Data assimilation in hydrologic routing

Impact of model error and sensor placement on flood forecasting

Mazzoleni, Maurizio; Chacon-Hurtado, Juan; Noh, Seong Jin; Seo, Dong Jun; Alfonso, Leonardo; Solomatine, Dimitri

DOI
10.1061/(ASCE)HE.1943-5584.0001656

Publication date
2018

Document Version
Final published version

Published in
Journal of Hydrologic Engineering

Citation (APA)
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), Article 04018018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001656

Data Assimilation in Hydrologic Routing: Impact of Model Error and Sensor Placement on Flood Forecasting

Maurizio Mazzoleni, Ph.D.[1]; Juan Chacon-Hurtado[2]; Seong Jin Noh, Ph.D., A.M.ASCE[3]; Dong-Jun Seo, Ph.D., A.M.ASCE[4]; Leonardo Alfonso, Ph.D.[5]; and Dimitri Solomatine, Ph.D.[6]

Abstract: Diverse hydrologic and hydraulic models of varying complexities have been proposed in the past few decades to accurately predict the water levels and discharges along rivers. Among them, the hydrologic routing models are widely used because of their simplicity, minimal data, and computational requirements. Due to their simplified assumptions, however, they are subject to various sources of uncertainty. To reduce their predictive uncertainty and improve their operational forecast abilities, data assimilation techniques have been proposed to update the states and/or parameters of the mathematic models by integrating real-time river observations with them. However, the characterization of the model errors and the location of the sensors used for data assimilation have an important effect on the model performance. The main objective of this study was to assess the effect of sensor placement and the errors of both the model and the boundary conditions on the assimilation of flow observations in the distributed hydrologic routing models. A Muskingum-Cunge routing model was applied first to a synthetic river reach with a rectangular cross section and then to a more complex natural river, the Bacchiglione River in Italy, with varying geometry of the river cross sections. The Kalman filter was used to assimilate the flow observations. Synthetic and real-world experiments were carried out. The results showed an improved model performance after the assimilation of the flow observations (e.g., a Nash index higher than 0.9 in the synthetic river and 0.85 in the Bacchiglione River); however, the procedure was sensitive to the model error and the locations of the sensors. In particular, when the model error was larger than the boundary condition error, it was suggested to place the sensors in the lower part of the river reach to maximize the model improvement at the river outlet. On average, the model performance was improved by 14% in terms of the Nash index when the sensor was located in the upstream part of the reaches of the Bacchiglione River instead of in the downstream part. Sensors placed in the upper part of the reaches enabled the improved skills to persist for additional lead time of up to 6 h for the forecasting of the water level at the reach outlet. This study presented a method that allowed identifying the optimal locations of the sensors and thus helped to improve the flood forecasts. **DOI: 10.1061/(ASCE)HE.1943-5584.0001656.**

Author keywords: Data assimilation; Sensor positioning; Hydrologic routing model; Flood forecasting.

Introduction

In the past few decades, the negative impact of floods has drastically increased worldwide (European Environment Agency 2005; Di Baldassarre et al. 2010; Aerts et al. 2014). Flood events in Europe, such as the 2013 Elbe floods and the 2013 United Kingdom floods, were considered national crises and are estimated to have caused approximately 15 and 6.5 billion Euro in damages, respectively (European Environmental Agency 2005). Moreover, due to the combined effects of rapid urbanization, the growth of the population in the proximity of floodplains, and increasing the flood levels due to climate change and sea-level rise, this trend is expected to worsen in the coming years (Hinkel et al. 2014; Jongman et al. 2014).

Nonstructural measures such as flood forecasting in early warning systems are used to provide timely and accurate forecasts in order to reduce the impact of floods on urbanized areas (Todini et al. 2005; McLauglin 2002). However, the hydrological and hydraulic models used in these systems contain uncertainties originating from errors in the observations, input, model parameters, and model structure (Pappenberger et al. 2006; Goetzinger and Bardossy 2008; Alfonso and Tefferi 2015). In order to reduce such uncertainties, data assimilation approaches have been actively used in the water system models in the past few decades (WMO 1992; Refsgard 1997) to optimally update the model states, inputs, or parameters in response to the real-time observations coming into the model (Robinson et al. 1998; Moradkhani et al. 2005; McLaughlin 2002; Liu and Gupta 2007; Reichle 2008; Sakov et al. 2010). A number of authors have used data assimilation approaches based on correcting model errors by means of data-driven methods (Babovic and Fuhrman 2002; Vojinovic et al. 2003; Abebe and Price 2004). A number of studies have been carried out to test various schemes of updating the states of the hydraulic models by means of water-depth measurements from remote sensing

[1]Lecturer, Integrated Water Systems and Governance, IHE Delft Institute for Water Education, Westvest 7, 2611 AX, Delft, Netherlands (corresponding author). E-mail: m.mazzoleni@un-ihe.org

[2]Ph.D. Fellow, Integrated Water Systems and Governance, IHE Delft Institute for Water Education, Westvest 7, 2611 AX, Delft, Netherlands. E-mail: j.chaconhurtado@un-ihe.org

[3]Assistant Professor, Dept. of Civil Engineering, Univ. of Texas at Arlington, Nedderman Hall 417, Arlington, TX 76010. E-mail: seongjin.noh@gmail.com

[4]Professor, Dept. of Civil Engineering, Univ. of Texas at Arlington, Nedderman Hall 417, Arlington, TX 76010. E-mail: djseo@uta.edu

[5]Senior Lecturer, IHE Delft Institute for Water Education, Westvest 7, 2611 AX, Delft, Netherlands. E-mail: l.alfonso@un-ihe.org

[6]Professor, Integrated Water Systems and Governance, IHE Delft Institute for Water Education, Westvest 7, 2611 AX, Delft, Netherlands; Professor, Water Resources Section, Delft Univ. of Technology, Mekelweg 2, 2628 CD, Delft, Netherlands; Water Problems Institute of RAS, Gubkina, 3, Moscow 117971, Russia. E-mail: d.solomatine@un-ihe.org

Note. This manuscript was submitted on June 27, 2017; approved on November 21, 2017; published online on March 28, 2018. Discussion period open until August 28, 2018; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699.

(e.g., Andreadis et al. 2007; Neal et al. 2009; Hostache et al. 2010; Matgen et al. 2010; Biancamaria et al. 2011; Giustarini et al. 2011; Mason et al. 2012; García-Pintado et al. 2013; Andreadis and Schumann 2014; Yan et al. 2015) and in situ sensors (Neal et al. 2007; Kim et al. 2013; Ricci et al. 2011; Romanowicz et al. 2006; Neal et al. 2012; Noh et al. 2013; Li et al. 2015; Wang et al. 2016).

However, only a few studies showed the effect of flow assimilation into the hydrologic routing models. Liu et al. (2008) presented an application of the maximum likelihood ensemble filter (MLEF) for a hydrologic channel routing model based on the variable three-parameter Muskingum model. Errors in the inflow and outflow observations, and uncertainties in the initial conditions and Muskingum parameters, were considered. Similarly, Lee et al. (2008) applied a one-dimensional variational assimilation (VAR) method in order to integrate the real-time streamflow observations into a three-parameter Muskingum model. Rakovec et al. (2012) analyzed the assimilation into the routing states (updating the frequency/network density).

One of the problems in flood forecasting and modeling in general is that water managers often face the problem of a limited sampling budget for locating the sensors. To this end, Alfonso et al. (2010) presented various methodologies to optimally locate the water-level and discharge sensors according to the informational content they can produce, while avoiding information overlaps. In addition, Alfonso and Price (2012) presented an alternative for the sampling location based on the value of the information, in which the consequences of decisions were taken into account. A comprehensive review of sensor location methods was provided by Mishra and Coulibaly (2009) and Chacón-Hurtado et al. (2017); the latter included a generic framework guiding the process of sensor placement on the basis of various criteria. However, the knowledge about the effect of sensor locations on data assimilation is limited. Recently, Mazzoleni et al. (2017b) investigated the effect of the location of the sensors in a semidistributed hydrological model in the case of different errors in the observations. However, the authors did not account for error in the boundary conditions and the model itself. It is worth noting the previous studies by the same authors, cited in this paper, on the assimilation of crowdsourced observations into hydrological and hydraulic models. The only common point between the authors' previous studies and this study is the use of the same case studies and models.

Along with the issue of the correct sensor locations, the proper definition of the error of the model and the boundary conditions has a considerable effect on the data assimilation performance. Input to a hydraulic model usually comes from a hydrological model that is subject to various sources of uncertainty, such as the uncertain precipitation, model structure, and model parameters. However, to the best of the authors' knowledge, none of the previous studies have analyzed how the different model errors and sensor network configurations affect the flood forecasts carried out by a simplified hydraulic model such as a hydrologic routing model.

The objective of this study was to assess the effect of the sensor location and the errors in the boundary conditions and the hydrologic routing model on the improvement of flood forecasting, by assimilating the streamflow observations. Assimilation performances are not only affected by the sensor locations but also are also strictly dependent on the quantification of both the boundary conditions and the model errors. Sets of distributed sensors along the river systems were considered to provide the data to feed the hydrological routing. A single river reach (synthetic river) and a more complex river network (the Bacchiglione River in Italy) were used as case studies. Because the real-time flow observations from the distributed physical sensors were not available, synthetic observations were used instead. A Muskingum-Cunge routing model was applied to both the synthetic and the natural rivers, and the Kalman filter was used to assimilate the synthetic flow observations. The focus of the experiment conducted on the synthetic river was to assess the model improvement when the streamflow observations were assimilated from the hypothetical gauges in the case of different errors in the boundary conditions and in the model. For the Bacchiglione River, only the effect of the different gauge locations on the assimilation performance was assessed.

Methodology

The methods used to assimilate the distributed observations, including the definition of the hydrologic routing model and its state-space representation, are described in this section. Also included are a review of the Kalman filter and the characterization of the observation error, followed by the objective measures to assess the model.

Hydrologic Routing Model: Muskingum-Cunge

Over a number of years, various methods for hydrological and hydraulic routing have been proposed. However, it is always difficult to objectively assess the performance of one method in comparison with another one without a full implementation in a case study. This study used the Muskingum-Cunge routing (MC) model (Cunge 1969; Koussis 1983; Ponce and Chaganti 1994; Ponce and Lugo 2001; Todini 2007) to propagate the river flow along the considered rivers. Complete solutions (namely, fully one-dimensional hydrodynamic models) are usually complex in their structure and expensive in their running time, and they do not necessarily lead to better modeling results in the data assimilation contexts for short lead times (Mazzoleni 2017) but yield results comparable with those of the hydraulic routing models (Viessmann and Lewis 2003). The results of the MC model are also comparable to, for example, the dynamic and kinematic wave models for steep slopes (Barati et al. 2013), and they are able to represent fairly well the dynamics of the flow in canals for high flows over long distances in nonmeandric rivers (Heatherman 2008). Finally, the MC model is widely used by practitioners; therefore, this method was deemed relevant and of interest to a large audience.

