Can assimilation of crowdsourced streamflow observations in hydrological modelling improve flood prediction?

M. Mazzoleni[1], M. Verlaan[2], L. Alfonso[1], M. Monego[3], D. Norbiato[3], M. Ferri[3] and D. P. Solomatine[1,4]

[1]{UNESCO-IHE Institute for Water Education, Delft, The Netherlands}

[2]{Deltares, Delft, The Netherlands}

[3]{Alto Adriatico Water Authority, Venice, Italy}

[4]{Delft University of Technology, Water Resources Section, Delft, The Netherlands}

Correspondence to: M. Mazzoleni (m.mazzoleni@unesco-ihe.org)

Abstract

Monitoring stations have been used for decades to properly measure hydrological variables and better predict floods. To this end, methods to incorporate such observations into mathematical water models have also being developed, including data assimilation. Besides, in recent years, the continued technological improvement has stimulated the spread of low-cost sensors that allow for employing crowdsourced and obtain observations of hydrological variables in a more distributed way than the classic static physical sensors allow. However, such measurements have the main disadvantage to have asynchronous arrival frequency and variable accuracy. For this reason, this is one of the first studies that aims to demonstrate that crowdsourced streamflow observations can improve flood prediction if integrated in hydrological models. Two different types of hydrological models, applied to four case studies, are considered. Realistic (albeit synthetic) streamflow observations are used to represent crowdsourced streamflow observations in both case studies. Consistent results are found across the all case studies. It is found that the accuracy of the observations influences the model results more than the actual (irregular) moments in which the streamflow observations are assimilated into the hydrological models. This study demonstrates how networks of low-cost sensors can complement traditional networks of physical sensors and improve the accuracy of flood forecasting.

1 Introduction

Observations of hydrological variables measured by physical sensors have been increasingly integrated into mathematical models by means of model updating methods. The use of these techniques allows for the reduction of intrinsic model uncertainty and improves the flood forecasting accuracy (Todini et al., 2005). The main idea behind model updating techniques is to either update model input, states, parameters or outputs as new observations become available (Refsgaard, 1997; WMO, 1992). Input update is the classical method used in operational forecasting as uncertainties of the input data can be considered as the main source of uncertainty (Bergström, 1991; Canizares et al., 1998; Todini et al., 2005). Regarding the state updating, Kalman filtering approaches such as Kalman filter (Kalman, 1960), extended Kalman filter (Aubert et al., 2003; Kalman, 1960; Madsen and Cañizares, 1999; Verlaan, 1998) or Ensemble Kalman filter (EnKF, Evensen, 2006) are ones of the most used when new observations are available.

Due to the complex nature of the hydrological processes, spatially and temporally distributed measurements are needed in the model updating procedures to ensure a proper flood prediction (Clark et al., 2008; Mazzoleni et al., 2015; Rakovec et al., 2012). However, traditional physical sensors require proper maintenance and personnel which can be very expensive in case of a vast network. For this reason, the technological improvement led to the spread of low-cost sensors used to measure hydrological variables such as water level or precipitation in a distributed way. An example of such sensors, defined in the following as “social sensor”, is a smart-phone camera used to measure the water level at a staff gauge with an associate QR code used to infer the spatial location of the measurement (see Figure 1). The main advance of using these type of sensors is that they can be used not only by technicians but also by regular citizens, and that due to their reduced cost a more spatially distributed coverage can be achieved. The idea of designing such alternative networks of low-cost social sensors and using the obtained crowdsourced observations is the base of the EU-FP7 WeSenseIt project (2012-2016), which also sponsors this research. Various other projects have also been initiated in order to assess the usefulness of crowdsourced observations inferred by low-cost sensors owned by citizens. For instance, in the project CrowdHydrology (Lowry and Fienen, 2013), a method to monitor stream stage at designated gauging staffs using crowd source-based text messages of water levels is developed using untrained observers. Cifelli et al. (2005) described a community-based network of volunteers (CoCoRaHS), engaged in collecting precipitation measurements of rain,

hail and snow. An example of hydrological monitoring, established in 2009, of rainfall and streamflow values within the Andean ecosystems of Piura, Peru, based on citizen observations is reported in Célleri et al. (2009). Degrossi et al. (2013) used a network of wireless sensors in order to map the water level in two rivers passing by Sao Carlos, Brazil. Recently, the iSPUW Project is aims to integrate data from advanced weather radar systems, innovative wireless sensors and crowdsourcing of data via mobile applications in order to better predict flood events in the urban water systems of the Dallas-Fort Worth Metroplex (ISPUW, 2015; Seo et al., 2014). Other examples of crowdsourced water-related water-related information include the so-called Crowdmap platform for collecting and communicating the information about the floods in Australia in 2011 (ABC, 2011), and informing citizens about the proper time for water supply in an intermittent water system (Alfonso, 2006; Au et al., 2000; Roy et al., 2012). A detailed and interesting review of the examples of citizen science applications in hydrology and water resources science is provided by Buytaert et al. (2014). In this review study, the potential of citizen science, based on robust, cheap, and low-maintenance sensing equipment, to complement more traditional ways of scientific data collection for hydrological sciences and water resources management is explored. In order to study the challenges and opportunities in the integration of hydrologically-oriented citizen science in water resources management, four case studies from remote mountain region (e.g. the Peruvian Andes) are considered.

The traditional hydrological observations from physical sensors have a well-defined structure in terms of frequency and accuracy. On the other hand, crowdsourced observations are provided by citizens with varying experience of measuring environmental data and little connections between each other, and the consequence is that the low correlation between the measurements might be observed. So far, in operational hydrology practice, the added value of crowdsourced data it is not integrated into the forecasting models but just used to compare the model results with the observations in a post-event analysis. This can be related to the intrinsic variable accuracy, due to the lack of confidence in the data quality from such heterogeneous sensors, and the variable life-span of the crowdsourced observations.

Regarding data quality, Bordogna et al. (2014) and Tulloch and Szabo (2012) stated that quality control mechanisms should consider contextual conditions to deduce indicators about reliability (expertise level), credibility (volunteer group) and performance of volunteers such as accuracy, completeness and precision level. Bird et al. (2014) addressed the issue of data quality in conservation ecology by means of new statistical tools to assess random error and bias in such

observations. Cortes et al. (2014) evaluated data quality by distinguishing the in-situ data collected between a volunteer and a technician and comparing the most frequent value reported at a given location. They also gave some range of precision according to the rating scales. With in-situ exercises, it might be possible to have an indication of the reliability of data collected (expertise level). However, this indication does not necessarily lead to a conclusion of high, medium or low accuracy every time a streamflow observation of a contributor is received. In addition, such approach is not enough at operational level to define accuracy in data quality. In fact, every time a crowdsourced observation is received in real-time, the reliability and accuracy of observations should be identified. To do so, one possible approach could be to filter out the measurements following a geographic approach which defines semantic rules governing what can occur at a given location (e.g. Vandecasteele and Devillers, 2013). Another approach could be to compare measurements collected within a pre-defined time-window in order to calculate the most frequent value, the mean and the standard deviation.

Regarding the variable life-span, crowdsourced observations can be defined as *asynchronous* because do not have predefined rules about the arrival frequency (the observation might be sent just once, occasionally or at irregular time steps which can be smaller than the model time step) and accuracy. In a recent paper, Mazzoleni et al. (2015) presented results of the study of the effects of distributed synthetic streamflow observations having synchronous intermittent temporal behaviour and variable accuracy in a semi-distributed hydrological model. It has been shown that the integration of distributed uncertain intermittent observations with single measurements coming from physical sensors would allow for the further improvements in model accuracy. However, we have not considered the possibility that the asynchronous observations might be coming at the moments not coordinated with the model time steps. A possible solution to handle asynchronous observations in time with EnKF is to assimilate them at the moments coinciding with the model time steps (Sakov et al., 2010). However, as these authors mention, this approach requires the disruption of the ensemble integration, the ensemble update and a restart, which may not feasible for large-scale forecasting applications. Continuous approaches, such as 3D-Var or 4D-Var methods, are usually implemented in oceanographic modeling in order to integrate asynchronous observations at their corresponding arrival moments (Derber and Rosati, 1989; Huang et al., 2002; Macpherson, 1991; Ragnoli et al., 2012). In fact, oceanographic observations are commonly collected at not pre-determined, or asynchronous, times. For this reason, in variational data assimilation, the past asynchronous observations are simultaneously used to minimize the cost function that measures the weighted difference

between background states and observations over the time interval, and identify the best estimate of the initial state condition (Drecourt, 2004; Ide et al., 1997; Li and Navon, 2001). In addition to the 3D-Var and 4D-Var methods, Hunt et al. (2004) proposed a Four Dimensional Ensemble Kalman Filter (4DEnKF) which adapts EnKF to handle observations that have occurred at non-assimilation times. In this method the linear combinations of the ensemble trajectories are used to quantify how well a model state at the assimilation time fits the observations at the appropriate time. Furthermore, in case of linear dynamics 4DEnKF is equivalent to instantaneous assimilation of the measured data (Hunt et al., 2004). Similarly to 4DEnKF, Sakov et al. (2010) proposed the Asynchronous Ensemble Kalman Filter (AEnKF), a modification of the EnKF, mainly equivalent to 4DEnKF, used to assimilate asynchronous observations (Rakovec et al., 2015). Contrary to the EnKF, in the AEnKF current and past observations are simultaneously assimilated at a single analysis step without the use of adjoint model. Yet another approach to assimilate asynchronous observations in models is the so-called First-Guess at the Appropriate Time (FGAT) method. Like in 4D-Var, the FGAT compares the observations with the model at the observation time. However, in FGAT the innovations are assumed constant in time and remain the same within the assimilation window (Massart et al., 2010). Having reviewed all the described approaches, in this study we have decided to use a straightforward and pragmatic method, due to the linearity of the hydrological models implemented in this study, similar to the AEnKF to assimilate the asynchronous crowdsourced observations.

The main objective of this novel study is to assess the potential use of crowdsourced observations within hydrological modelling. In particular, the specific objectives of this study are to a) assess the influence of different arrival frequency of the crowdsourced observations and their related accuracy on the assimilation performances in case of a single social sensor; b) to integrate the distributed low-cost social sensors with a single physical sensor to assess the improvement in the flood prediction performances in an early warning system. The methodology is applied in the Brue (UK), Sieve (Italy), Alzette (Luxemburg) and Bacchiglione (Italy) catchments, considering lumped and semi-distributed hydrological models respectively. Due to the fact that streamflow observations from social sensors are not available in the Brue, Sieve and Alzette catchments while in the Bacchiglione catchment the sensors are being recently installed, the synthetic time series, asynchronous in time and with random accuracy, that imitate the crowdsourced observations, are generated and used.

The study is organized as follows. Firstly, the case studies and the datasets used are presented. Secondly, the hydrological models used are described. Then, the procedure used to integrate the crowdsourced observations is reported. Finally, the results, discussion and conclusions are presented.

2 Case studies and datasets

In this paper we choose four different case studies in order to validate the obtained results for areas having diverse topographical and hydrometeorological features and represented by two different hydrological models. The Brue, Sieve and Alzette catchments are considered because of the availability of precipitation and streamflow data, while the Bacchiglione river is one of the official case studies of the WeSenseIt Project (Huwald et al., 2013), which is funding this research.

