi.org/10.1016/j.jhydrol.2017.03.019

Sun, L., Seidou, O., Nistor, I., & Liu, K. (2015). Review of the Kalman type hydrological data assimilation. *Hydrological Sciences Journal*. https://doi.org/10.1080/02626667.2015.1127376

Todini, E. (2007). A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach. *Hydrology and Earth System Sciences*, *11*(5), 1645–1659. https://doi.org/10.5194/hess-11-1645-2007

U.S. Geological Survey, (2016). National Water Information System data available on the world wide web (USGS Water Data for the Nation), accessed [November 13, 2017], at URL [http://waterdata.usgs.gov/nwis/].

van Meerveld, H. J. I., Vis, M. J. P., & Seibert, J. (2017). Information content of stream level class data for hydrological model calibration. *Hydrology and Earth System Sciences*, *21*, 4895–4905. https://doi.org/10.5194/hess-21-4895-2017

Walker, D., Forsythe, N., Parkin, G., & Gowing, J. (2016). Filling the observational void: Scientific value and quantitative validation of hydrometeorological data from a community based monitoring programme. *Journal of Hydrology*, *538*, 713–725. https://doi.org/10.1016/j.jhydrol.2016.04.062

Walker, J. P., & Houser, P. R. (2005). *Hydrologic data assimilation. Advances in Water Science Methodologies* (Vol. 1, pp. 25–48). Londres: Taylor & Francis.

Wang, R. Q., Mao, H., Wang, Y., Rae, C., & Shaw, W. (2018). Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. *Computers & Geosciences*, *111*, 139–147. https://doi.org/10.1016/j.cageo.2017.11.008

Weeser, B., Stenfert Kroese, J., Jacobs, S. R., Njue, N., Kemboi, Z., Ran, A., et al. (2018). Citizen science pioneers in Kenya—A crowdsourced approach for hydrological monitoring. *Science of the Total Environment*, *631–632*, 1590–1599.

Willner, S. N., Levermann, A., Zhao, F., & Frieler, K. (2018). Adaptation required to preserve future high-end river flood risk at present levels. *Science Advances*, *4*, eaao1914.

WMO (World Meteorological Organization) (1992). Simulated real-time intercomparison of hydrological models. Geneva, Switzerland: World Meteorological Organization, WMO Operational Hydrology Report.

Yang, P., & Ng, T. L. (2017). Gauging through the crowd: A crowd-sourcing approach to urban rainfall measurement and storm water modeling implications. *Water Resources Research*, *53*, 9462–9478. https://doi.org/10.1002/2017WR020682

Yu, D., Yin, J., & Liu, M. (2016). Validating city-scale surface water flood modelling using crowd-sourced data. *Environmental Research Letters*, *11*, 124011. https://doi.org/10.1088/1748-9326/11/12/124011