The MC model is based on the numerical solution of a kinematic wave model applied to a prismatic river reach between the upstream and downstream sections, as detailed in, for example, Todini (2007). It assumes a linear relationship between a channel's storage and its inflow and outflow discharge. The formulation of the Muskingum-Cunge routing approach can be represented

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

where

$$C_1 = \frac{c\Delta t + 2\Delta x\varepsilon}{2\Delta x(1-\varepsilon) + c\Delta t}$$
$$C_2 = \frac{c\Delta t - 2\Delta x\varepsilon}{2\Delta x(1-\varepsilon) + c\Delta t}$$
$$C_3 = \frac{2\Delta x(1-\varepsilon) - c\Delta t}{2\Delta x(1-\varepsilon) + c\Delta t} \tag{2}$$

t and j = temporal and spatial discretization, respectively; Δx and Δt = spatial and temporal increments, respectively; c = wave celerity; and ε = coefficient

$$\varepsilon = \frac{1}{2}\left(1 - \frac{q}{c\Delta x S_0}\right) \tag{3}$$

where q = unit-width, or specific, discharge; and S_0 = channel bottom slope. The wave celerity can be estimated as proposed by Todini (2007).

In order to apply the data assimilation method for updating the states, it is necessary to express Eq. (1) in a state-space form, that is, in an ordinary differential equation in which the flow at time step $t+1$ along the river is obtained as a function of the flow at time step t. For the state-space representation, this work used the approach proposed by Georgakakos et al. (1990). In this approach, Eq. (1) was converted into a routing states equation that described the changes in the system-states vector $\mathbf{x}$ that responded to the input $\mathbf{I}$

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{I}_t + w_t \quad w_t \sim N(0, \mathbf{M}_m t) \tag{4}$$

where $\mathbf{x}_t = (Q_1^t, Q_2^t, \ldots Q_j^t, \ldots, Q_N^t)$ is the $n_{\text{states}} \times 1$ vector of the model states (streamflow in m^3/s) where n_{states} is the number of discrete reaches into which the river is divided; $\mathbf{I}_t = (Q_L^t, Q_L^{t+1})$ is the 2×1 input vector in which Q_0 is the discharge at the upstream boundary condition; and w_t = uncertainty (due, for example, to an inadequate model structure) represented by the normal distribution with zero mean and error covariance $\mathbf{M_m}$ at time t. The state-transition and input-transition matrices $\mathbf{A}(n_{\text{states}} \times n_{\text{states}})$ and $\mathbf{B}(n_{\text{states}} \times 2)$ are given by Georgakakos et al. (1990)

$$\mathbf{A} = \begin{bmatrix} C_{1,3} & 0 & \ldots & 0 \\ C_{2,1} + C_{2,2}C_{1,3} & C_{2,3} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \prod_{j=3}^{N=1} C_{j,2}(C_{2,1} + C_{2,2}C_{1,3}) & \cdots & C_{N-1,3} & \vdots \\ \prod_{j=2}^{N=1} C_{j,2}(C_{2,1} + C_{2,2}C_{1,3}) & \cdots & \cdots & C_{N,3} \end{bmatrix} \tag{5}$$

$$\mathbf{B} = \begin{bmatrix} C_{1,1} & C_{1,2} \\ C_{2,2}C_{1,1} & C_{2,2}C_{1,2} \\ \vdots & \vdots \\ \prod_{j=2}^{N} C_{j,2}C_{1,1} & \prod_{j=2}^{N} C_{j,2}C_{1,2} \end{bmatrix} \tag{6}$$

The associated observation equation, which relates the observations to the system states

$$z_{t+1} = \mathbf{H}_t^T \mathbf{x}_t + v_t \quad v_t \sim N(0, \mathbf{R_t}) \tag{7}$$

where z_t is an $n_{obs} \times 1$ matrix representing the flow along the river channel at time $t+1$; v = uncertainty of the measurements represented by normal distribution with zero mean and covariance $\mathbf{R}$; and $\mathbf{H}$ = $n_{obs} \times n_{states}$ output matrix. Because the positions of the flow observations change according to the locations of the sensors, the matrix $\mathbf{H}$ changes accordingly. The Manning's equation is used to estimate the water depth (WD) based on the observed river cross sections.

Kalman Filter and the MC Model

Among the data assimilation techniques, the Kalman filter (KF) is one of the most widely known methods to assimilate, in an efficient recursive way, potentially noisy observations into dynamic linear systems (Kalman 1960). Liu et al. (2012) provides a detailed review of the status, progress, challenges and opportunities in advancing DA for operational hydrologic forecasting. Due to the linear nature of the MC model, the Kalman filter (Kalman 1960) is used to assimilate the streamflow observations and to assess the improvement in the model predictions. Under the assumption of a linear stochastic model, Gaussian error, and an unbiased model states estimate, the KF incorporates all uncertain observations, resulting in the best states estimate with minimum variance of the model error. The KF method consists of two steps: the forecast and the update (or analysis). In the forecast step, the forecast-states matrix is prescribed by Eq. (5).

The $n_{states} \times n_{states}$ model error covariance matrix (Georgakakos et al. 1990) is

$$\mathbf{P}_t^- = \mathbf{A}\mathbf{P}_{t-1}^+\mathbf{A}^T + \mathbf{B}\mathbf{M}_b\mathbf{B}^T + \mathbf{M}_m t \tag{8}$$

where $\mathbf{M}_b$ is the 2×2 covariance error matrix of the boundary conditions; superscript—indicates the forecast covariance; and superscript + indicates the update matrix. When an observation becomes available at the time step t, the prior model states and model error covariance matrix are updated using the analysis equations

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + \mathbf{K}_t(z_t^o - \mathbf{H}_t\mathbf{x}_t^-) \tag{9}$$

$$\mathbf{P}_t^+ = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_t^- \tag{10}$$

where the $(n_{states} \times n_{obs})$ Kalman gain is

$$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{H}_t^T(\mathbf{H}_t\mathbf{P}_t^-\mathbf{H}_t^T + \mathbf{R}_t)^{-1} \tag{11}$$

$\mathbf{x}^+$ = updated (or analysis) model states matrix; $\mathbf{R}$ = diagonal matrix $(n_{obs} \times n_{obs})$ representing the observational error assuming independence among the sensors used; and z^0 = new observation. Because the observations vector z^0 in the MC model is expressed in terms of the river flow, the simulated values of WD, assuming a steady and uniform flow, are converted into discharge using the Manning equation for the natural river cross section available.

One of the main limitations in the KF implementation is the subjective determination of model errors, as discussed in Liu et al. (2012), Sun et al. (2016), and Mazzoleni (2017). Puente and Bras (1987) claimed that the proper error quantification of the model is even more important than the selection of the DA methods.

Observational Error

In addition to an accurate identification of the boundary conditions and model errors, the appropriate estimation of the observations error is an important issue to address, as it influences the KF performance. The observational error is assumed to be normally distributed with zero mean and standard deviation (Weerts and El Serafy 2006; Clark et al. 2008; Mazzoleni et al. 2015)

$$\sigma_t^j = \alpha_Q \cdot Q_{t,obs}^j \tag{12}$$

where the observation error's standard deviation is assumed to be the product between the observed flow Q_{obs} and the coefficient α_Q, which includes the measurements and rating curve uncertainties, and is considered equal to 0.1 (Weerts and El Serafy 2006; Clark et al. 2008; Mazzoleni et al. 2015).

Performance Measures

In order to assess the model performance in relation to the observed values of WD, several traditional statistical measures were used.

One of the most widely used measures in hydrology is the Nash-Sutcliffe efficiency (NSE) index (Nash and Sutcliffe 1970) which compares the simulated and observed quantities

$$NSE = 1 - \frac{\sum_{i=1}^{N} (WD_i^m - WD_i^o)^2}{\sum_{i=1}^{N} (WD_i^m - \overline{WD_i}^o)^2} \quad (13)$$

where WD_i^m = simulated water depth in the ith time step; $WD_i{}^\circ$ = observed water depth; $\overline{WD_i}^o$ = average observed water depth; and N = number of pairs of simulated and observed water depths. An NSE of 1 represents a perfect model simulation, whereas an NSE of 0 indicates that the simulated streamflow is only as skillful as the mean of the observed water depth.

The Pearson's correlation coefficient (R) is used to measure the linear correlation between the two variables, the simulated and observed water depth

$$R = \frac{Cov(WD^m, WD^o)}{\sigma(WD^m)\sigma(WD^o)} \quad (14)$$

where $\sigma(WD^m)$ and $\sigma(WD^o)$ = standard deviations of the simulated and observed water depths, respectively. The values of R close to 1 indicate a strong positive correlation between the two variables.

Last, the bias index, or bias, measures the tendency of the simulated water depth to be higher or lower than the observed water depth (comparing their means)

$$Bias = \frac{\sum_{i=1}^{N} WD_i^m}{\sum_{i=1}^{N} WD_i^o} \quad (15)$$

Values greater than 1 indicate the overestimation of the water depth, and values smaller than 1 represent an overall underestimation.

Case Studies and Data Sets

Two case studies are considered in this section. First, a synthetic river with predefined cross sections and hydraulic features is introduced to study the effect of boundary conditions and model errors in the assimilation of distributed flow observations. Second, the Bacchiglione River is used to evaluate the results achieved for the synthetic river and to demonstrate the feasibility of assimilating real-time observations to improve flood forecasting.

Synthetic River

A synthetic river was used to test the assimilation of flow observations by means of the Kalman filter into a Muskingum-Cunge routing model. For this purpose, a rectangular channel was used, with an increasing width $B(m)$ equal to $50 + S_f x$, where x was the distance along the river and S_f the bed slope, a Manning's coefficient n of 0.035, and a total channel length L equal to 50 km.

The synthetic flow observations along the synthetic river were generated using the time series of the recorded streamflow values as the perfect boundary condition for the MC model. To estimate the simulated flow values, the simulated flow time series calculated with a conceptual hydrological lumped model of the linear reservoir (Szilagyi and Szollosi-Nagi 2010) was used as the boundary condition. The observed flow value and meteorological forcing used for the MC model were two flood events that occurred in the River Brue, United Kingdom (Mazzoleni et al. 2015): flood event A (from November 8 to 29, 1994) and flood event B (from October 28 to November 16, 1994). The focus of this study was the MC routing model rather than the hydrological model. For more details about the hydrological model, refer to Mazzoleni et al. (2015).

Bacchiglione River

The second study area was located in the upstream part of the Bacchiglione River basin in northeastern Italy. The Bacchiglione River, from the most downstream point up to Vicenza, has an average width of approximately 40 m, a slope of approximately 0.5%, and a total length of approximately 50 km. In the considered river reach, no backwater effects were present. On the basis of the geometric characteristics of the river channel, the cross sections in the Bacchiglione River could be assumed to have a rectangular shape. The main urban areas were located in the downstream part of the study area, in particular, close to the Vicenza city. In the study area, there were three main tributaries of the Bacchiglione River, upstream of Vicenza. On the eastern side was the Timonchio River and on the western side were the Leogra and the Orolo, shown in Fig. 1 as reaches 1, 2, and 5, respectively.