2.1 Brue catchment

The first case study is located in the Brue catchment (Figure 2), in Somerset, with a drainage area of about 135 km^2 at the catchment outlet in Lovington. Using the SRTM DEM with the 90m resolution it is possible to derive the topographical characteristics, streamflow network and the consequent time of concentration, by means of the Giandotti equations (Giandotti, 1933), which is about 10 hours. The hourly precipitation (49 rainfall stations) and streamflow data used in this study are supplied by the British Atmospheric Data Centre from the HYREX (Hydrological Radar Experiment) project (Moore et al., 2000; Wood et al., 2000). The average precipitation value in the catchment is estimated using the Ordinary Kriging (Matheron, 1963).

2.2 Sieve catchment

The second case study is the Sieve catchment (see Figure 2), a tributary of the Arno River located in the Central Italian Apennines, Italy. The catchment has a drainage area of about 822km^2 with a length of 56 km and it covers mostly hills and mountainous areas with an average elevation of 470 m above sea level. The time of concentration of the Sieve catchment is about 12 hours. Hourly discharge observations are provided by the Centro Funzionale di Monitoraggio Meteo Idrologico-Idralico of the Tuscany Region at the outlet section of the catchment at Fornacina. The mean areal precipitation is calculated by Thiessen polygon method using 11 rainfall stations (Solomatine and Dulal, 2003).

2.3 Alzette catchment

The third case study is the Alzette catchment, located in the large part of the Grand-Duchy in Luxembourg. The drainage area of the catchment is about 288km^2 and the river has a length of 73 km along France and Luxembourg. The catchment covers cultivated land, grassland, forestland and urbanized land (Fenicia et al., 2007). Thiessen polygon method is used for averaging the series at the individual stations and calculate hourly rainfall series (Fenicia et al., 2007), while streamflows data are available measured at the Hesperange gauging station.

2.4 Bacchiglione catchment

The last case study is the upstream part of the Bacchiglione River basin, located in the North-East of Italy, and tributary of the River Brenta which flows into the Adriatic Sea at the South of the Venetian Lagoon and at the North of the River Po delta. The study area has an overall extent and river length of about 400 km2 and 50 km (Ferri et al., 2012). The main urban area located in the downstream part of the study area is Vicenza. The analysed part of the Bacchiglione River has four main tributaries. On the Western side the confluences with the Bacchiglione are the Leogra, the Orolo and the Retrone River, whose junction is located in the urban area itself. In Figure 3 the Retrone River it is not shown since it does not influence the water level measured at the gauged station of Vicenza (Ponte degli Angeli in Figure 3). On the Eastern side there is the Timonchio River (see Figure 3). The Alto Adriatico Water Authority (AAWA) has implemented an Early Warning System to properly forecast the possible future flood events. Recently, within the activities of the WeSenseIt Project (Huwald et al., 2013), , one physical sensor and three staff gauges complemented by a QR code (social sensor, as represented in Figure 1) were installed in the Bacchiglione River to measure the water level. In particular, the physical sensor is located at the outlet of the Leogra catchment while the three social sensors are located at the Timonchio, Leogra and Orolo catchments outlet respectively (see Figure 3).

2.5 Datasets

Three flood events for each one of the four described catchments are considered to assess the assimilation of crowdsourced observations in hydrological modelling. The observed precipitation values are treated as the “perfect forecasts” and are fed into the hydrological model. For the Brue catchment, a 2 years’ time series (June 1994 to May 1996) of observed

streamflow and precipitation data are available for model calibration and validation. On the other hand, for the Sieve catchment only 3 months of hourly runoff discharge and precipitation data (December 1959 to February 1960) are available (Solomatine and Dulal, 2003). For the Alzette catchment, two-year hourly data (July 2000 to June 2002) are used for the model calibration and validation (Fenicia et al., 2007).

In case of Bacchiglione catchment, three flood events occurred in 2013, 2014 and 2016 are considered. In particular, the one of 2013 had high intensity and resulted in several traffic disruptions at various locations upstream Vicenza. For flood forecasting, AAWA uses the 3-day weather forecast as the input to the hydrological model. The observed values of streamflow and water level at Ponte degli Angeli are used to assess the performance of the hydrological model.

3 Hydrological modelling

3.1 Lumped model

A lumped conceptual hydrological model is implemented to estimate the flood hydrograph at the outlet section of the Brue, Sieve and Alzette catchments. The choice of the model is based on previous studies performed on the Brue catchment in case of assimilation of streamflow observations from dynamic sensors (Mazzoleni et al., 2015). Direct runoff is used as input in the conceptual model and assessed by means of the Soil Conservation Service Curve Number (SCS-CN) method (Mazzoleni et al., 2015). The average value of CN within the catchment is calibrated by minimizing the difference between the simulated volume and observed quickflow, using the method proposed by Eckhardt (2005), at the outlet section.

The main module of the hydrological model is based on the Kalinin-Milyukov-Nash (KMN), Szilagyi and Szollosi-Nagy (2010), equation:

$$Q(t) = \frac{1}{k} \cdot \frac{1}{(n-1)!} \int_{t_0}^{t} \left(\frac{\tau}{k} \right)^{n-1} \cdot e^{-\tau/k}(\tau) \cdot I(t-\tau) \cdot d\tau \quad (1)$$

where I is the model forcing (in this case direct runoff), n (number of storage elements) and k (storage capacity expressed in hours) are the two parameters of the model and Q is the model output (streamflow). In this study, the parameter k is assumed as a linear function between the time of concentration, assessed using the Giandotti equation (Giandotti, 1933) and a coefficient

246 c_k. Szilagyi and Szollosi-Nagy (2010) derived the discrete state-space system of Eq. (1) that is
247 used in this study in order to apply the data assimilation (DA) approach (Mazzoleni et al., 2014,
248 2015).

249 The model calibration is performed maximizing the NSE and correlation between the simulated
250 and observed value of discharge, at the outlet point of the Brue, Sieve and Alzette catchments,
251 using historical time series. The results of such calibration provided a value of the parameters
252 n and c_k equal to 4 and 0.026, 1 and 0.0055, and 1 and 0.00064 for the Brue, Sieve and Alzette
253 catchments respectively.

254 ## 3.2 Semi-distributed model

255 The hydrological and routing models used in this study are based on the early warning system
256 implemented by the AAWA and described in Ferri et al. (2012). One the main goal of this study
257 is also to test our methodology using synthetic observations to then apply it, in the framework
258 of the WeSenseIt Project, on the existing early warning system implemented by AAWA on the
259 Bacchiglione catchment.

260 In the schematization of the Bacchiglione catchment, the location of physical and social sensors
261 corresponds to the outlet section of three main sub- catchment s, Timonchio, Leogra and Orolo,
262 while the remaining sub-catchments are considered as inter-catchment. For both sub-
263 catchments and inter-catchments, a conceptual hydrological model, described below, is used to
264 estimate the outflow hydrograph. The outflow hydrograph of the three main sub-catchments is
265 considered as upstream boundary conditions of a hydraulic model used to estimate water level
266 in the main river channel (see Figure 3), while the outflow from the inter-catchment is
267 considered as internal boundary condition to account for their corresponding drained area. In
268 the following, a brief description of the main components of the hydrological and routing
269 models is provided.

270 The input for the hydrological model consists of precipitation only. The hydrological response
271 of the catchment is estimated using a hydrological model that considers the routines for runoff
272 generation and a simple routing procedure. The processes related to runoff generation (surface,
273 sub-surface and deep flow) are modelled mathematically by applying the water balance to a
274 control volume representative of the active soil at the sub-catchment scale. The water content
275 Sw in the soil is updated at each calculation step dt using the following balance equation:

$$Sw_{t+dt} = Sw_t + P_t - R_{sur,t} - R_{sub,t} - L_t - ET_t \quad (2)$$

where P and ET are the components of precipitation and evapotranspiration, while R_{sur}, R_{sub} and L are the surface runoff, sub-surface runoff and deep percolation model states respectively (see Figure 3). The surface runoff is expressed by the equation based on specifying the critical threshold beyond which the mechanism of dunnian flow (saturation excess mechanism) prevails:

$$R_{sur,t} = \begin{cases} C \cdot \left(\frac{Sw_t}{Sw_{\max}} \right) \cdot P_t \Rightarrow P(t) \le f = \frac{Sw_{\max} \cdot (Sw_{\max} - Sw_t)}{(Sw_{\max} - C \cdot Sw_t)} \\ P_t - (Sw_{\max} - Sw_t) \Rightarrow P_t > f \end{cases}. \quad (3)$$

where C is a coefficient of soil saturation obtained by calibration, and Sw_{max} is the content of water at saturation point which depends on the nature of the soil and on its use.

The sub-surface flow is considered proportional to the difference between the water content $Sw(t)$ at time t and that at soil capacity S_c:

$$R_{sub,t} = c \cdot (Sw_t - S_c). \quad (4)$$

while the estimated deep flow is evaluated according to the expression proposed by Laio et al. (2001):

$$L_t = \frac{K_s}{e^{\beta \cdot \left(1 - \frac{S_c}{Sw_{\max}}\right)} - 1} \cdot \left(e^{\beta \cdot \left(\frac{Sw_t - S_c}{Sw_{\max}} \right)} - 1 \right). \quad (5)$$

where, K_s is the hydraulic conductivity of the soil in saturation conditions, β is a dimensionless exponent characteristic of the size and distribution of pores in the soil. The evaluation of the real evapotranspiration is performed assuming it as a function of the water content in the soil and potential evapotranspiration, calculated using the formulation of Hargreaves and Samani (1982).

Knowing the values of R_{sur}, R_{sub} and L, it is possible to model the surface Q_{sur}, sub-surface Q_{sub} and deep flow Q_g routed contributes according to the conceptual framework of the linear reservoir at the closing section of the single sub-catchment. In particular, in case of Q_{sur} the value of the parameter k, which is a function of the residence time in the catchment slopes, is estimated relating the slopes velocity of the surface runoff to the average slopes length L. However, one of difficulties involved is the proper estimation of the surface velocity, which

should be calculated for each flood event (Rinaldo and Rodriguez-Iturbe, 1996). According to Rodríguez-Iturbe et al. (1982), such velocity is a function of the effective rainfall intensity and event duration. In this study, the estimate of the surface velocity is performed using the relation between velocity and intensity of rainfall excess proposed in Kumar et al. (2002). In this way it is possible to estimate the average time travel and the consequent parameter k. However, such formulation is applied in a lumped way for a given sub-catchment. As reported in McDonnell and Beven (2014) more reliable and distributed models should be used to reproduce the spatial variability of the residence times within the catchment over the time. That is why, in the advanced version of the model implemented by AAWA, in each sub-catchment the runoff propagation is carried out according to the geomorphological theory of the hydrologic response. In such model, the overall catchment travel time distributions is considered as nested convolutions of statistically independent travel time distributions along sequentially connected, and objectively identified, smaller sub-catchments. The parameter k assumes different values for each time step as the rainfall changes. In fact, the variability of residence time is considered according to Rodríguez-Iturbe et al. (1982) by assuming the surface velocity as a function of the effective rainfall intensity (Kumar et al., 2002). Anyway, the correct estimation of the residence time should be derived considering the latest findings reported in McDonnell and Beven (2014). In case of Q_{sub} and Q_g the value of k is calibrated comparing the observed and simulated discharge at Vicenza as previously described.