Data Sets

In order to evaluate the proposed methodology, the three flood events that occurred in May 2013 (event 1), November 2014 (event 2), and February 2016 (event 3) were considered. Event 1 was a high-intensity event that resulted in several traffic disruptions. Both the forecasted and the measured precipitation time series were available for the considered flood events. These time

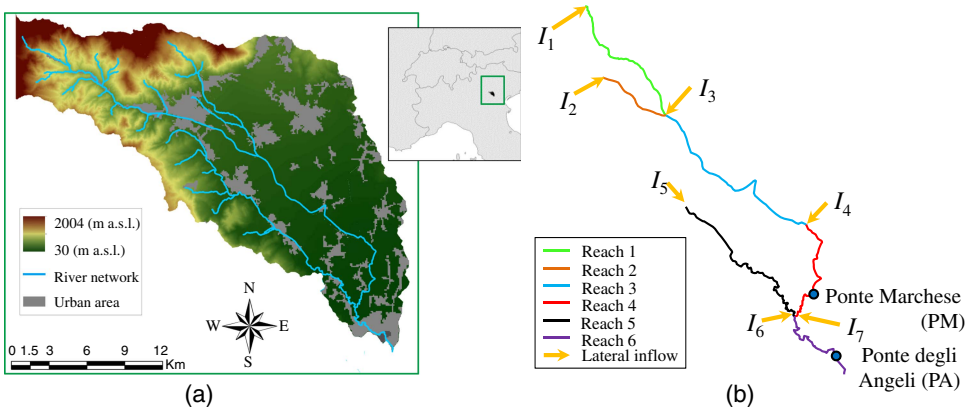

Fig. 1. (a) Location of the Bacchiglione basin; (b) river network and structure of the hydraulic model

J. Hydrol. Eng., 2018, 23(6): 04018018

series of the precipitation were used to prescribe the boundary conditions for the MC model. In the synthetic experiments, the observed precipitation was used as input to the hydrological model to estimate the synthetic observed boundary conditions and to estimate the resulting synthetic observed flow values along the reaches of the Bacchiglione River. The forecasted precipitation was used to estimate the simulated flow.

Model Implementation

In the Bacchiglione Basin, an operational flood early warning system (FEWS) was developed by the Alto Adriatico Water Authority (AAWA) to properly forecast flood events (Ferri et al. 2012). The AAWA uses a forecast horizon of three days for flood forecasting purposes. In this system, modeling is carried out by a cascade of hydrologic and hydraulic models (from *MIKE11* software) used to predict the water levels in Vicenza and the adjacent urbanized areas. *MIKE11* solves the Saint Venant equations in the case of an unsteady flow, on the basis of an implicit finite difference scheme proposed by Abbott and Ionescu (1967). This model assumes a small bottom slope, so that the water depth is less in comparison with the wave length, and a flow parallel to the bottom; that is, the vertical acceleration is ignored and the hydrostatic pressure variation in the vertical direction is assumed (Kamel 2008).

This study used the model chain described in Mazzoleni et al. (2017a), in which a simplified version of the original hydrological model of the AAWA was implemented and where *MIKE11* was replaced by MC. As previously discussed, the MC-based hydrologic routing was chosen not only for its minimum data and computational requirements but also for the wide application of hydrologic routing in real-time forecasting and large-scale modeling (Gochis et al. 2013; Rakovec et al. 2016). However, the MC models do not provide the water level along the river but only the streamflow. On the other hand, because of the nonlinearity of the *MIKE11* model, a nonlinear filter such as the ensemble Kalman filter (as opposed to the linear Kalman filter) needs to be used to assimilate the flow observations. Unfortunately, one of the drawbacks of such a filter is the high computational costs, which may be prohibitive for the assimilation of the water level from multiple sensors in real-time flood forecasting.

Fig. 1 shows the domain of the hydrologic routing model implemented for the Bacchiglione basin. The river was divided into different reaches according to the locations of the internal boundary conditions. The outputs of the conceptual hydrological model were used as the upstream and internal boundary conditions of the MC model, at the locations denoted by the orange arrows. More details about the hydrological model are provided in Mazzoleni et al. (2017b).

Calibration and Validation

As mentioned previously, the MC model required the estimation of the Manning coefficient n in order to calculate the WD along the river reach using the available natural cross sections. The calculated WD was compared with the observed one to evaluate this study's approach. The calibration of the parameter n was performed by minimizing the error between the observed and the simulated rating curves at Vicenza, with Δt and Δx equal to 900 s and 1,000 m, respectively, during the flood event 1. Using this approach, a realistic calibrated value of n equal to 0.08 was found. This roughness value was justified by the physiographic conditions of the channel within the urban area of Vicenza.

With the parameter n a known value, it was possible to evaluate the results by comparing the calibrated MC model with the hydraulic model already implemented by AAWA (*MIKE 11*) for the flood event 1. The hydrographs shown in Fig. 2 indicate a good fit between the flow Q and the water depth WD at two different

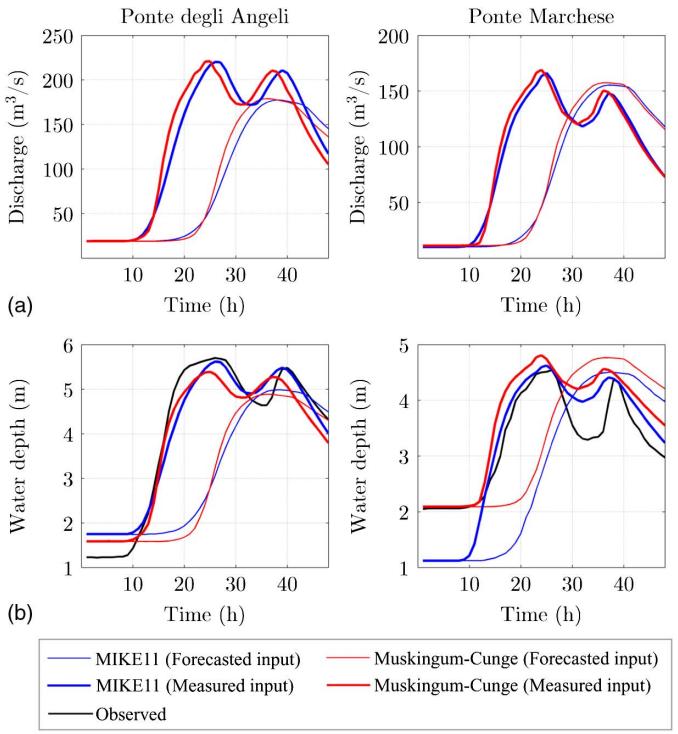

Fig. 2. Comparison between (a) observed and simulated flow and (b) water depth values obtained at Ponte degli Angeli and Ponte Marchese, using *MIKE11* and the MC model in the case of the forecasted and measured input of flood event 1

locations; the values were obtained using *MIKE11* and the MC model, respectively. However, the overprediction of WD is indicated at the Ponte Marchese (PM) in both cases in which measured and forecasted precipitation were used for the hydrological model. This may have resulted from the use of a simplified approach (the Manning equation) to estimate the WD and from the way the parameter n was calibrated. In fact, the optimal value of n was estimated on the basis of the observed rating curve in Vicenza and not at the PM, which may have introduced some bias into the estimation of n. Fig. 2 also shows the observed WD values recorded at the gauging stations of the PM and the Ponte degli Angeli (PA).

Experimental Setup

Two groups of experiments were carried out. In the synthetic experiment, the effects of different sensor placements and diverse assumptions on errors in the boundary conditions and in the model were assessed on the synthetic river and the Bacchiglione River. In the second experiment, DA real-world experiments considering the observed WD records from two existing sensors were carried out on the Bacchiglione River.

Synthetic Experiments

The synthetic experiments were carried out in the synthetic river and the Bacchiglione River. In both cases, the synthetic flow observations were used because data from physical sensors was not available at the time of this study. This made this work an observation system simulation experiment (OSSE). OSSEs are commonly used, for example, in satellite DA and in meteorology to estimate synthetic true states and measurements (Arnold and Dey 1986; Errico et al. 2013; Errico and Privé 2014;

Verlaan and Sumihar 2016). The method used to generate these synthetic observations is described in the following two subsections.

Synthetic River

The flow simulation and subsequent assimilation of the distributed observations were carried out for a hypothetical river over a 50-km reach of rectangular cross sections with varying magnitudes of model error. The spatial discretization of the MC model was set equal to 1,000 m, and dt was $0.9dx$. Given the observed and simulated time series of the upstream boundary conditions, it was possible to estimate the covariance error $\mathbf{M_b}$ for both considered flood events, A and B. The model error covariance matrix $\mathbf{M_m}$ was estimated at each time step as a function of the model states at each cross section

$$\mathbf{M}_{m,t} = (\alpha_m \cdot \mathbf{x}_t)^2 \tag{16}$$

where $\mathbf{x_t}$ = vector of model states at time t without any model update; and α_m = parameter used to define the model error. In the experiment, three different scenarios of model and boundary condition errors were considered. With the boundary condition error assumed fixed, the α_m was varied in order to simulate the conditions described in the following three scenarios:

- Scenario 1: $\mathbf{M_m} \approx \mathbf{M_b}$. The value of α_m was set to 0.35 for both flood events, A and B;
- Scenario 2: $\mathbf{M_m} \gg \mathbf{M_b}$. The MC model was considered to be the main source of error in flood propagation ($\alpha_m = 0.8$);
- Scenario 3: $\mathbf{M_m} \ll \mathbf{M_b}$. The model error was considered negligible with respect to the error in the boundary condition ($\alpha_m = 0.01$).

In these scenarios, data coming from a single location was assimilated.

Bacchiglione River

The focus of this experiment was to understand how the assimilation of the distributed (synthetic) observations from the hypothetical static physical sensors might impact the MC model results at the outlet point at Ponte degli Angeli (PA) in Vicenza during three different flood events. Various hypothetical placements of the sensors along the six river reaches of the Bacchiglione River were considered, in order to study the sensitivity of the model results to the sensor placement.

Various lead times were used in order to evaluate the predictive capability of the MC model in assimilating the streamflow observations at different locations. The updating frequency was equal to the observation interval.

Fixed boundary and model errors were considered. In particular, the errors in the boundary conditions $\mathbf{M_b}$ for reaches 1, 2 and 5 in the headwater catchments were calculated by comparing the observed and the simulated hydrographs derived using the hydrological model. For reaches 3, 4 and 6, $\mathbf{M_b}$ consisted of the error in the flow from the upstream reaches and that from the interbasin estimated with the hydrological model. Fig. 3 shows that $\mathbf{M_b}$ in reach 3 ($\mathbf{M_{b,3}}$) was a function of the error covariance matrix of the flows at the outlets of reaches 1 and 2 ($\mathbf{M_{out,1}}$ and $\mathbf{M_{out,2}}$) and the flow in I_3, $\mathbf{M_{I3}}$. In case of reach 3, it was assumed that the error in the boundary condition, $\mathbf{M_b}$, in reaches 1 and 2 was larger than the model error $\mathbf{M_{m,1}}$ and $\mathbf{M_{m,2}}$. Therefore, it could be assumed that $\mathbf{M_{out,1}}$ and $\mathbf{M_{out,2}}$ had the same magnitude as $\mathbf{M_{b,1}}$ and $\mathbf{M_{b,2}}$. At that point, it was possible to estimate $\mathbf{M_{b,3}}$ as the maximum values among $\mathbf{M_{out,1}}$, $\mathbf{M_{out,2}}$, and $\mathbf{M_{I3}}$. The same procedure was followed for the reaches 4 and 6, as shown in Fig. 3. However, in the case of the available observed streamflow values at the reach outlet, $\mathbf{M_{out}}$ needed to be estimated as a function among the observed and simulated values in the same way it was estimated with reaches 1, 2, and 5.

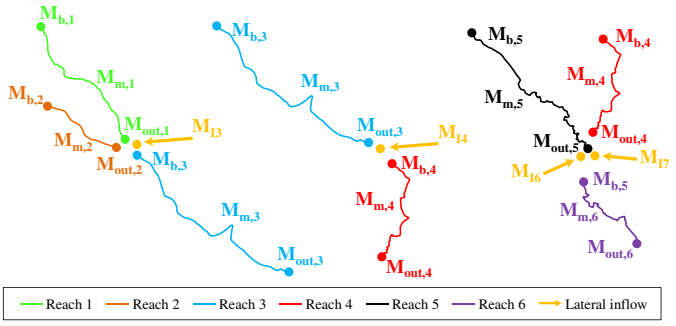

Fig. 3. Representation of the interconnections between model errors and boundary conditions errors for the different river reaches of the Bacchiglione River model

Real-World Experiment

In the real-world experiment carried out on the Bacchiglione River, the flow observations were assimilated from the existing physical locations of the sensors at the PM and PA for the case of event 1. In this experiment, only flood event 1 was considered, due to its higher magnitude in comparison with the other two flood events. This experiment aimed to validate the results of the synthetic experiments in the cases of only two locations of the sensors. Table 1 provides a summary of the synthetic and real-world experiments previously described.

Results and Discussion

Synthetic Experiments

Synthetic River

Fig. 4 shows the simulated hydrographs with and without DA against the synthetic observed hydrograph at two particular river sections. It was assumed that the location of the sensor, or the assimilation point (AP), was 15 km from the upstream boundary of the reach of the synthetic river during the two flood events, A and B. Analysis of the results showed that accurate forecasting was achieved via the MC and the KF. The DA had its main impact downstream from the AP, whereas a small update was observed upstream from the AP. This can be traced to the distributed structure of the MC model and the KF. However, no indication of the effect of the model and boundary conditions error was provided at this step.

Fig. 5 shows the difference between the observed and the simulated WD in time and space (the first row), the Kalman gain $\mathbf{K}$ (the second row), and the model error covariance matrix $\mathbf{P}$ (the third row) during flood event B in scenario 1 for three different APs, at 15, 30, and 45 km from upstream to downstream. These results clearly show the distributed effect of the KF applied on the MC. As expected, the maximum value of the Kalman gain $\mathbf{K}$ was achieved at the AP, because the KF updating effects tended to be propagated both upstream and downstream. However, as previously demonstrated, the difference between the observed and simulated WD was larger upstream from the AP than downstream, because the MC model did not account for the backwater effects. Fig. 5 shows that the symmetric model states covariance matrix $\mathbf{P}$, obtained at the model time step of 100 h, had its smallest values at

Table 1. Summary of the Experimental Setup Proposed in This Study

Setup	Synthetic experiment		Real world experiment
River	Synthetic	Bacchiglione	Bacchiglione
Model	Muskingum-Cunge	Muskingum-Cunge	Muskingum-Cunge
Length	50 km	Approximately 40 km	Approximately 40 km
Cross sections	Rectangular	Natural	Natural
Flood events	2 (event A, event B)	3 (event 1, event 2, event 3)	1 (event 1)
Error scenarios	Boundary error $\approx$ model error Boundary error $\gg$ model error Boundary error $\ll$ model error	Real boundary and model error	Real boundary and model error
Assimilated observations	Water level at three single sensors locations	Water levels at different sensors locations at 6 river reaches	Sensors located at Ponte Marchese (PM) and Ponte degli Angeli (PA)
Lead time	1 h	Up to 8 h	Up to 8 h

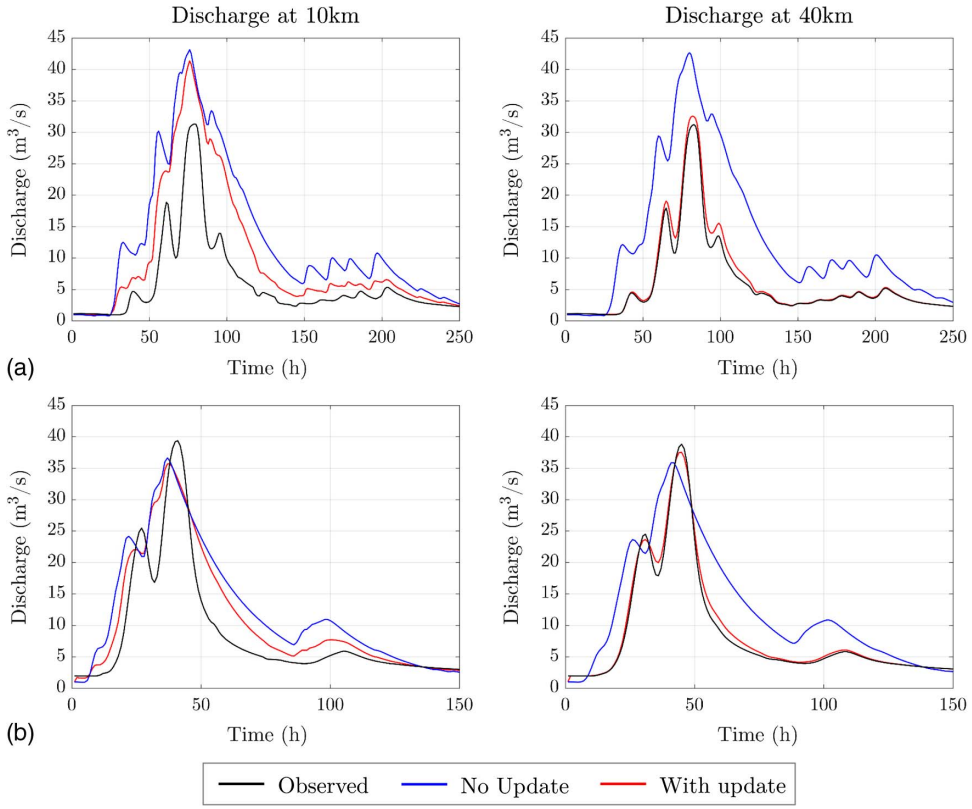

Fig. 4. Flow hydrographs obtained at two river sections (at 10 and 40 km) during (a) event A and (b) event B, assimilating observed water depth at 15 km from the upstream boundary conditions

the APs. However, such points were not the optimal sensor locations because they did not correspond to the minimum value of $\mathbf{P}$ among all the possible locations of the sensors. In addition, the optimal location of the sensor could depend on the accuracy requirements of the users along the reach. Although not aimed explicitly at optimizing the locations of the sensors, the results of this study could be used for network design, to complement the recent studies in the research area (Alfonso and Price 2012; Alfonso et al. 2010; Kollat et al. 2011).

The previous results were obtained in the case of a comparable model and boundary errors. However, these errors can differ on the basis of the accuracies of the input information and of the hydraulic model itself. For that reason, in the next analysis, the opposite values of the model $\mathbf{M_m}$ and the boundary $\mathbf{M_b}$ errors are selected to investigate their effects on the DA performances. Fig. 6 shows the NSE of the simulated flow for each section of the river for the APs at 15, 30, and 45 km. A large error in the boundary conditions (Scenario 3) provided higher NSE values than a large model error (Scenario 2). In Scenario 2, a change in the NSE values localized at the AP could be detected, whereas in Scenario 3, the NSE behavior was continuous and smooth in space. In both scenarios, downstream from the AP, the NSE improved up to an asymptotic value. The asymptote increased as the AP approached the reach outlet in the case of Scenario 2. This study demonstrated that when the model error was smaller than the boundary condition error (Scenario 3), the sensor should have been located closer to the upstream boundary in order to achieve an improvement over the entire river reach. Moreover, locating the sensor upstream would have allowed additional response time in predicting the flow and the WD at the outlet of the river reach, as demonstrated in the next

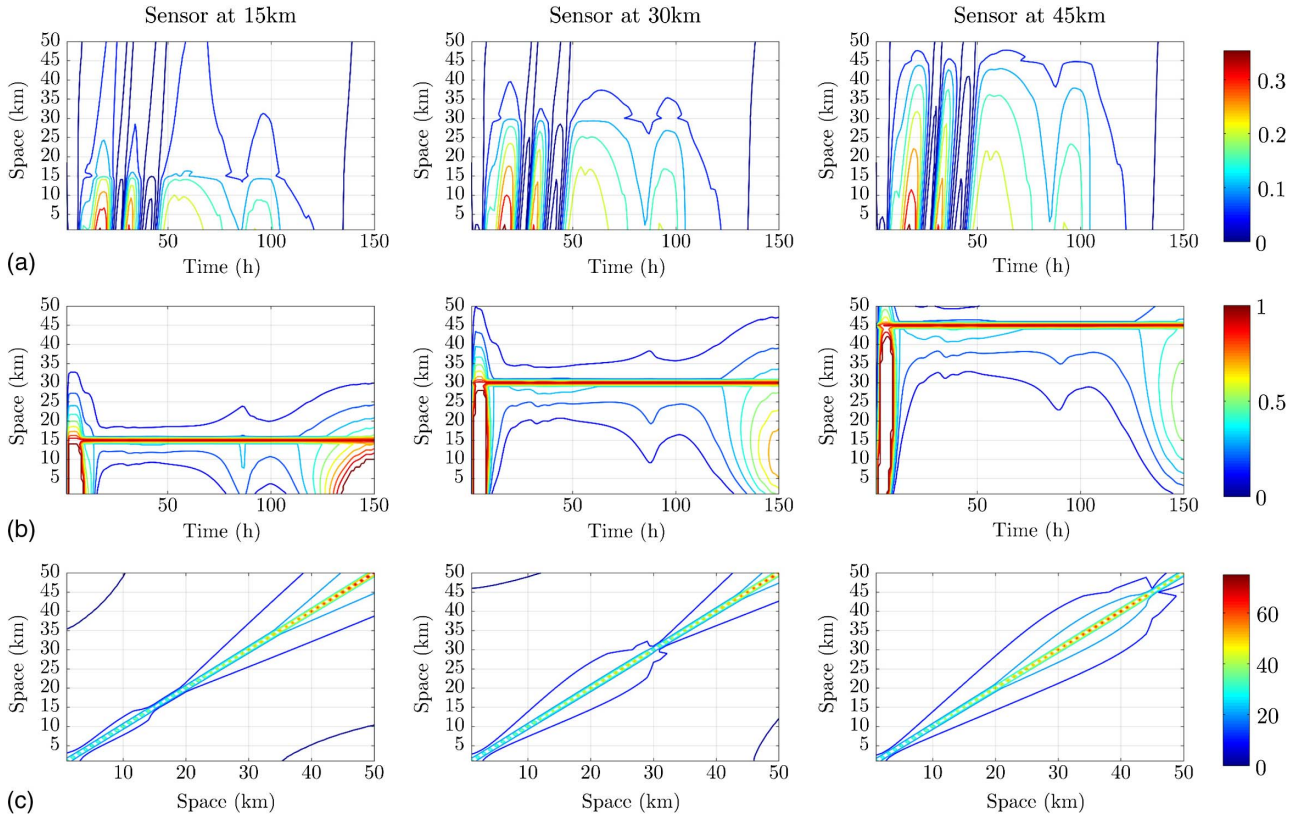

Fig. 5. (a) Spatial-temporal difference between observed and simulated (with update) water depth; (b) spatial-temporal value of the Kalman gain K; (c) model error covariance matrix $P(\mathrm{m}^6/\mathrm{s}^2)$ at 100 h; all graphs refer to flood event B in Scenario 1

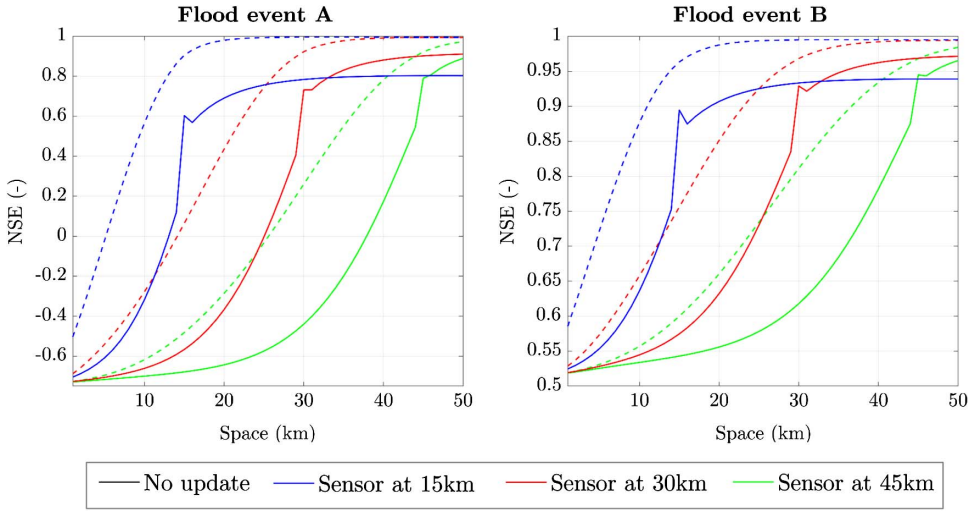

Fig. 6. NSE values achieved comparing observed and simulated water depth along the synthetic river during flood event A and B, assimilating flow observations at 15, 30, and 45 km; Scenario 2 is represented by continuous line and Scenario 3 by dashed line

section. On the other hand, when the model error was higher than the boundary condition error (Scenario 2), it was preferable to locate the sensor far from the upstream boundary condition, because the model itself was more uncertain.