In the early warning system implemented by AAWA in the Bacchiglione catchment, the flood propagation along the main river channel is represented one-dimensional hydrodynamic model, MIKE 11 (DHI, 2005). This model solves the Saint Venant Equations in case of unsteady flow based on an implicit finite difference scheme proposed by Abbott and Ionescu (1967). However, in order to reduce the computational time required by the analysis performed in this study MIKE11 is replaced by a hydrological routing Muskingum-Cunge model (see, e.g. Todini 2007), considering river cross-sections as rectangular for the estimation of hydraulic radios, wave celerity and the other hydraulic variables.

Calibration of the hydrological and hydrodynamic model parameters is performed by AAWA, and described in Ferri et al. (2012), considering the time series of precipitation from 2000 to 2010 in order to minimize the root mean square error between observed and simulated values of water level at Ponte degli Angeli gauged station. In order to stay as close as possible to the

early warning system implemented by AAWA, we used the same calibrated model parameters proposed by Ferri et al. (2012).

4 Data assimilation procedure

4.1 Kalman Filter

In Data Assimilation (DA) it is typically assumed that the dynamic system can be represented in the state-space as follows:

$$\mathbf{x}_t = M(\mathbf{x}_{t-1}, \vartheta, I_t) + w_t \qquad w_t \sim N(0, \mathbf{S}_t). \tag{6}$$

$$\mathbf{z}_t = H(\mathbf{x}_t, \vartheta) + v_t \qquad v_t \sim N(0, R_t). \tag{7}$$

where, $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ are state vectors at time t and t-1, M is the model operator that propagates the states $\mathbf{x}$ from its previous condition to the new one as a response to the inputs I_t, while H is the operator which maps the model states into output $\mathbf{z}_t$. The system and measurements errors w_t and v_t are assumed to be normally distributed with zero mean and covariance $\mathbf{S}$ and R. In a hydrological modelling system, these states can represent the water stored in the soil (soil moisture, groundwater) or on the earth surface (snow pack). These states are one of the governing factors that determine the hydrograph response to the inputs into the catchment.

In case of the linear systems used in this study, the discrete state-space system of Eq. (1) can be represented as follows (Szilagyi and Szollosi-Nagy, 2010):

$$\mathbf{x}_t = \mathbf{\Phi}\mathbf{x}_{t-1} + \mathbf{\Gamma} I_t + w_t. \tag{8}$$

$$Q_t = \mathbf{H}\mathbf{x}_t + v_t. \tag{9}$$

where t is the time step, $\mathbf{x}$ is vector of the model states (stored water volume in m^3), $\mathbf{\Phi}$ is the state-transition matrix (function of the model parameters n and k), $\mathbf{\Gamma}$ is the input-transition matrix, $\mathbf{H}$ is the output matrix, and I and Q are the input (forcing) and model output (discharge in this case). For example, for n=3 the matrix $\mathbf{H}$ is expressed as $\mathbf{H} = \begin{bmatrix} 0 & 0 & k \end{bmatrix}$. Expressions for matrices $\mathbf{\Phi}$ and $\mathbf{\Gamma}$ can be found in Szilagyi and Szollosi-Nagy (2010).

For the Bacchiglione model, the preliminary sensitivity analysis on the model states (soil content S and the storage water x_{sur}, x_{sub} and x_L related to Q_{sur}, Q_{sub} and Q_g) is performed in order to decide on which of the states to update. The results of this analysis (shown in the next section) pointed out that the stored water volume x_{sur} (estimated using Eq. (8) with n=1, H=k

and I_t replaced by R_{sur}) is the most sensitive state and for this reason we decided to update only this state.

The Kalman Filter (KF, Kalman, 1960) is a mathematical tool which allows estimating, in an efficient computational (recursive) way, the state of a process which is governed by a linear stochastic difference equation. KF is optimal under the assumption that the error in the process is Gaussian; in this case KF is derived by minimizing the variance of the system error (error in state) assuming that the model state estimate is unbiased. In an attempt to overcome these limitations, various variants of the Kalman filter, such as the extended Kalman filter (EKF), unscented Kalman filter and ensemble Kalman filter (EnKF) have been proposed.

Kalman filter procedure can be divided in two steps, namely forecast equations, (Eqs. (10) and (11)), and update (or analysis) equations (Eqs. (12), (13) and (14)):

$$\mathbf{x}_t^- = \mathbf{\Phi}\mathbf{x}_{t-1}^+ + \mathbf{\Gamma}\mathbf{I}_t . \tag{10}$$

$$\mathbf{P}_t^- = \mathbf{\Phi}\mathbf{P}_{t-1}^+\mathbf{\Phi}^T + \mathbf{S}. \tag{11}$$

$$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_t^-\mathbf{H}^T + R\right)^{-1}. \tag{12}$$

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + \mathbf{K}_t\left(Q_t^o - \mathbf{H}\mathbf{x}_t^-\right). \tag{13}$$

$$\mathbf{P}_t^+ = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}\right)\mathbf{P}_t^- . \tag{14}$$

where $\mathbf{K}_t$ is the Kalman gain matrix, $\mathbf{P}$ is the error covariance matrix and Q^0 is the new observation. The prior model states $\mathbf{x}$ at time t are updated, as the response to the new available observations, using the analysis equations Eqs. (12) to (14). This allows for estimation of the updated states values (with superscript +) and then assessing the background estimates (with superscript –) for the next time step using the time update equations Eqs. (10) and (11). The proper characterization of the model covariance matrix $\mathbf{S}$ is a fundamental issue in Kalman filter. In this study, in order to evaluate the effect of assimilating crowdsourced observations, the model is considered more accurate than the observations. For this reason, a different value of the covariance matrix $\mathbf{S}$ is considered for each case study. In fact, a covariance matrix $\mathbf{S}$ with diagonal values of $1 m^6/s^2$, $25 m^6/s^2$ and $1 m^6/s^2$ are considered for the Brue, Sieve and Alzette catchments. The bigger value of $\mathbf{S}$ in the Sieve catchment is due to the higher flow magnitude in such catchment if compared to the other two. A sensitivity analysis of model performances depending on the value of $\mathbf{S}$ is reported in the Results section. In case of the Bacchiglione

catchment, **S** is estimated, for each given flood event, as the variance between observed and simulated flow values.

4.2 Assimilation of asynchronous streamflow observations with irregular accuracy

In most of the hydrological applications of DA, observations from physical sensors are integrated into water models at a regular, synchronous, time step. However, a social sensor can be used by different operators, having different accuracy, to measure water level at a specific point. For this reason, social sensors provide crowdsourced observations which are asynchronous in time and with a higher degree of uncertainty than the one of observations from physical sensors. In particular, crowdsourced observations have three main characteristics: a) irregular arrival frequency (asynchronicity); b) random accuracy; c) random number of observations received by the static device within two model time steps.

As described in the Introduction, various methods have been proposed in order to include asynchronous observations in models. Having reviewed them, in this study we are proposing a somewhat simpler DA approach for integrating Crowdsourced Observations into hydrological models (DACO). This method is based on the assumption that the change in the model states and in the error covariance matrices within the two consecutive model time steps t_0 and t (observation window) is linear, while the inputs are assumed constant. All the data received during the observation window are assimilated in order to update the model states and output at time t. Therefore, assuming that one observation would be available at time t_0^*, the first step of such a filter (A in Figure 4) is the definition of the model states and error covariance matrix at t_0^* as:

$$\mathbf{x}_{t_0^*}^- = \mathbf{x}_{t_0}^+ + \left(\mathbf{x}_t^- - \mathbf{x}_{t_0}^+\right).\frac{t_0^* - t_0}{t - t_0}. \quad (15)$$

$$\mathbf{P}_{t_0^*}^- = \mathbf{P}_{t_0}^+ + \left(\mathbf{P}_t^- - \mathbf{P}_{t_0}^+\right).\frac{t_0^* - t_0}{t - t_0} \quad (16)$$

The second step (B in Figure 4) is the estimation of the updated model states and error covariance matrix, as the response to the streamflow observation $Q_{t_0^*}^o$. The estimation of the posterior values of $\mathbf{x}_{t_0^*}^-$ and $\mathbf{P}_{t_0^*}^-$ is performed by Eqs. (13) and (14) respectively. The Kalman gain is estimated by Eq. (12), where the prior values of model states and error covariance matrix

417 at t_0^* are used. Knowing the posterior value $\mathbf{x}^+_{t_0^*}$ and $\mathbf{P}^+_{t_0^*}$ it is possible to predict the value of
418 states and covariance matrix at one model step ahead, t^* (C in Figure 4) using the model forecast
419 equations Eqs. (10) and (11).

420 The last step (D in Figure 4) is the estimation of the interpolated value of **x** and **P** at time step
421 t. This is performed by means of a linear interpolation between the current values of **x** and **P** at
422 t_0^* and t^*:

423 $$\tilde{\mathbf{x}}^-_t = \mathbf{x}^-_{t_0^*} + \left(\mathbf{x}^-_{t^*} - \mathbf{x}^+_{t_0^*}\right) \cdot \frac{t - t_0^*}{t^* - t_0^*} . \quad (17)$$

424 $$\tilde{\mathbf{P}}^-_t = \mathbf{P}^-_{t_0^*} + \left(\mathbf{P}^-_{t^*} - \mathbf{P}^+_{t_0^*}\right) \cdot \frac{t - t_0^*}{t^* - t_0^*} . \quad (18)$$

425 The symbol ~ is added on the new matrices **x** and **P** in order to differentiate them from the
426 original forecasted values in t. Assuming that a new streamflow observation is available at an
427 intermediate time t_1^* (between t_0^* and t), the procedure is repeated considering the values at t_0^*
428 and t for the linear interpolation. Then, in case when no more observations are available, the
429 updated value of $\tilde{\mathbf{x}}^-_t$ is used to predict the model states and output at t+1 (Eqs. (10) and (11)).
430 Finally, in order to account for the intermittent behaviour of such observations, the approach
431 proposed by Mazzoleni et al. (2015) is applied. In this method, the model states matrix **x** is
432 updated and forecasted when observations are available, while without observations the model
433 is run using Eq. (10) and covariance matrix **P** propagated at the next time step using Eq. (11)

434 ## 4.3 Observation accuracy

435 In this section, the uncertainty related to the streamflow crowdsourced observations is
436 characterised. The observational error is assumed to be the normally distributed noise with zero
437 mean and given standard deviation:

438 $$\sigma^Q_t = \alpha_t \cdot Q^{true}_t \quad (19)$$

439 where the coefficient α is related to the degree of uncertainty of the measurement (Weerts and
440 El Serafy, 2006).