Fig. 7 shows the discontinuous behavior of the NSE values in Scenario 2. In Scenario 2 (model error > boundary error), the model update, expressed in terms of Kalman gain **K**, was localized at the AP with a small improvement at the boundary conditions. On the other hand, in Scenario 3 (model error < boundary error), the largest value of **K** (the maximal gain by the KF) for the AP at 15 km was achieved at the boundary location and the gain was propagated downstream, generating the spatially continuous update shown in Fig. 6. Similar results were obtained when the *WD* observations were assimilated at 30 and 45 km.

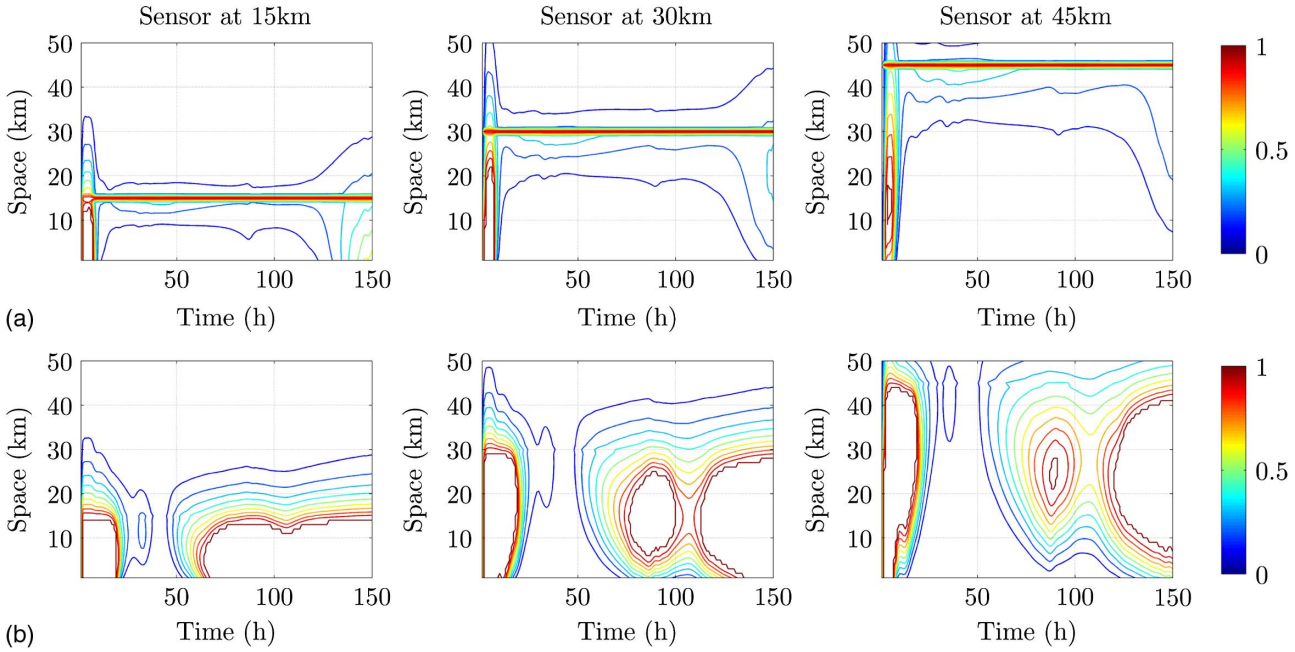

Fig. 7. Spatial-temporal variation of the Kalman gain for (a) Scenario 2; (b) Scenario 3 for flood event B, assimilating water depth at three locations (15, 30, and 45 km from upstream) along the synthetic river

Bacchiglione River

The results discussed in the preceding section were valid for a single river reach. The Bacchiglione River case is an example of a more complex river network. The boundary condition errors in reaches 1, 2, and 5, referred to as $\mathbf{M_{b,1}}$, $\mathbf{M_{b,2}}$, and $\mathbf{M_{b,5}}$, respectively, were estimated by comparing the hydrographs calculated using observed and forecast precipitation. Then, the flow observations from the hypothetical static physical sensors assumed to be at various locations were assimilated into the MC model. Fig. 8 shows the sensitivity of the model performances at the PA to the assimilation of the flow observations in different river reaches for the three considered flood events. Overall, the assimilation within reaches 1, 2, and 5 did not provide any additional improvement in comparison with the model results with no update. On the other hand, accurate forecasting was made by assimilating the flow observations via the KF for the reaches 3, 4, and 6. This could be related to the fact that these reaches were located upstream from the reaches 3, 4, and 6, providing a lower contribution in the overall model improvement at the PA. Among the remaining reaches, reach 6, located in the downstream part of the catchment, was the one that allowed the best model update. It is noteworthy that the impact of the sensor positioning was evaluated without assimilating all the observations at the same time. If all information were considered, the impact of certain additional measurements might be marginal or redundant.

In the previous analyses, the observations were assimilated at the middle point of each river reach. Fig. 9 shows the NSE values achieved by placing the hypothetical sensors at different locations along reaches 3, 4, and 6. As demonstrated above, the largest improvement for each river reach was obtained when the sensor was located in its upper part, because the boundary conditions had larger errors than the model. During flood event 1, placing the sensors toward the outlet of reach 6 significantly reduced the predictive skill by as much as 12%. For reaches 3 and 4, placing the sensors 5 km downstream slightly deteriorated the model results by as much as 1 and 4%, respectively. Table 2 provides the overall degradation in model performance from the upstream boundary to the outlet. One of the effects of the higher boundary condition error was that the assimilation of the observations from the sensors located at the downstream river reaches did not always lead to a higher NSE in comparison with the assimilation of observations from the upstream reaches. For example, the assimilation at 5 km from the upstream boundary of reach 4 provided a higher model improvement than the assimilation in the last section of reach 6 (Fig. 9).

Improvement in the predictive skill as a function of the lead time varied, depending on the sensor locations. The flow observations were assimilated up to the time of the forecast, after which the model ran in forecasting mode, that is, without updating. Fig. 10 shows the predictive skill expressed in terms of the NSE for all flood events as a function of the lead time. Various lines indicate the varying locations within the single river reach in which the observations were assimilated. For example, the black line in reach 3 indicates that the observation was assimilated at 2 km from upstream, that is, at the second cross section of that reach (because Δx is equal to 1,000 m). This figure shows the fast reduction in predictive skill for reach 6. This is due to the short travel time along reach 6 and consequently the short memory of the assimilation process. Considering the average flow velocity of 1 m/s and the 5-km length of the reach, the MC model loses the effect of assimilation after 1.5 h.

On the other hand, reach 3 provided the longest memory in the system due to an average travel time of approximately 5.5 h, even though the improvement was not as large as that for reach 6. Adding the sensor in reach 3 allowed it to gain as much as 2 and 4 h of lead time in comparison with the assimilation of observations for reaches 4 and 6, respectively. However, if only reach 3 is considered, the improvement in lead time due to the addition of a sensor did not change between the locations of 2 and 10 km from the upstream boundary. The choice of the optimal locations of the sensors should therefore be made to optimize both the NSE and the lead time, depending on the limited sampling budget. It is noteworthy that the low NSE associated with the MC results without updating was due to the underforecasted precipitation that resulted in underprediction of the discharge and the WD. Additional analyses should be carried out to assess the effect of assimilating distributed observations in the case of overforecasted precipitation.

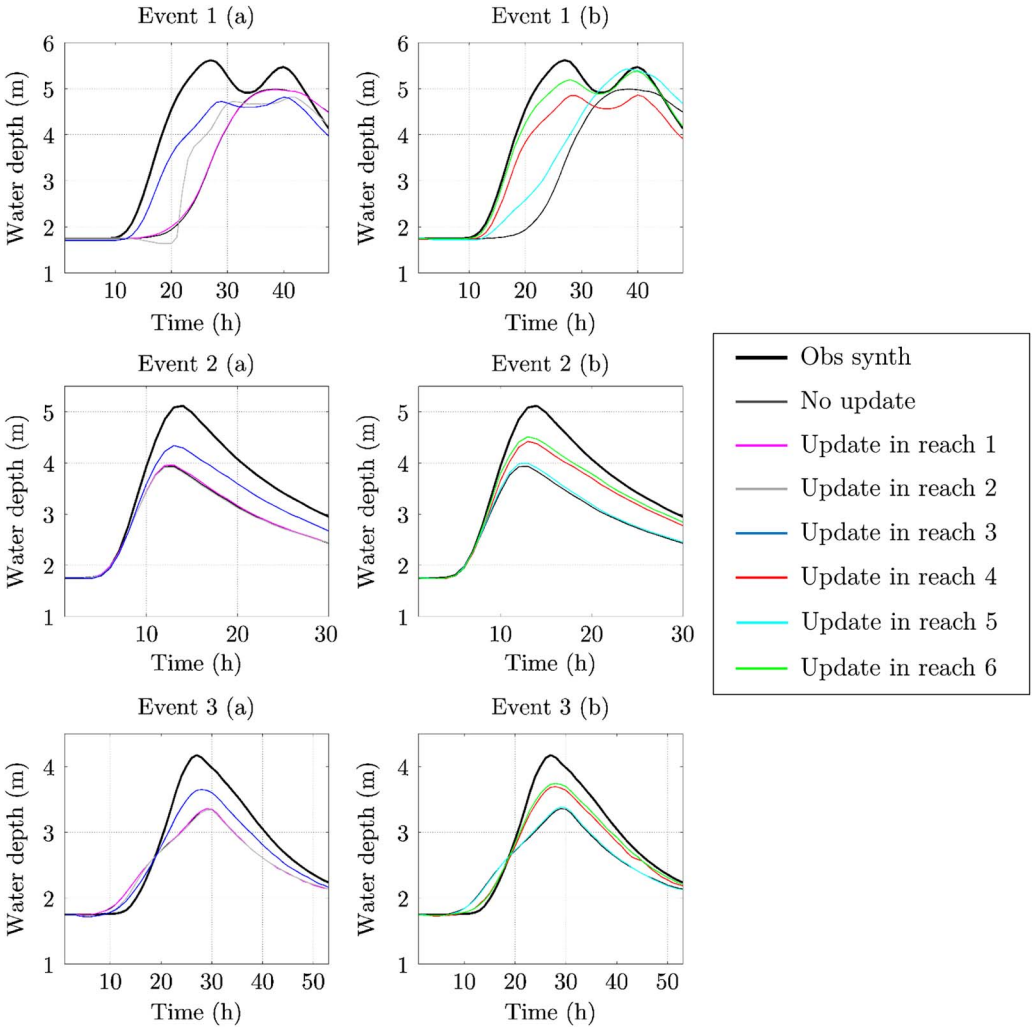

Fig. 8. Observed and simulated hydrographs at Ponte degli Angeli in the case of assimilation of water depth from sensors at the first upstream cross section in the reaches of the Bacchiglione River

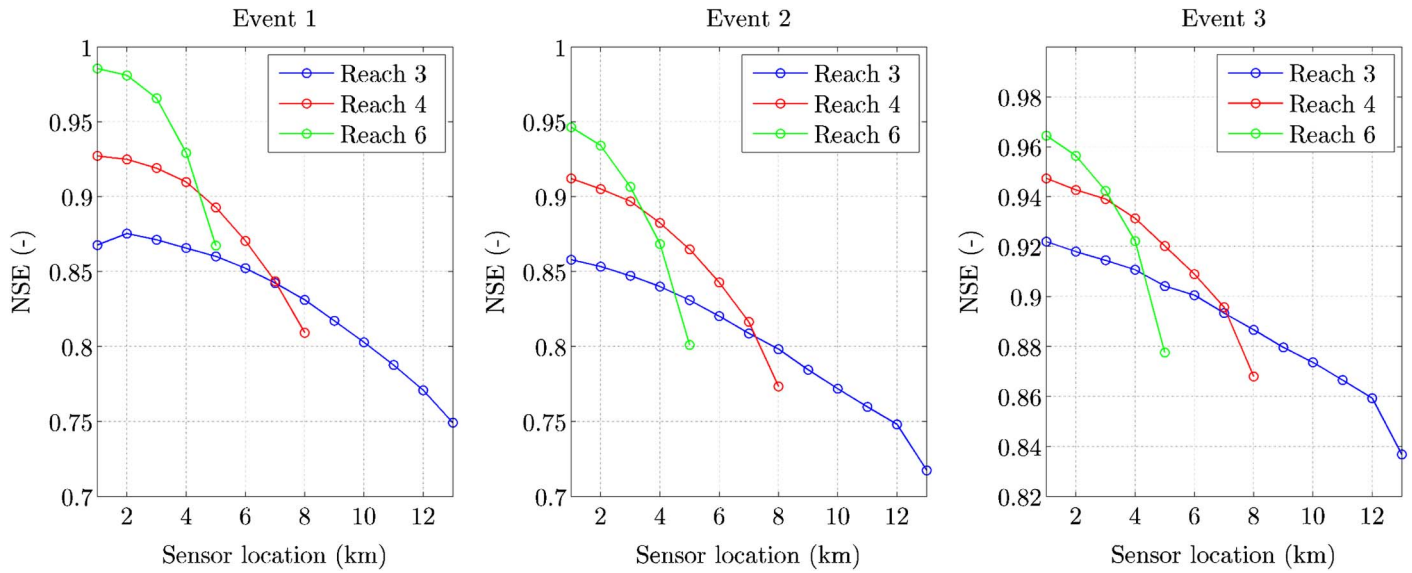

Fig. 9. NSE values obtained comparing simulated and observed water depth at Ponte degli Angeli, assimilating water depth data at different sensor locations (from upstream) along the reaches 3, 4, and 6

Table 2. NSE Values Obtained Assimilating Flow Observations in the First (Upstream) and Last (Downstream) Section of Reaches 3, 4, and 6 of the Bacchiglione River during Flood Events 1, 2, and 3

Reaches	Location	Event 1	Event 2	Event 3
3	Upstream	0.868	0.858	0.922
	Downstream	0.749	0.748	0.837
	Improvement (%)	−0.158	−0.196	−0.102
4	Upstream	0.927	0.912	0.947
	Downstream	0.809	0.773	0.868
	Improvement (%)	−0.146	−0.179	−0.091
6	Upstream	0.986	0.946	0.965
	Downstream	0.867	0.801	0.878
	Improvement (%)	−0.136	−0.181	−0.099

Real World Experiment

In contrast with the previous experiment, in this experiment the real-world *WD* observations from the sensors at the PM and the PA were assimilated. The observations were first converted to flow values and then assimilated into the hydrologic routing model to improve the flood forecasting at the PA. The performances were evaluated by comparing the observed and the simulated *WD* at the Bacchiglione outlet (the PA) during flood event 1 (which had a higher magnitude in comparison with flood events 2 and 3). Fig. 11 shows that the data assimilation tended to underpredict the observed *WD*, and that high correlations (above 0.85) were found between the observed and the simulated *WD*. Fig. 11 shows the Taylor diagram that plots the standard deviation, correlation coefficient R, and root mean squared difference of the

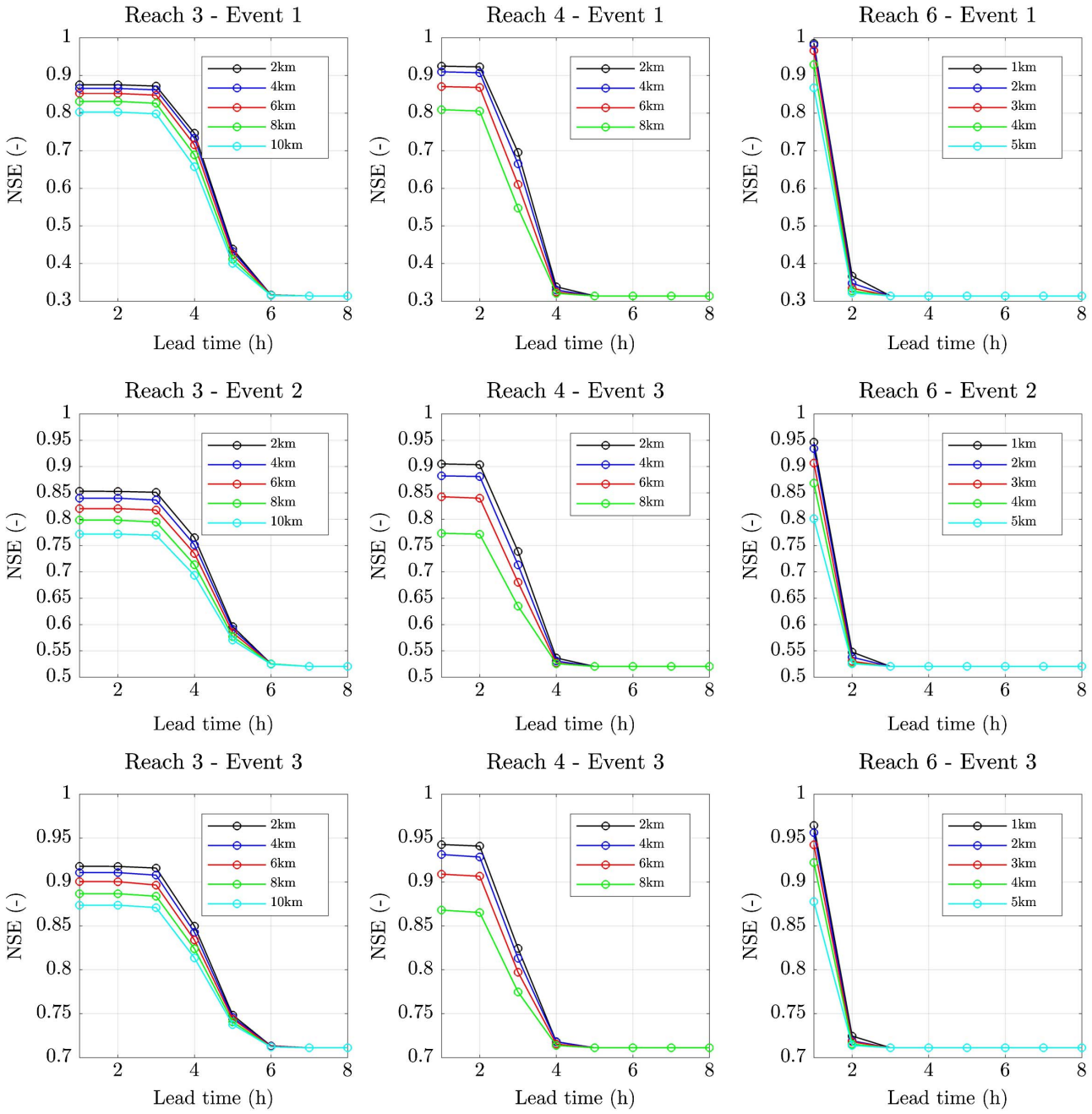

Fig. 10. NSE values obtained for different lead times, assimilating water depth at different sensor locations from upstream, along reaches 3, 4, and 6 during three flood events

J. Hydrol. Eng., 2018, 23(6): 04018018

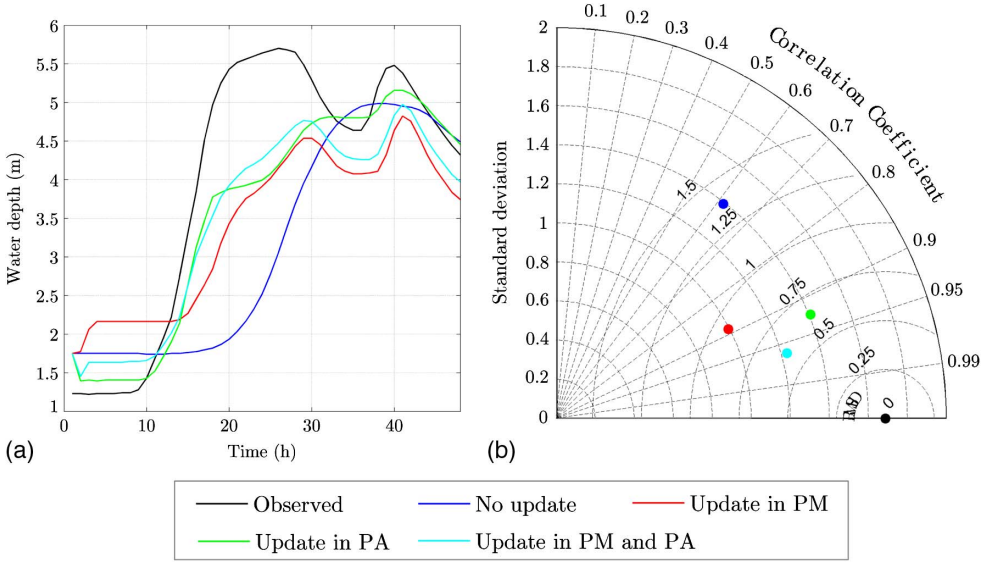

Fig. 11. (a) Observed and simulated water depth hydrographs; (b) Taylor diagram representing the statistics of the hydrographs obtained in the real-world experiment along the Bacchiglione River

simulated and updated hydrographs, and hence summarizes these results. The figure shows that assimilating the observations at the PM was not as effective as assimilating the observations at the PA. Assimilating the observations at both the PA and the PM showed an improvement in the correlation and a reduction in the root mean square error. Assimilation of the observed flow at the PM reduced the simulated *WD*, which in combination with uncertain internal boundary conditions I6 and I7 (Fig. 1), resulted in the underestimation at the PA, as reported in the hydrographs and the Taylor diagram shown in Fig. 11. Finally, Fig. 12 shows the benefits, in terms of the increased predictive skill and lead time, of assimilating the observations upstream of the targeted forecast point (PA). Although assimilation at the PA was useful in improving the NSE, the R, and the bias for short lead times, assimilating the observations at the PM helped to maximally exploit the memory of the system and improve model performance for longer lead times up to 4 h (the travel time from the PM to the PA).