441 One of the main and obvious issues in citizen-based observations is to maintain the quality
442 control of the water observations (Cortes et al., 2014; Engel and Voshell, 2002). In Introduction

a number of methods to estimate (calibrate) the model of observational uncertainty have been referred to. In this study coefficient α is assumed a random variable uniformly distributed between 0.1 and 0.3, so we leave more thorough investigation of uncertainty level of the crowdsourced data for future studies. Cortes et al. (2014) argue (and this is a reasonable suggestion) that the uncertainty of a measurement provided by a well-trained technician is smaller than the one coming from a normal citizen. For this reason we assumed that the maximum value of α is three times higher than the uncertainty coming from the physical sensors. The value of Q^{true} is the streamflow value measured at a asynchronous time step and it is described in the next section.

5 Experimental setup

In this section, two sets of experiments are performed in order to test the proposed method and assess the benefit to integrate crowdsourced observations, asynchronous in time and with variable accuracy, in real-time flood forecasting.

In the first set of experiments, called "Experiments 1", assimilation of streamflow observations at one social sensor location is carried out to understand the sensitivity of the employed hydrological model (KMN) under various scenarios of such observations.

In the second set of experiments, called "Experiments 2", the distributed observations coming from social and physical sensors, at four locations within the Bacchiglione catchment, are considered, with the aim of assessing the improvement in the flood forecasting accuracy. The social sensors, showed in Figure 1, were installed in the summer of 2014 within the framework of the WeSenseIt project.

5.1 Experiments 1: Assimilation of crowdsourced observations from one social sensor

The focus of Experiments 1 is to study the performance of the hydrological model (KMN) assimilating crowdsourced observations, having lower arrival frequencies than the model time step and random accuracies, coming from a social sensor located at the outlet point of the Brue, Sieve and Alzette catchments.

Due to the fact that crowdsourced observations are not available in these case studies at the moment of this study, realistic synthetic streamflow observations having different

characteristics are generated. For this reason, observed hourly streamflow observations at the catchments outlet are interpolated to represent observations coming at arrival frequency higher than hourly. A similar approach, termed "observing system simulation experiment" (OSSE), is commonly used in meteorology to estimate synthetic "true" states and measurements by introducing random errors in the state and measurement equations (Arnold and Dey, 1986; Errico et al., 2013; Errico and Privé, 2014). OSSEs have the advantage of making it possible to directly compare estimates to "true" states and they are often used for validating DA algorithms.

To analyse all possible combinations of arrival frequency, number of observations within the observation window (1 hour) and accuracy, a set of scenarios are considered (Figure 5), changing from regular arrival frequency of observations with high accuracy (scenario 1) to random and chaotic asynchronous observations with variable accuracy (scenario 11). In each scenario a varying the number of observations from 1 to 100 is considered. It is worth noting that in case of one observation per hour and regular arrival time, scenario 1 corresponds to the case of physical sensors with an observation arrival frequency of one hour.

Scenario 2 corresponds to the case of observations having fixed accuracy (α equal to 0.1) and irregular arrival moments, but in which at least one observation coincides with the model time step. In particular, scenario 1 and 2 are exactly the same in case of one observation available within the observation window since it is assumed that the arrival frequency of that observation has to coincide with the model time step. On the other hand, the arrival frequency of the observations in scenario 3 is assumed to be random and observations might not arrive at the model time step.

Scenario 4 considers observations with regular frequency but random accuracy at different moments within the observation window, whereas in scenario 5 observations have irregular arrival frequency and random accuracy. In all the previous scenarios the arrival frequency, the number and accuracy of the observations are assumed to be periodic, i.e. repeated between consecutive observation windows along all the time series. However such periodic repetitiveness might not occur in real-life, and for this reason, a non-periodic behaviour is assumed in scenarios 6, 7, 8 and 9. The non-periodicity assumptions of the arrival frequency and accuracy are the only factors that differentiate scenarios 6, 7, 8 and 9 from the scenarios 2, 3, 4, and 5 respectively. In addition, the non-periodicity of the number of observations within the observation window is introduced in scenario 10.

Finally, in scenario 11 the observations, in addition to all the previous characteristics, might have an intermittent behaviour, i.e. not being available for one or more observation windows.

5.2 Experiments 2: Spatially distributed physical and social sensors

Synthetic hourly streamflow observations are calculated using measured precipitation recorded during the 2013, 2014 and 2016 flood events (post-event simulation) as input in the hydrological model of the Bacchiglione catchment. Interpolated streamflow observations having characteristics reported in scenarios 10 and 11, in Experiments 1, are generated due to the unavailability of crowdsourced observations at the moment of this study. In order to evaluate the model performances, observed and simulated streamflows are compared, for different lead times.

Streamflow observations from physical sensors are assimilated in the hydrological model of AMICO system at an hourly frequency, while crowdsourced observations from social sensors are assimilated using the DACO method previously described. The updated hydrograph estimated by the hydrological model is used as the input into Muskingum-Cunge model used to propagate the flow downstream, to the gauged station at Ponte degli Angeli, Vicenza.

The main goal of Experiments 2 is to understand the contribution of distributed crowdsourced observations to the improvement of the flood prediction at a specific point of the catchment, in this case at Ponte degli Angeli. For this reason, five different experimental settings are introduced, and represented in Figure 6, corresponding to different types of employed sensors.

Firstly, only the observations coming from the physical sensor at the Leogra sub-catchment are used to update the hydrological model of sub-catchment B (setting A). Secondly, in setting B, the model improvement in case of assimilation of crowdsourced observations at the same location of setting A is analysed. In setting C only the distributed crowdsourced observations within the catchment are assimilated into the hydrological model. Then, setting D accounts for the integration of crowdsourced and physical observations, contrary to the setting C where the physical sensors is dropped in favour of the social sensor at Leogra. Finally, setting E considers the complete integration between physical and social sensors in Leogra, Timonchio and Orolo sub-catchments.

6 Results

6.1 Experiments 1: Influence of crowdsourced observations on flood forecasting

The observed and simulated hydrographs at the outlet section of the Brue, Sieve and Alzette catchments with and without the model update (considering hourly streamflow observations) are reported in Figure 7 for nine different flood events in case of 1 hour lead time. As expected, it can be seen that the updated model tends to better represent the flood events than model without updating in all the case studies. However, such improvement it is closely related to the value of the matrix **S**. In fact, increasing the value of **S**, i.e. assuming a less accurate model, force the model towards the observations because more accurate than the model itself. For this reason, a sensitivity analysis on the influence of the matrix **S** on the assimilation of crowdsourced observations in case of scenario 1, i.e. coming and assimilated at regular time steps within the observation windows, is reported in Figure 8. The results of Figure 8 are related to the first flood event of the Brue, Sieve and Alzette catchments. Increasing the number of observations within the observation window results in an improvement of the NSE for different value of model error. However, such improvement becomes negligible for a given threshold value of streamflow observation, which is a function of the considered flood event. This means that the additional observations do not add information useful for improving the model performance. Overall, increasing the value of the model error **S** tends to increase NSE values as mentioned before. For this reason, in order to better evaluate the effect of assimilating crowdsourced observations, a small value of **S**, i.e. model more accurate than observations, is assumed.

In case of scenario 1 the arrival frequency is set as regular for different model runs, so the moments and accuracy in which the observations became available is always the same for any model run. However, for the other scenarios, the irregular moment in which the observations becomes available within the observation window and their accuracy are randomly selected and change according to the different model runs. This reflects in a random model performances, and consequent NSE. In order to remove such random behaviour, different model runs (100 in this case) are carried out, assuming different random values of arrival and accuracy (coefficient α) during each model run, for a given number of observations and lead time. The NSE value is estimated for each model run, so μ(NSE) and σ(NSE) represent the mean and standard deviation of the different values of NSE.

In case of scenarios 2 and 3 (represented using warm, red and orange, colours in Figure 9 and Figure 10 for lead time equal to 24 hours), i.e. random arrival frequency with fixed/controlled accuracy, the average values of NSE, μ(NSE), are smaller but comparable with the ones obtained in case of scenario 1 for all the considered flood events and case studies. In particular, scenario 3 has lower μ(NSE) than scenario 2. This can related to the fact that both scenarios has random arrival frequency, however, in scenario 3 observations are not provided at the model time step, as opposed to scenario 2. From Figure 10, higher values of σ(NSE), can be observed in case of scenario 3. Scenario 2 has the lowest standard deviation for low values of discharge observations due to the fact that the arrival frequency has to coincide with the model time step and this tends to stabilize the NSE. In particular, for increasing number of observations σ(NSE) tends to decrease. However, constant trend of σ(NSE) can be observed, due to particular characteristics of the flood events, in case of the flood event 1 of Sieve and flood event 2 and 3 of Alzette. It is worth nothing that scenario 1 has null standard deviation due to the fact that observations are assumed coming at the same moment with the same accuracy for the all 100 model runs.

In scenario 4, represented using cold blue colour, observations are considered coming at regular time steps but having random accuracy. Figure 9 shows that μ(NSE) values are lower in case of scenario 4 than scenarios 2 and 3. This can be related to the higher influence of observations accuracy if compared to arrival frequency. High variability in the model performances of model, especially for low values of crowdsourced observations, it can be observed in scenario 4 (Figure 10).

The combined effects of random arrival frequency and observation accuracy is represented in scenario 5 using a magenta colour (i.e. the combination of warm and cold colours used for scenarios 2, 3 and 4) in Figure 9 and Figure 10. As expected, this scenario is the one with the lower μ(NSE) and higher σ(NSE) values if compared to the previous ones.

The remaining scenarios, from 6 to 9, are equivalent to the ones from 2 to 5 with the only difference that are non-periodic in time. For this reason, in Figure 9 and Figure 10, scenarios from 6 to 9 have the same colour of scenarios 2 to 5 but indicated with dashed line in order to underline their non-periodic behaviour. Overall it can be observed that non-periodic scenarios have similar μ(NSE) values to their corresponding periodic scenario. However, their smoother μ(NSE) trends are because of lower σ(NSE) values which means that model performances are less dependent to the non-periodic nature of the crowdsourced observations than their period

595 behaviour. Table 1 shows the NSE values and model improvement obtained for the different
596 experimental scenarios during the different flood events. Overall, small improvements are
597 obtained when NSE is already high for 1 crowdsourced observation as in case of the Sieve
598 catchment during flood event 2 or the Alzette catchment in the event 2. Moreover, it can be
599 seen that lower improvement is achieved in case of scenarios where arrival frequency is random
600 and accuracy fixed if compared to those scenarios (4, 5, 8 and 9) where arrival frequency is
601 regular and accuracy is random.