Conclusions and Further Research

The main objective of this study was to assess the effects of the locations of the sensors and the specifications of both the model and the boundary conditions error matrices on the assimilation of distributed flow observations. For this purpose, a Muskingum-Cunge (MC) routing model was implemented along a synthetic river reach and the Bacchiglione River. Then, the Kalman filter was applied to the MC model to assimilate the flow observations. Synthetic and real-world experiments were conducted.

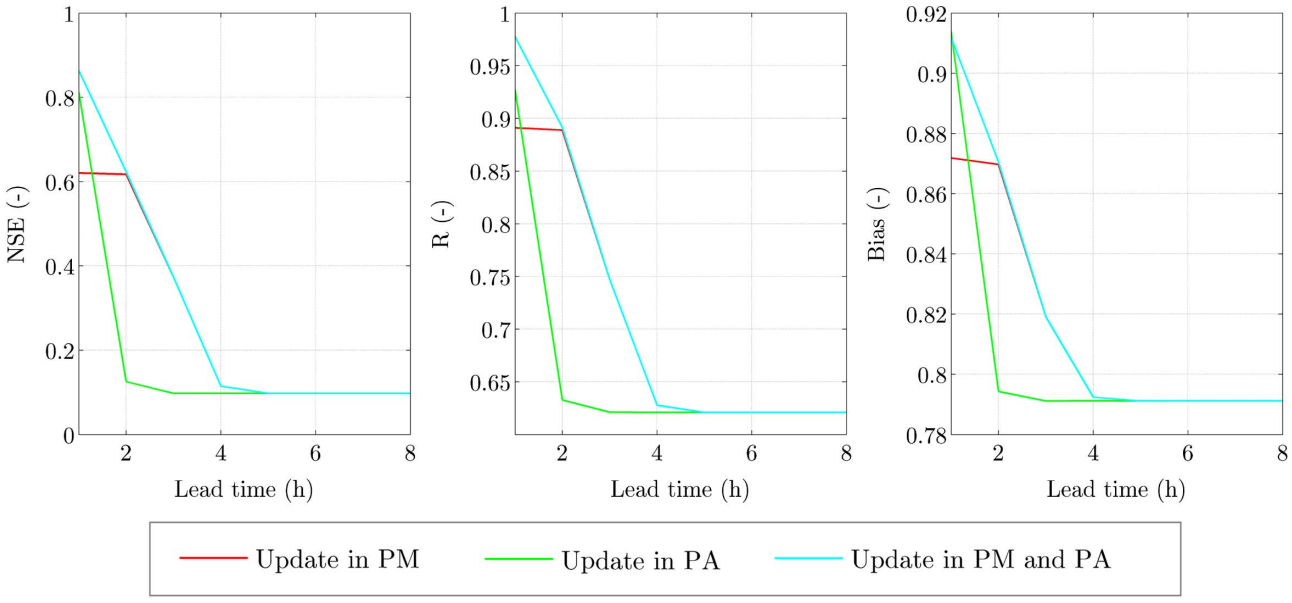

Fig. 12. NSE values obtained for different lead times at PA during the real-world experiments along the Bacchiglione River

This research demonstrated that both the location of sensors used to generate data for model updating purposes and the characterization of the errors in the boundary conditions and in the model itself had a significant effect on the model performance, in particular for different lead time values. The results of this study are not limited to the intuitive statement that observations far from the outlet sections contribute to improved skills for a longer lead time, whereas observations at the outlet lead to higher model improvement but for a shorter lead time.

The results related to the synthetic river showed that data assimilation induced an update along the whole river reach because of the distributed nature of the Kalman filter, whereas due to the MC model structure, the updating effect was more significant at the downstream than at the upstream. It was demonstrated that the magnitude of a model error affected the assimilation performance. On the one hand, a large error in the boundary condition (e.g., input from the hydrological model) tended to better improve the water profile when the assimilation point was closer to the boundary location and a smooth update was achieved along the river reach. On the other hand, for the model errors larger than the boundary errors, a localized update was obtained at the sensor locations and the good model performances were achieved if the sensor was located close to the reach outlet. That is the reason, in this last case, that it might be suggested to locate the sensor downstream of the river reach to maximize the model improvement at the river outlet.

The results obtained for the Bacchiglione River showed that only the assimilation within the main river channel (reaches 3, 4, and 6) provided additional improvement to the model results at the outlet (Ponte degli Angeli) in comparison with the model with no update. Because the upstream boundary conditions had a higher error than the routing model itself, high values of the NSE were obtained from assimilating the flow observations coming from the upstream part of each river reach, as previously demonstrated. The reach closer to the downstream outlet (reach 6) provided the best model performances. However, reach 6 tended to lose the assimilation effect faster than reaches 3 and 4, due to its shorter travel time.

For this reason, the choice of the optimal location of the sensors should be a compromise between the best NSE value and the best forecasting capability of the model. In addition, the choice of the optimal location of the sensors should be made considering the limited sampling budget and the accessibility and safety conditions of the monitored area. This study does not explicitly aim to optimize the locations of the sensors. However, these results provide an important insight that can be taken into account as an additional criterion to improve and complement the existing design and evaluation monitoring network design methodologies. This paper aims to provide the water managers of the Bacchiglione River all the required information to rank the potential sampling locations on the basis of an improved model performance and the required lead time, without considering the sampling budget and accessibility of the monitored area.

Despite the encouraging results achieved, there remain some limitations in the presented methodology; hence there is room for further improving the flow routing prediction by data assimilation. First, the effects of the sensor locations on the flood forecasting capabilities might change, depending on the structure and characteristics of the hydraulic model [as pointed out in Thiboult and Anctil (2015) and Thiboult et al. (2016)] and the DA method. Both the Muskingum-Cunge and the Kalman filter are linear methods, so that the conclusions of this study cannot be generalized to highly nonlinear models, which, for example, incorporate dynamic waves and water-quality models (Wang et al. 2016). For this reason, additional analyses with nonlinear models should be performed in order to sort out the general findings that are less affected by the methods and modeling domains.

Second, an obvious limitation of this study is the small number of case studies and the small sampling size of the considered flood events, which makes the results of this research very case-specific. For this reason, the results of this study should be validated by applying the same methodology to different rivers that have diverse characteristics. In fact, because of the practical limitations for sensor location (e.g., sensors are usually located at the bridges where flow disturbance is minimal), the conclusion that upstream locations could bring more benefits may not be valid for other rivers.

Third, the synthetic observations were used mainly because of the lack of distributed flow data in the two case studies. Such synthetic observations may be biased, as highlighted in Viero (2018). For that reason, additional analyses considering real-time observations should be carried out to assess the impact of the DA methods in the real world.

Finally, the temporal and spatial correlation of the model error was not considered; therefore, additional analysis, for example, the correlation within the model error matrix using autoregressive models, can be recommended.

Acknowledgments

This research was partly funded by the European FP7 Project WeSenseIt (Citizen Observatory of Water), grant agreement 308429. Some research components were developed in the framework of grant 17-77-30006 of the Russian Science Foundation and the Hydroinformatics research fund of IHE Delft. Data used were supplied by the British Atmospheric Data Centre from the NERC Hydrological Radar Experiment Dataset (http://www.badc.rl.ac.uk/data/hyrex/) and by the Alto Adriatico Water Authority. Support for Seongjin Noh and Dong-Jun Seo was provided by the National Science Foundation under Grant CyberSEES-1442735 (principal investigator Dong-Jun Seo, University of Texas at Arlington). This support is gratefully acknowledged.

References

Abbott, M. B., and Ionescu, F. (1967). "On the numerical computation of nearly horizontal flows." *J. Hydraul. Res.*, 5(2), 97–117.

Abebe, A., and Price, R. (2004). "Information theory and neural networks for managing uncertainty in flood routing." *J. Comput. Civ. Eng.*, 10.1061/(ASCE)0887-3801(2004)18:4(373), 373–380.

Aerts, J. C. J. H., Botzen, W. J. W., Emanuel, K., Lin, N., de Moel, H., and Michel-Kerjan, E. O. (2014). "Evaluating flood resilience strategies for coastal megacities." *Science*, 344(6183), 473–475.

Alfonso, L., Lobbrecht, A., and Price, R. (2010). "Information theory based approach for location of monitoring water level gauges in polders." *Water Resour. Res.*, 46(3), W03528.

Alfonso, L., and Price, R. (2012). "Coupling hydrodynamic models and value of information for designing stage monitoring networks." *Water Resour. Res.*, 48(8), W08530.

Alfonso, L., and Tefferi, M. (2015). "Effects of uncertain control in transport of water in a river-wetland system of the Low Magdalena River, Colombia." *Transport of water versus transport over water*, C. Ocampo-Martinez and R. R. Negenborn, eds., Springer International, Cham, Switzerland, 131–144.

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., and Alsdorf, D. E. (2007). "Prospects for river discharge and depth estimation through assimilation of swathaltimetry into a raster-based hydrodynamics model." *Geophys. Res. Lett.*, 34(10), L10403.

Andreadis, K. M., and Schumann, G. J. P. (2014). "Estimating the impact of satellite observations on the predictability of large-scale hydraulic models." *Adv. Water Resour.*, 73, 44–54.

Arnold, C. P., and Dey, C. H. (1986). "Observing-systems simulation experiments: Past, present, and future." *Bull. Am. Meteorol. Soc.*, 67(6), 687–695.

Babovic, V., and Fuhrman, D. (2002). "Data assimilation of local model error forecasts in a deterministic model." *Int. J. Numer. Methods Fluids*, 39(10), 887–918.

Barati, R., Akbari, H. G., and Rahimi, S. (2013). "Flood routing of an unmanaged river basin using Muskingum-Cunge model; field application and numerical experiments." *Caspian J. Appl. Sci. Res.*, 2, 8–20.

Biancamaria, S., et al. (2011). "Assimilation of virtual wide swath altimetry to improve arctic river modelling." *Remote Sens. Environ.*, 115(2), 373–381.

Chacon-Hurtado, J. C., Alfonso, L., and Solomatine, D. P. (2017). "Rainfall and streamflow sensor network design: A review of applications, classification, and a proposed framework." *Hydrol. Earth Syst. Sci.*, 21(6), 3071–3091.

Clark, M. P., et al. (2008). "Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observation to update states in a distributed hydrological model." *Adv. Water Resour.*, 31(10), 1309–1324.

Cunge, J. A. (1969). "On the subject of a flood propagation computation method (Muskingum method)." *J. Hydraul. Res.*, 7(2), 205–230.

Di Baldassarre, G., Montanari, A., Lins, H., Koutsoyiannis, D., Brandimarte, L., and Bloeschl, G. (2010). "Flood fatalities in Africa: From diagnosis to mitigation." *Geophys. Res. Lett.*, 37(22), L22402.

Errico, R. M., et al. (2013). "Development and validation of observing-system simulation experiments at NASA's Global Modeling and Assimilation Office." *Q. J. R. Meteorol. Soc.*, 139(674), 1162–1178.

Errico, R. M., and Privé, N. C. (2014). "An estimate of some analysis-error statistics using the Global Modeling and Assimilation Office observing-system simulation framework." *Q. J. R. Meteorol. Soc.*, 140(680), 1005–1012.

European Environment Agency. (2005). "Vulnerability and adaptation to climate change: Scoping report." Copenhagen, Denmark.

Ferri, M., Monego, M., Norbiato, D., Baruffi, F., Toffolon, C., and Casarin, R. (2012). "AMICO: La piattaforma previsionale per i bacini idrografici del Nord Est Adriatico (I)." *Proc., XXXIII Conf. of Hydraulics and Hydraulic Engineering*, B. Bacchi, R. Ranzi, and M. Tomirotti, eds., Brescia, Italy, 10.

García-Pintado, J., Neal, J. C., Mason, D. C., Dance, S. L., and Bates, P. D. (2013). "Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling." *J. Hydrol.*, 495, 252–266.

Georgakakos, A. P., Georgakakos, K. P., and Baltas, E. A. (1990). "A state-space model for hydrologic river routing." *Water Resour. Res.*, 26(5), 827–838.

Giustarini, L., et al. (2011). "Assimilating SAR-derived water level data into a hydraulic model: A case study." *Hydrol. Earth Syst. Sci.*, 15(7), 2349–2365.

Gochis, D. J., Yu, W., and Yates, D. N. (2013). "The WRF-Hydro model technical description and user's guide." ⟨http://www.ral.ucar.edu/projects/wrf_hydro/⟩ (Nov. 21, 2017).

Goetzinger, J., and Bardossy, A. (2008). "Generic error model for calibration and uncertainty estimation of hydrological models." *Water Resour. Res.*, 44(12), W00B07.

Heatherman, W. J. (2008). "Flood routing on small streams: A review of the Muskingum-Cunge, cascading reservoirs and full dynamic solutions." Ph.D. dissertation, Univ. of Kansas, Lawrence, KS, 373.

Hinkel, J., et al. (2014). "Coastal flood damage and adaptation costs under 21st century sea-level rise." *Proc. Natl. Acad. Sci. U. S. A.*, 111(9), 3292–3297.

Hostache, R., Lai, X., Monnier, J., and Puech, C. (2010). "Assimilation of spatially distributed water levels into a shallow-water flood model. II: Use of a remote sensing image of Mosel River." *J. Hydrol.*, 390(3–4), 257–268.

Jongman, B., et al. (2014). "Increasing stress on disaster-risk finance due to large floods." *Nat. Clim. Change*, 4(4), 264–268.

Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems." *J. Basic Eng.*, 82(1), 35–45.

Kamel, A. H. (2008). "Application of a hydrodynamic MIKE 11 model for the Euphrates River in Iraq." *Slovak J. Civ. Eng.*, 2, 1–7.

Kim, Y., Tachikawa, Y., Shiiba, M., Kim, S., Yorozu, K., and Noh, S. J. (2013). "Simultaneous estimation of inflow and channel roughness using 2D hydraulic model and particle filters." *J. Flood Risk Manage.*, 6(2), 112–123.

Kollat, J. B., Reed, P. M., and Maxwell, R. M. (2011). "Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics." *Water Resour. Res.*, 47(2), W02529.

Koussis, A. D. (1983). "Accuracy criteria in diffusion routing: A discussion." *J. Hydraul. Div.*, 10.1061/(ASCE)0733-9429(1983)109:5(803), 803–806.

Lee, H., Liu, Y., He, M., Demargne, J., and Seo, D. J. (2008). "Variational assimilation of streamflow into the three-parameter Muskingum routing model for improved operational river flow forecasting." *EGU Fall Meeting Abstracts*, Vienna, Austria.

Li, Y., Ryu, D., Western, A. W., and Wang, Q. J. (2015). "Assimilation of stream discharge for flood forecasting: Updating a semidistributed model with an integrated data assimilation scheme." *Water Resour. Res.*, 51(5), 3238–3258.

Liu, Y., Lee, H., Seo, D. J., Brown, J., Corby, R., and Howieson, T. (2008). "Ensemble data assimilation for channel flow routing to improve operational hydrologic forecasting." *AGU Fall Meeting Abstracts*, San Francisco.

Liu, Y. Q., et al. (2012). "Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities." *Hydrol. Earth Syst. Sci.*, 16(10), 3863–3887.

Liu, Y. Q., and Gupta, H. V. (2007). "Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework." *Water Resour. Res.*, 43(7), 1944–7973.

Mason, D. C., Schumann, G. J. P., Neal, J. C., Garcia-Pintado, J., and Bates, P. D. (2012). "Automatic near real-time selection of flood water levels from high resolution synthetic aperture radar images for assimilation into hydraulic models: A case study." *Remote Sens. Environ.*, 124, 705–716.

Matgen, P., et al. (2010). "Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the particle filter: Proof of concept." *Hydrol. Earth Syst. Sci.*, 14(9), 1773–1785.

Mazzoleni, M. (2017). "Improving flood prediction assimilating uncertain crowdsourced data into hydrologic and hydraulic models." Ph.D. thesis, CRC Press, Leiden, Netherlands.

Mazzoleni, M., et al. (2017a). "Can assimilation of crowdsourced data in hydrological modelling improve flood prediction?." *Hydrol. Earth Syst. Sci.*, 21(2), 839–861.

Mazzoleni, M., Alfonso, L., Chacon-Hurtado, J., and Solomatine, D. P. (2015). "Assimilating uncertain, dynamic and intermittent streamflow observations in hydrological models." *Adv. Water Resour.*, 83, 323–339.

Mazzoleni, M., Alfonso, L., and Solomatine, D. P. (2017b). "Influence of spatial distribution of sensors and observation accuracy on the assimilation of distributed streamflow data in hydrological modeling." *Hydrol. Sci. J.*, 62(3), 389–407.

McLaughlin, D. (2002). "An integrated approach to hydrologic data assimilation: Interpolation, smoothing and filtering." *Adv. Water Resour.*, 25(8–12), 1275–1286.

MIKE11 [Computer software]. Danish Hydraulic Institute, Hørsholm, Denmark.

Mishra, A. K., and Coulibaly, P. (2009). "Developments in hydrometric network design: A review." *Rev. Geophys.*, 47(2), RG2001.

Moradkhani, H., Hsu, K. L., Gupta, H., and Sorooshian, S. (2005). "Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter." *Water Resour. Res.*, 41(5), 1–17.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models. I: A discussion of principles." *J. Hydrol.*, 10(3), 282–290.

Neal, J., Atkinson, P. M., and Hutton, C. W. (2007). "Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements." *J. Hydrol.*, 336(3–4), 401–415.

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., and Pappenberger, F. (2009). "A data assimilation approach to discharge estimation from space." *Hydrol. Process.*, 23(25), 3641–3649.
Neal, J. C., Atkinson, P. M., and Hutton, C. W. (2012). "Adaptive space-time sampling with wireless sensor nodes for flood forecasting." *J. Hydrol.*, 414–415, 136–147.
Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S. (2013). "Ensemble Kalman filtering and particle filtering in a lag-time window for short-term streamflow forecasting with a distributed hydrologic model." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000751, 1684–1696.
Pappenberger, F., Matgen, P., Beven, K. J., Henry, J.-B., Pfister, L., and Fraipont de, P. (2006). "Influence of uncertain boundary conditions and model structure on flood inundation predictions." *Adv. Water Resour.*, 29(10), 1430–1449.
Ponce, V. M., and Chaganti, P. V. (1994). "Variable-parameter Muskingum-Cunge revisited." *J. Hydrol.*, 162(3–4), 433–439.
Ponce, V. M., and Lugo, A. (2001). "Modeling looped ratings in Muskingum-Cunge routing." *J. Hydrol. Eng.*, 10.1061/(ASCE)1084-0699(2001)6:2(119), 119–124.
Puente, C. E., and Bras, R. L. (1987). "Application of nonlinear filtering in the real time forecasting of river flows." *Water Resour. Res.*, 23(4), 675–682.
Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L. (2016). "Improving the realism of hydrologic model functioning through multivariate parameter estimation." *Water Resour. Res.*, 52(10), 7779–7792.
Rakovec, O., Weerts, A. H., Hazenberg, P. F., Torfs, P. J. J., and Uijlenhoet, R. (2012). "State updating of a distributed hydrological model with ensemble kalman filtering: Effects of updating frequency and observation network density on forecast accuracy." *Hydrol. Earth Syst. Sci.*, 16(9), 3435–3449.
Refsgard, J. C. (1997). "Validation and intercomparison of different updating procedures for real-time forecasting." *Nordic Hydrol.*, 28(2), 65–84.
Reichle, R. H. (2008). "Data assimilation methods in the Earth sciences." *Adv. Water Res.*, 31(11), 1411–1418.
Ricci, S., Piacentini, A., Thual, O., Le Pape, E., and Jonville, G. (2011). "Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting." *Hydrol. Earth Syst. Sci.*, 15(11), 3555–3575.
Robinson, A. R., Lermusiaux, P. F. J., and Sloan, N. Q. (1998). "Data assimilation." *Sea*, 10, 541–594.
Romanowicz, R. J., Young, P. C., and Beven, K. J. (2006). "Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom." *Water Resour. Res.*, 42(6), W06407.
Sakov, P., Evensen, G., and Bertino, L. (2010). "Asynchronous data assimilation with the EnKF." *Tellus A*, 62(1), 24–29.
Sun, L., Seidou, O., Nistor, I., and Liu, K. (2016). "Review of the Kalman-type hydrological data assimilation." *Hydrol. Sci. J.*, 61(13), 2348–2366.
Szilagyi, J., and Szollosi-Nagy, A. (2010). *Recurvive streamflow forecasing: A state space approach*, CRC Press, London.
Thiboult, A., and Anctil, F. (2015). "On the difficulty to optimally implement the ensemble Kalman filter: An experiment based on many hydrological models and catchments." *J. Hydrol.*, 529(3), 1147–1160.
Thiboult, A., Anctil, F., and Boucher, M.-A. (2016). "Accounting for three sources of uncertainty in ensemble hydrological forecasting." *Hydrol. Earth Syst. Sci.*, 20, 1809–1825.
Todini, E. (2007). "A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach." *Hydrol. Earth Syst. Sci.*, 11(5), 1645–1659.
Todini, E., et al. (2005). "ACTIF Best practice paper-understanding and reducing uncertainty in flood forecasting." *Int. Conf. Innovation, Advances and Implementation of Flood Forecasting Technology*, Tromsø, Norway.
Verlaan, M., and Sumihar, J. (2016). "Observation impact analysis methods for storm surge forecasting systems." *Ocean Dyn.*, 66(2), 221–241.
Viero, D. P. (2018). "Comment on Can assimilation of crowdsourced data in hydrological modelling improve flood prediction? by Mazzoleni et al. (2017)." *Hydrol. Earth Syst. Sci.*, 22(1), 171–177.
Viessman, W., and Lewis, G. L. (2003). *Introduction to hydrology*, 5th Ed., Pearson Academic, New York.
Vojinovic, Z., Vojislav, K., and Babovic, V. (2003). "Hybrid approach for modeling wet weather response in wastewater systems." *J. Water Resour. Plann. Manage.*, 10.1061/(ASCE)0733-9496(2003)129:6(511), 511–521.
Wang, X., Zhang, J., and Babovic, B. (2016). "Improving real-time forecasting of water quality indicators with combination of process-based models and data assimilation technique." *Ecol. Indic.*, 66, 428–439.
Weerts, A. H., and El Serafy, G. Y. H. (2006). "Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models." *Water Resour. Res.*, 42(9), W09403.
World Meteorological Organization. (1992). "Simulated real-time intercomparison of hydrological models." *Operational Hydrology Rep. 38*, Geneva.
Yan, K., Di Baldassarre, G., Solomatine, D. P., and Schumann, G. J.-P. (2015). "A review of low-cost space-borne data for flood modelling: Topography, flood extent and water level." *Hydrol. Process.*, 29(15), 3368–3387.