602 In the previous analysis, model improvements are expressed only in terms of NSE. However,
603 statistics such as NSE only explain the overall model accuracy and not the real
604 increases/decreases in prediction error. Therefore, increases in model accuracy due to the
605 assimilation of crowdsourced observations have to be presented in different ways as increased
606 accuracy of flood peak magnitudes and timing. For this reason, additional analyses are carried
607 out to assess the change in flood peak prediction considering 3 peaks occurred during flood
608 event 2 in Brue catchment (see Figure 7). Errors in the flood peak timing, Err_t, and intensity,
609 Err_I, are estimated as:

610 $$Err_t = t_P^o - t_P^S . \quad (6)$$

611 $$Err_I = \frac{Q_P^o - Q_P^S}{Q_P^o} . \quad (7)$$

612 where t_p^o and t_p^s are the observed and simulated peak time (hours), while Q_p^o and Q_p^s are the
613 observed and simulated peak discharge (m^3/s). From the results reported in Figure 11,
614 considering 12-hours lead time, it can be observed that, overall, errors reduction in peak
615 prediction is achieved for increasing number of crowdsourced observations. In particular,
616 assimilation of crowdsourced observations has more influence in the reduction of the peak
617 intensity rather than peak timing. In fact, a small reduction of Err_t of about 1 hour is obtained
618 even increasing the number of observations. In both Err_I and Err_t the higher error reduction is
619 obtained considering fixed observation accuracy and random arrival frequency (e.g. scenarios
620 1, 2, 3, 6 and 7). In fact, smaller Err_I error values are obtained in case of scenario 1, while
621 scenarios 5 and 9 are the ones that show the lowest improvement in terms of peak prediction.
622 These conclusions are very similar to the previous ones obtained analysing only NSE as model
623 performance measures.

The combination of all the previous scenarios is represented by scenario 10, where number of crowdsourced observations changing at in each observation windows are considered. In scenario 11 the intermittent nature of crowdsourced observations is accounted as well. The μ(NSE) and σ(NSE) values of these scenarios obtained for the considered flood events are showed in Figure 12. It can be observed that scenarios 10 tends to provide higher μ(NSE) and lower σ(NSE) values, for a given flood event, if compared to scenarios 11. In fact, intermittency in crowdsourced observations tends to reduce model performance and increase variability of NSE values for random configuration of arrival frequency and observations accuracy. In particular, σ(NSE) tends to be constant for increasing number of observations.

6.2 Experiments 2: Influence of distributed physical and social sensors

Three different flood events occurred in the Bacchiglione catchment are used within Experiments 2. Figure 13 shows the observed and simulated streamflow value at the outlet section of Vicenza. In particular, two simulated time series of streamflow are calculated using as input for the hydrological model the measured and forecasted time series of precipitation (provided by AAWA). Overall, an underestimation of the observed discharge can be observed using forecasted input while the results achieved used measured precipitation tend to well represent the observations. In order to find out what model states leads to a maximum increase of the model performance, a preliminary sensitivity analysis is performed. The four model states, x_S, x_{sur}, x_{sub} and x_L, related to Sw, Q_{sur}, Q_{sub} and Q_g, are perturbed by ±20% around the true state value using the uniform distribution, every time step from the initial time step up to the perturbation time (PT). No correlation between time steps is considered. After PT, the model realizations are run without perturbation in order to assess the perturbation effect on the system memory. No assimilation, and consequent model update, is performed at this step. From the results reported in Figure 14, related to the flood event 1, it can be observed that the model state x_{sur} is the most sensitive states if compared to the other ones. In addition, the perturbations of all the states seem to affect the model output even after the PT (high system memory). For this reason, in this experiments, only the model state x_{sur} is updated by means of the DACO method.

Scenarios 10 and 11, described in the previous sections, are used to represent the irregular and random behaviour of the crowdsourced observations assimilated in the Bacchiglione catchment.

Figure 15 and Figure 16 show the results obtained from the experiments settings represented in Figure 6 in case of physical and crowdsourced observations during three different flood events.

Three different lead time values are considered. Different model runs (100) are performed to account for the effect induced by the random arrival frequency and accuracy of the crowdsourced observations within the observation window as described above. From Figure 15, in which observations have the same characteristics of previous scenario 10, it can be seen that the assimilation of observations from the physical sensor in the Leogra sub-catchment (Setting A) provides a better flood prediction at Ponte degli Angeli if compared to the assimilation of a small number of crowdsourced observations provided by a social sensor in the same location (Setting B). In particular, Figure 15 show that, depending on the flood event, the same NSE values achieved with assimilation of physical observations (hourly frequency and high accuracy) can be obtained by assimilating between 10 and 20 crowdsourced observations per hour in case of 4 hours lead time. Such number of crowdsourced observations tends to increase for increasing values of lead times. In case of intermittent observations (Figure 16) the overall reduction of NSE is such that even with a high number of observations (even higher than 50 per hour) the NSE is always lower than the one obtained assimilating physical observations for any lead time.

In case of Setting C, it can be observed for all three flood events that distributed social sensors in Timonchio, Leogra and Orolo sub-catchments allow for obtaining higher model performances than the one achieved with only one physical sensor (see Figure 15). However, in case of flood event 3 this is valid only for low lead time values. In fact, for 8 and 12 hours lead time values, the contribution of crowdsourced observations tend to decrease in favour of physical observations from the Leogra sub-catchment. This effect is predominant in case of intermittent crowdsource observations, scenario 11. In this case, Setting C has higher μ(NSE) values than Setting A only during flood event 1 and for lead time values equal to 4 and 8 hours (see Figure 16).

It is interesting to note that in case of Setting D, during flood event 1, the μ(NSE) is higher than Setting C for low number of observations. However, with higher number of observations, Setting C is the one providing the best model improvement for low lead time values. In case of intermittent observations, it can be noticed that the Setting D provides always higher improvement than Setting C. For flood event 1, the best model improvement is achieved in case of Setting E, i.e. fully integrating physical sensor with distributed social sensors. On the other hand, during flood events 2 and 3 Setting D shows higher improvements than Setting E. In case of intermittent observations the difference between Setting D and E tends to reduce for all the

flood events. Overall, settings D and E are the ones which provided the highest μ(NSE) in both scenario 10 and 11. This demonstrates the importance of integrating existing network of physical sensors (Setting A) with social sensors providing crowdsourced observations in order to improve flood predictions.

Figure 17 shows the standard deviation of the NSE obtained for the different settings in case of 4 hours lead time. Similar results are obtained for the 3 considered flood events. In case of Setting A, σ(NSE) is equal to zero since observations are coming from physical sensor at regular time steps. Higher σ(NSE) values are obtained in case of Setting B, while including different crowdsourced observations (Setting C) tend to decrease the value of σ(NSE). It can be observed that σ(NSE) decreases for high values of crowdsourced observations. As expected, the lowest values of σ(NSE) are achieved including the physical sensor in the DA procedure (Setting D and E). Similar considerations can be drawn in case of intermittent observations, where higher and more perturbed σ(NSE) values are obtained.

7 Discussion

Assimilation of crowdsourced observations is performed in four different case studies considering only one social sensor location in the Brue, Sieve and Alzette catchments, and distributed social and physical sensors in the Bacchiglione catchment.

In the first three catchments, different characteristics of the crowdsourced observations are represented by means of 11 scenarios. Nine different flood events are used to assess the beneficial use in assimilating crowdsourced observations in hydrological model to improve flood forecasting.

Overall, assimilation of crowdsourced observations improves model performances in all the considered case studies. In particular, there is a limit in the number of crowdsourced observations for which satisfactory model improvements can be achieved and for which additional observations become redundant. This asymptotic behaviour when extra information is added has also been observed using other metrics by Krstanovic and Singh (1992), Ridolfi et al. (2014), Alfonso et al. (2013)), among others. From Figure 9 it can be seen that, in all the considered catchments, increasing the number of model error induces an increase of such asymptotic value with a consequent reduction of observations needed to improve model performances. For this reason, a small value of model error is assumed in this study. In addition, it is not possible to define a priori number of observations needed to improve model due to the

718 different model behaviour for a given flood event in case of no update. In fact, as reported in
719 Table 1 and Figure 8, flood events with high NSE values even without update tends to achieve
720 the asymptotic values of NSE for small number of observations (e.g. flood event 1 in Brue and
721 2 in Sieve), while more observations are needed for flood events having low NSE without
722 update. However, in case of these case studies and during these nine flood events, an indicative
723 value of 10 observations can be considered to achieve a good model improvement.

724 Figure 9 and Figure 10 show the μ(NSE) and σ(NSE) values for the scenarios 2 to 9. Figure 9
725 demonstrate that in case of irregular arrival frequency and constant accuracy (e.g. scenarios 2,
726 3, 6 and 7) the NSE is higher than in case of scenarios in which accuracy is variable and arrival
727 frequency fixed (e.g. scenarios 4, 5, 8 and 9). These results point out that the model performance
728 is more sensitive to the accuracy of the observations than to the moment in time at which the
729 streamflow observations become available. Overall, σ(NSE) tends to decrease for the high
730 number of observations. The combined effects of irregular frequency and uncertainty are
731 reflected in scenario 5 which has the lower mean and higher standard deviation of NSE if
732 compared to the first four scenarios. However, it can be observed from scenarios 2 to 5 that the
733 trend it is not as smooth as the one obtained with scenario 1. This can be related to the fact that
734 NSE may vary with varying arrival frequency and observations accuracy.

735 An interesting fact is that passing from periodic to non-periodic scenarios the standard deviation
736 σ(NSE) is significantly reduced, while μ(NSE) remains the same but with a smoother trend. A
737 non-periodic behaviour of the observations, common in real life, helps to reduce the fluctuation
738 of the NSE generated by the random behaviour of streamflow observations. Finally, the results
739 obtained for scenarios 10 and 11 are showed in Figure 12. The assimilation of irregular number
740 of observations in scenario 10, in each observation window, seems to provide the same μ(NSE)
741 than the ones obtained with scenario 9. One of the main outcome is that the intermittent nature
742 of the observations (scenario 11) induces a drastic reduction of the NSE and an increase in its
743 noise in both considered flood events. All these previous results are consistent across the
744 considered catchments.

745 In case of the Bacchiglione catchment, the physical and crowdsourced observations are
746 assimilated within a hydrological model to improve the poor flow prediction in Vicenza for the
747 three considered flood events. In fact, such predictions are affected by an underestimation of
748 the 3-days rainfall forecast used as input in flood forecasting practice in this area.

One of the main outcomes of these analyses is that the replacement of a physical sensor (Setting A) for a social sensor at only one location (settings B) does not improve the model performance in terms of NSE for small number of crowdsourced observations. Figure 15 and Figure 16 show that distributed locations of social sensors (setting C) can provide higher values of NSE than a single physical sensor, even for low number of observations, in case of regular observations (scenario 10). In case of flood event 1, Setting C provides better model improvement than Setting D for low lead time values and high number of observations. This can be due to the fact that the physical sensor at Leogra provides constant improvement, for a given lead time, while the social sensor tends to achieve better results with a higher number of observations. This dominant effect of the social sensor, in case of high number of observations, tends to increase for the higher lead times. On the other hand, for intermittent observations (scenario 11) such effect decreases in particular for flood events 2 and 3.

Integrating physical and social sensors (Setting D and E) induces the highest model improvements for all the three flood events. In case of flood event 1, assimilation from Setting E it appears to provide better results than assimilation from Setting D. Opposite results are obtained in case of flood events 2 and 3. In fact, the high μ(NSE) values of setting D, when compared to setting E during such flood events, can be due to the fact that flood events 2 and 3 are characterized by one main peak and similar shape while flood event 1 has two main peaks. From Figure 17 it can be seen that assimilation of crowdsourced observations from distributed social sensors tend to reduce the variability of the NSE coefficient in both scenarios 10 and 11.

8 Conclusions

This innovative study demonstrates that crowdsourced observations, asynchronous in time and with variable accuracy, can improve flood prediction if integrated in hydrological models. Such observations are assumed to be inferred using low-cost social sensors as, for example, staff gauge connected to a QR code on which people can read the water level indication and send the observations via a mobile phone application. This type of social sensor is tested within the framework of the WeSenseIt FP7 Project. Four different case studies, the Brue (UK), Sieve (Italy), Alzette (Luxemburg) and Bacchiglione (Italy) catchments, are considered, and the two types of hydrological models are used. In the Experiments 1 (Brue, Sieve and Alzette catchments) the sensitivity of the model results to the different frequencies and accuracies of the crowdsourced observations coming from a hypothetical social sensor at the catchments outlet is assessed. On the other hand, in the Experiments 2 (Bacchiglione catchment), the

influence of the combined assimilation of crowdsourced observations, coming from a distributed network of social sensors, and existing streamflow observations from physical sensors, used in the Early Warning System implemented by AAWA, is evaluated. Due to the fact that crowdsourced streamflow observations are not yet available in all case studies, realistic synthetic observations with various characteristics of arrival frequency and accuracy are introduced.

Overall, we demonstrated that the results we have obtained are very similar in terms of model behaviour assimilating asynchronous observations in all cases studies.

In Experiments 1 it is found that increasing the number of crowdsourced observations within the observation window increases the model performance even if these observations have irregular arrival frequency and accuracy. Therefore, observations accuracy affects the average value of NSE more than the moment in which these observations are assimilated. The noise in the NSE is reduced when the assimilated observations are considered having non-periodic behaviour. In addition, the intermittent nature of the observations tends to drastically reduce the NSE of the model for different values of lead times. In fact, if the intervals between the observations are too large then the abundance of crowdsourced data at other times and places is no longer able to compensate their intermittency.

Experiments 2 showed that, in the Bacchiglione catchment, the integration of observations from social sensors and single physical sensor can improve the flood prediction even in case of a small number of intermittent crowdsourced observations. In case of both physical and social sensors located at the same place the assimilation of crowdsourced observations give the same model improvement than the assimilation of physical observations only in case of high number and non-intermittent behaviour. Overall, the integration of existing physical sensors with a new network of social sensors can improve the model predictions, as shown in the Bacchiglione case study. We agree that the cases and models are indeed different, but the presented study demonstrated that the results obtained are very similar in terms of model behaviour assimilating asynchronous observations.

In our study we have obtained interesting results, however, this work has still certain limitations. Firstly, the proposed method used to assimilate crowdsourced observations is applied to the linear parts of hydrological models, so the proposed methodology has to be tested on models with explicit non-linearities. Secondly, while quite realistic synthetic streamflow observations have been used in this study, the developed methodology was not tested on real-

life data (observations coming from actual social sensors) so there is a need to check if the drawn conclusions are still valid. Finally, advancing methods for a more accurate assessment of the data quality and accuracy of streamflow observations coming from social sensors need to be considered (e.g. developing a pre-filtering module aimed to select only observations having good accuracy while discarding the one with low accuracy).

The future work will be aimed at addressing the limitations formulated above, which would hopefully allow for a better characterisation of the crowdsourced observations (citizens observatories) and making them a more reliable source of data for model-based forecasting.

Acknowledgements

This research was partly funded in the framework of the EC FP7 Project WeSenseIt: Citizen Observatory of Water, grant agreement No. 308429. 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 (Italy). The Authors wish to thank the two anonymous reviewers the Editor for their insightful and useful comments.

References

Abbott, M. B. and Ionescu, F.: On The Numerical Computation Of Nearly Horizontal Flows, J. Hydraul. Res., 5(2), 97–117, doi:10.1080/00221686709500195, 1967.

ABC: ABC's crowdsourced flood-mapping initiative, ABCs Crowdsourced Flood-Mapp. Initiat. [online] Available from: http://www.abc.net.au/technology/articles/2011/01/13/3112261.htm (Accessed 20 January 2016), 2011.

Alberoni, P., Collier, C. and Khabiti, R.: ACTIF Best practice paper - Understanding and reducing uncertainty in flood forecasting, Proceeding Act. Conf., (1), 1–43, 2005.

Alfonso, L.: Use of hydroinformatics technologies for real time water quality management and operation of distribution networks. Case study of Villavicencio, Colombia, M.Sc. Thesis, UNESCO-IHE, Institute for Water Education, Delft, The Netherlands., 2006.

Alfonso, L., He, L., Lobbrecht, A. and Price, R.: Information theory applied to evaluate the discharge monitoring network of the Magdalena River, J. Hydroinformatics, 15(1), 211, doi:10.2166/hydro.2012.066, 2013.

Arnold, C. P. and Dey, C. H.: Observing-Systems Simulation Experiments: Past, Present, and Future, Bull. Am. Meteorol. Soc., 67(6), 687–695, doi:10.1175/1520-0477(1986)067<0687:OSSEPP>2.0.CO;2, 1986.

Au, J., Bagchi, P., Chen, B., Martinez, R., Dudley, S. A. and Sorger, G. J.: Methodology for public monitoring of total coliforms, Escherichia coli and toxicity in waterways by Canadian high school students, J. Environ. Manage., 58(3), 213–230, doi:10.1006/jema.2000.0323, 2000.

Aubert, D., Loumagne, C. and Oudin, L.: Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model, J. Hydrol., 280(1–4), 145–161, doi:10.1016/S0022-1694(03)00229-4, 2003.

Bergström, S.: Principles and confidence in hydrological modelling, Hydrol. Res., 22(2), 123–136, 1991.

Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J. F., Pecl, G. T., Barrett, N. and Frusher, S.: Statistical solutions for error and bias in global citizen science datasets, Biol. Conserv., 173, 144–154, doi:10.1016/j.biocon.2013.07.037, 2014.

Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M. and Rampini, A.: A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science, Inf. Sci., 258, 312–327, doi:10.1016/j.ins.2013.07.013, 2014.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., De BiÃ¨vre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D. M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G. and Zhumanova, M.: Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development, Front. Earth Sci., 2(October), 1–21, doi:10.3389/feart.2014.00026, 2014.

869 Canizares, R., Heemink, A. W. and Vested, H. J.: Application of advanced data assimilation
870 methods for the initialisation of storm surge models, J. Hydraul. Res., 36(4), 655–674,
871 doi:10.1080/00221689809498614, 1998.

872 Célleri, R., Buytaert, W., De Bièvre, B., Tobón, C., Crespo, P., Molina, J. and Feyen, J.:
873 Understanding the hydrology of tropical Andean ecosystems through an Andean Network of
874 Basins, [online] Available from: http://dspace.ucuenca.edu.ec/handle/123456789/22089
875 (Accessed 19 February 2016), 2009.

876 Cifelli, R., Doesken, N., Kennedy, P., Carey, L. D., Rutledge, S. A., Gimmestad, C. and Depue,
877 T.: The Community Collaborative Rain, Hail, and Snow Network: Informal Education for
878 Scientists and Citizens, Bull. Am. Meteorol. Soc., 86(8), 1069–1077, 2005.

879 Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J. and
880 Uddstrom, M. J.: Hydrological data assimilation with the ensemble Kalman filter: Use of
881 streamflow observations to update states in a distributed hydrological model, Adv. Water
882 Resour., 31(10), 1309–1324, doi:10.1016/j.advwatres.2008.06.005, 2008.

883 Cortes Arevalo, V. J., Charrière, M., Bossi, G., Frigerio, S., Schenato, L., Bogaard, T.,
884 Bianchizza, C., Pasuto, A. and Sterlacchini, S.: Evaluating data quality collected by volunteers
885 for first-level inspection of hydraulic structures in mountain catchments, Nat. Hazards Earth
886 Syst. Sci., 14(10), 2681–2698, doi:10.5194/nhess-14-2681-2014, 2014.

887 Degrossi, L. C., Do Amaral, G. G., da Vasconcelos, E. S. M., Albuquerque, J. P. and Ueyama,
888 J.: Using Wireless Sensor Networks in the Sensor Web for Flood Monitoring in Brazil, in
889 Proceedings of the 10th International ISCRAM Conference, Baden-Baden, Germany. [online]
890 Available from:
891 http://humanitariancomp.referata.com/wiki/Using_Wireless_Sensor_Networks_in_the_Sensor
892 _Web_for_Flood_Monitoring_in_Brazil (Accessed 10 February 2016), 2013.

893 Derber, J. and Rosati, A.: A Global Oceanic Data Assimilation System, J. Phys. Oceanogr.,
894 19(9), 1333–1347, doi:10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2, 1989.

895 DHI: MIKE FLOOD User Manual, 2005.

896 Drecourt, J.-P.: Data assimilation in hydrological modelling, Environment & Resources DTU.
897 Technical University of Denmark., 2004.

898 Eckhardt, K.: How to construct recursive digital filters for baseflow separation, Hydrol.
899 Process., 19(2), 507–515, doi:10.1002/hyp.5675, 2005.

900 Engel, S. R. and Voshell Jr, J. R.: Volunteer biological monitoring: can it accurately assess the
901 ecological condition of streams?, Am. Entomol., 48(3), 164–177, 2002.

902 Errico, R. M. and Privé, N. C.: An estimate of some analysis-error statistics using the Global
903 Modeling and Assimilation Office observing-system simulation framework, Q. J. R. Meteorol.
904 Soc., 140(680), 1005–1012, doi:10.1002/qj.2180, 2014.

905 Errico, R. M., Yang, R., Privé, N. C., Tai, K.-S., Todling, R., Sienkiewicz, M. E. and Guo, J.:
906 Development and validation of observing-system simulation experiments at NASA's Global

907 Modeling and Assimilation Office, Q. J. R. Meteorol. Soc., 139(674), 1162–1178,
908 doi:10.1002/qj.2027, 2013.

909 Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, 2nd ed. 2009 edition., Springer,
910 Place of publication not identified., 2006.

911 Fenicia, F., Solomatine, D. P., Savenije, H. H. G. and Matgen, P.: Soft combination of local
912 models in a multi-objective framework, Hydrol. Earth Syst. Sci. Discuss., 4, 91–123,
913 doi:10.5194/hessd-4-91-2007, 2007.

914 Ferri, M., Monego, M., Norbiato, D., Baruffi, F., Toffolon, C. and Casarin, R.: La piattaforma
915 previsionale per i bacini idrografici del Nord Est Adriatico (I), in Proc.XXXIII Conference of
916 Hydraulics and Hydraulic Engineering, p. 10, Brescia., 2012.

917 Giandotti, M.: Previsione delle piene e delle magre dei corsi d'acqua, Servizio Idrografico
918 Italiano, Rome., 1933.

919 Hargreaves, G.H. and Samani, Z.A.: Estimating potential evapotranspiration, J. Irrig. Drain.
920 Div., 108(3), 225–230, 1982.

921 Huang, B., Kinter, J. L. and Schopf, P. S.: Ocean data assimilation using intermittent analyses
922 and continuous model error correction, Adv. Atmospheric Sci., 19(6), 965–992,
923 doi:10.1007/s00376-002-0059-z, 2002.

924 Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J.
925 A. and Zimin, A. V.: Four-dimensional Ensemble Kalman Filtering, Tellus A, 56(4), 273–277,
926 doi:10.1111/j.1600-0870.2004.00066.x, 2004.

927 Huwald, H., Barrenetxea, G., de Jong, S., Ferri, M., Carvalho, R., Lanfranchi, V., McCarthy,
928 S., Glorioso, G., Prior, S., Solà, E., Gil-Roldàn, E., Alfonso, L., Wehn de Montalvo, U.,
929 Onencan, A., Solomatine, D. and Lobbrecht, A.: D1.11 Sensor technology requirement
930 analysis, Confidential Deliverable, The WeSenseIt Project (FP7/2007-2013 grant agreement no
931 308429)., 2013.

932 Ide, K., Courtier, P., Ghil, M. and Lorenc, A. C.: Unifed notation for data assimilation:
933 operational, sequential and variational, J. Meteorol. Soc. Jpn., 75(1B), 181–189, 1997.

934 ISPUW: iSPUW: Integrated Sensing and Prediction of Urban Water for Sustainable Cities,
935 [online] Available from: http://ispuw.uta.edu/nsf/ (Accessed 19 February 2016), 2015.

936 Kalman, R. E.: A new approach to linear filtering and prediction problems, J. Basic Eng., 82(1),
937 35–45, doi:10.1115/1.3662552, 1960.

938 Krstanovic, P. F. and Singh, V. P.: Evaluation of rainfall networks using entropy: II.
939 Application, Water Resour. Manag., 6(4), 295–314, doi:10.1007/BF00872282, 1992.

940 Kumar, R., Chatterjee, C., Lohani, A. K., Kumar, S. and Singh, R. D.: Sensitivity Analysis of
941 the GIUH based Clark Model for a Catchment, Water Resour. Manag., 16(4), 263–278,
942 doi:10.1023/A:1021920717410, 2002.

943 Laio, F., Porporato, A., Ridolfi, L. and Rodriguez-Iturbe, I.: Plants in water-controlled
944 ecosystems: active role in hydrologic processes and response to water stress: II. Probabilistic
945 soil moisture dynamics, Adv. Water Resour., 24(7), 707–723, doi:10.1016/S0309-
946 1708(01)00005-7, 2001.

947 Li, Z. and Navon, I. M.: Optimality of variational data assimilation and its relationship with the
948 Kalman filter and smoother, Q. J. R. Meteorol. Soc., 127(572), 661–683,
949 doi:10.1002/qj.49712757220, 2001.

950 Lowry, C. S. and Fienen, M. N.: CrowdHydrology: Crowdsourcing hydrologic data and
951 engaging citizen scientists, GroundWater, 51(1), 151–156, doi:10.1111/j.1745-
952 6584.2012.00956.x, 2013.

953 Macpherson, B.: Dynamic initialization by repeated insertion of data, Q. J. R. Meteorol. Soc.,
954 117(501), 965–991, doi:10.1002/qj.49711750105, 1991.

955 Madsen, H. and Cañizares, R.: Comparison of extended and ensemble Kalman filters for data
956 assimilation in coastal area modelling, Int. J. Numer. Methods Fluids, 31(6), 961–981,
957 doi:10.1002/(SICI)1097-0363(19991130)31:6<961::AID-FLD907>3.0.CO;2-0, 1999.

958 Massart, S., Pajot, B., Piacentini, A. and Pannekoucke, O.: On the merits of using a 3D-FGAT
959 assimilation scheme with an outer loop for atmospheric situations governed by transport, Mon.
960 Weather Rev., 138(12), 4509–4522, 2010.

961 Matheron, G.: Principles of geostatistics, Econ. Geol., 58(8), 1246–1266, 1963.

962 Mazzoleni, M., Alfonso, L. and Solomatine, D.: Effect Of Different Hydrological Model
963 Structures On The Assimilation Of Distributed Uncertain Observations, Int. Conf.
964 Hydroinformatics [online] Available from: http://academicworks.cuny.edu/cc_conf_hic/114,
965 2014.

966 Mazzoleni, M., Alfonso, L., Chacon-Hurtado, J. and Solomatine, D.: Assimilating uncertain,
967 dynamic and intermittent streamflow observations in hydrological models, Adv. Water Resour.,
968 83, 323–339, 2015.

969 McDonnell, J. J. and Beven, K.: Debates—The future of hydrological sciences: A (common)
970 path forward? A call to action aimed at understanding velocities, celerities and residence time
971 distributions of the headwater hydrograph, Water Resour. Res., 50(6), 5342–5350,
972 doi:10.1002/2013WR015141, 2014.

973 Moore, R. J., Jones, D. A., Cox, D. R. and Isham, V. S.: Design of the HYREX raingauge
974 network, Hydrol. Earth Syst. Sci., 4(4), 521–530, doi:10.5194/hess-4-521-2000, 2000.

975 Ragnoli, E., Zhuk, S., Donncha, F. O., Suits, F. and Hartnett, M.: An optimal interpolation
976 scheme for assimilation of HF radar current data into a numerical ocean model, in Oceans,
977 2012, pp. 1–5., 2012.

978 Rakovec, O., Weerts, A. H., Hazenberg, P., F. Torfs, P. J. J. and Uijlenhoet, R.: State updating
979 of a distributed hydrological model with ensemble kalman Filtering: Effects of updating
980 frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci.,
981 16(9), 3435–3449, doi:10.5194/hess-16-3435-2012, 2012.

Rakovec, O., Weerts, A. H., Sumihar, J. and Uijlenhoet, R.: Operational aspects of asynchronous filtering for flood forecasting, Hydrol. Earth Syst. Sci., 19(6), 2911–2924, doi:10.5194/hess-19-2911-2015, 2015.

Refsgaard, J. C.: Validation and Intercomparison of Different Updating Procedures for Real-Time Forecasting, Nord. Hydrol., 28(2), 65–84, doi:10.2166/nh.1997.005, 1997.

Ridolfi, E., Alfonso, L., Baldassarre, G. D., Dottori, F., Russo, F. and Napolitano, F.: An entropy approach for the optimization of cross-section spacing for river modelling, Hydrol. Sci. J., 59(1), 126–137, doi:10.1080/02626667.2013.822640, 2014.

Rinaldo, A. and Rodriguez-Iturbe, I.: Geomorphological Theory of the Hydrological Response, Hydrol. Process., 10(6), 803–829, doi:10.1002/(SICI)1099-1085(199606)10:6<803::AID-HYP373>3.0.CO;2-N, 1996.

Rodríguez-Iturbe, I., González-Sanabria, M. and Bras, R. L.: A geomorphoclimatic theory of the instantaneous unit hydrograph, Water Resour. Res., 18(4), 877–886, doi:10.1029/WR018i004p00877, 1982.

Roy, H. E., Pocock, M. J. O., Preston, C. D., Roy, D. B. and Savage, J.: Understanding Citizen Science and Environmental Monitoring, Final Report of UK Environmental Observation Framework., 2012.

Sakov, P., Evensen, G. and Bertino, L.: Asynchronous data assimilation with the EnKF, Tellus A, 62(1), 24–29, doi:10.1111/j.1600-0870.2009.00417.x, 2010.

Seo, D. ., Kerke, B., Zink, M., Fang, N., Gao, J. and Yu, X.: iSPUW: A Vision for Integrated Sensing and Prediction of Urban Water for Sustainable Cities., 2014.

Solomatine, D. P. and Dulal, K. N.: Model trees as an alternative to neural networks in rainfall—runoff modelling, Hydrol. Sci. J., 48(3), 399–411, doi:10.1623/hysj.48.3.399.45291, 2003.

Szilagyi, J. and Szollosi-Nagy, A.: Recursive Streamflow Forecasting: A State Space Approach - CRC Press Book., 2010.

Todini, E.: A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach, Hydrol. Earth Syst. Sci., 11, 1645–1659, 2007.

Tulloch, A. I. T. and Szabo, J. K.: A behavioural ecology approach to understand volunteer surveying for citizen science datasets, Emu, 112(4), 313, doi:10.1071/MU12009, 2012.

Vandecasteele, A. and Devillers, R.: Improving volunteered geographic data quality using semantic similarity measurements, ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., 1(1), 143–148, 2013.

Verlaan, M.: Efficient Kalman Filtering Algorithms for Hydrodynamic Models, PhD Thesis, Delft University of Technology, The Netherlands., 1998.

1017 Weerts, A. H. and El Serafy, G. Y. H.: Particle filtering and ensemble Kalman filtering for state
1018 updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42(9), 1–
1019 17, doi:10.1029/2005WR004093, 2006.

1020 WMO: Simulated real-time intercomparison of hydrological models, World Meteorological
1021 Organization., 1992.

1022 Wood, S. J., Jones, D. A. and Moore, R. J.: Accuracy of rainfall measurement for scales of
1023 hydrological interest, Hydrol. Earth Syst. Sci. Discuss., 4(4), 531–543, 2000.

1024

1025

1026

1027

Tables

Table 1. NSE improvements, from 1 to 50 crowdsourced observations, in case of different experimental scenarios during the nine flood events occurred in the Brue, Sieve and Alzette catchments.

Scenario	1	2	3	4	5	6	7	8	9
Brue - event 1	0.126	0.125	0.140	0.243	0.253	0.125	0.144	0.237	0.248
Brue - event 2	0.416	0.413	0.445	0.920	0.902	0.413	0.463	0.841	0.870
Brue - event 3	0.443	0.438	0.472	0.890	0.842	0.440	0.471	0.809	0.822
Sieve - event 1	0.250	0.246	0.228	0.271	0.221	0.247	0.225	0.263	0.237
Sieve - event 2	0.066	0.064	0.067	0.057	0.056	0.064	0.068	0.057	0.060
Sieve - event 3	0.629	0.623	0.632	1.085	1.045	0.625	0.634	1.019	0.995
Alzette - event 1	0.884	0.881	0.883	1.274	1.265	0.882	0.890	1.251	1.342
Alzette - event 2	0.137	0.135	0.135	0.120	0.121	0.134	0.147	0.119	0.135
Alzette - event 3	0.314	0.309	0.305	0.297	0.283	0.310	0.315	0.297	0.281

Figures

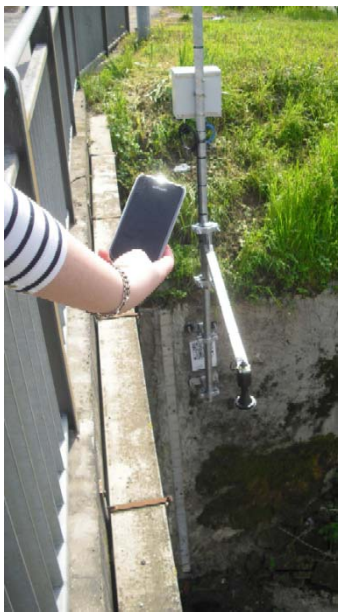

Figure 1. Example of a low-cost social sensor, and crowdsourced observations, implemented in the Bacchiglione River, Italy, under the WeSenseIt project

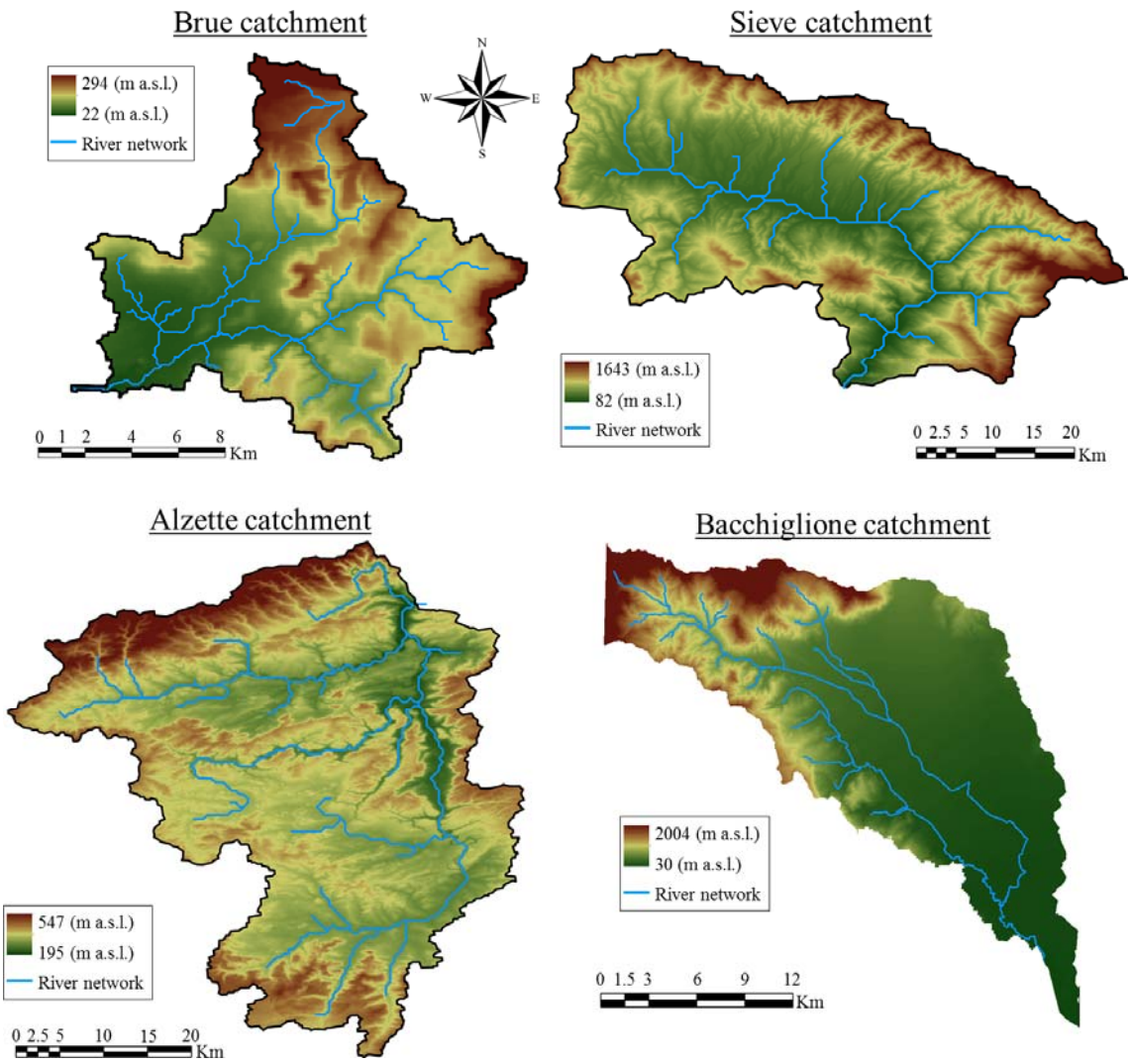

Figure 2. Representation of the four case studies considered in this study

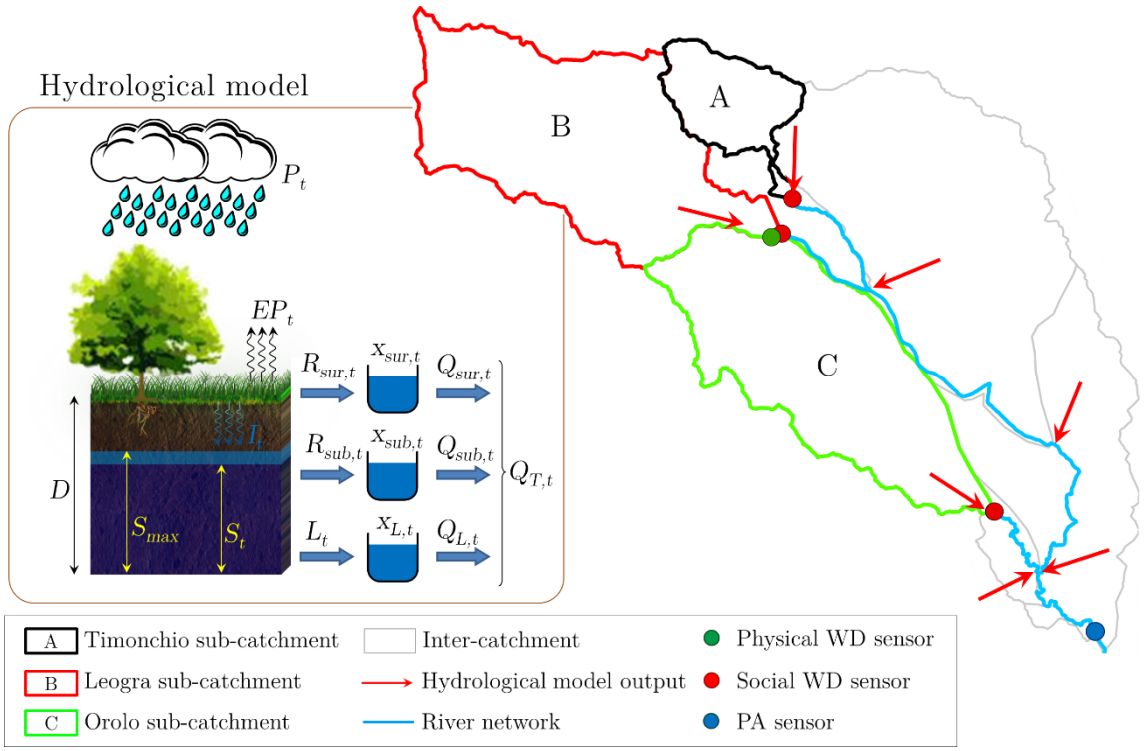

Figure 3. Structure of the early warning system AMICO and location of the physical, social and Ponte degli Angeli (PA) sensors implemented in the Bacchiglione catchment by the Alto Adriatico Water Authority

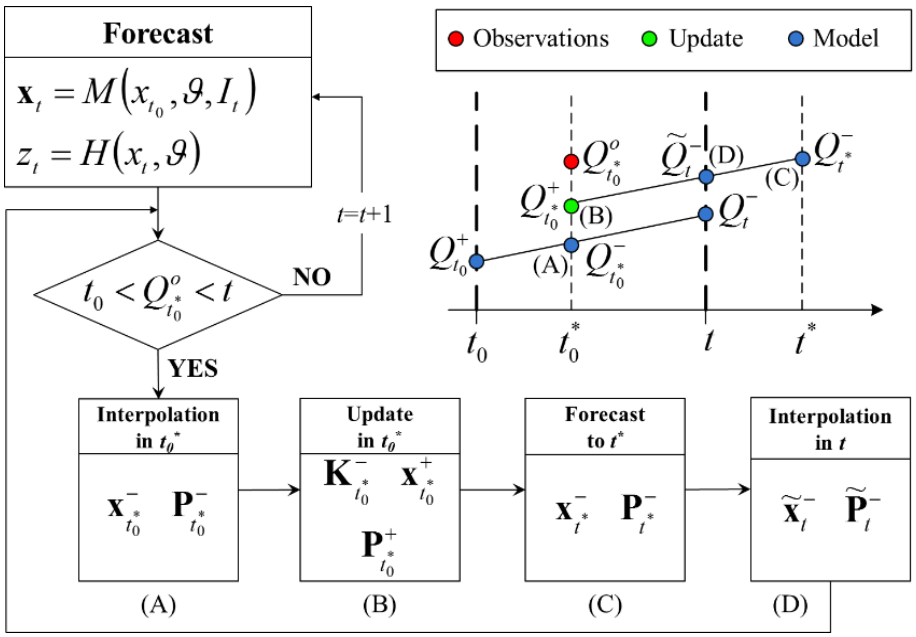

Figure 4. Graphical representation of the DACO method proposed in this study to assimilate crowdsourced asynchronous observations

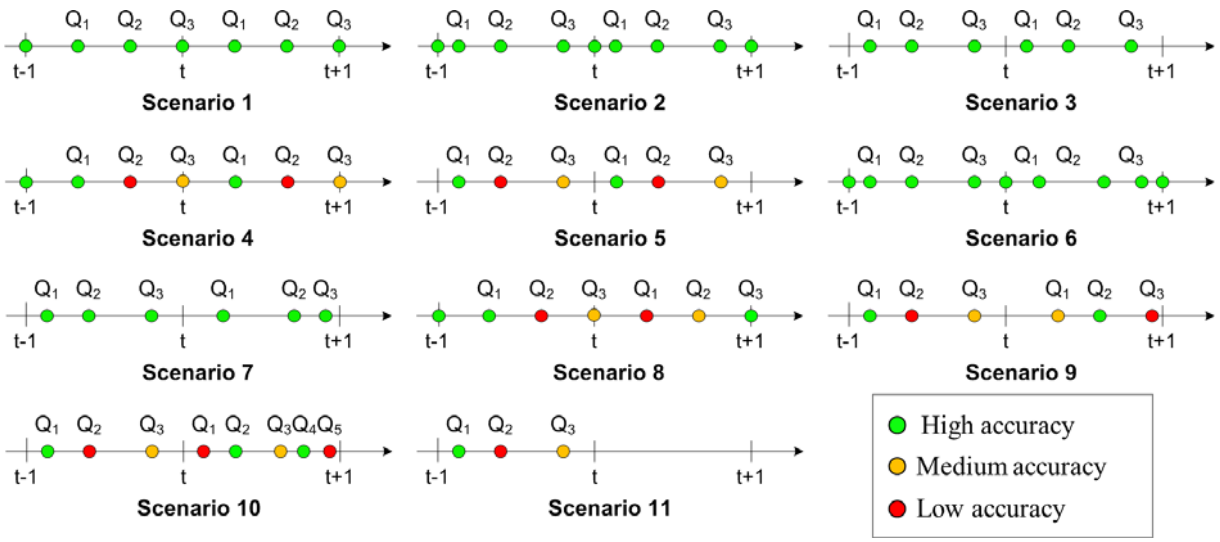

Figure 5. The experimental scenarios representing different configurations of arrival frequency, number and accuracy of the streamflow observations

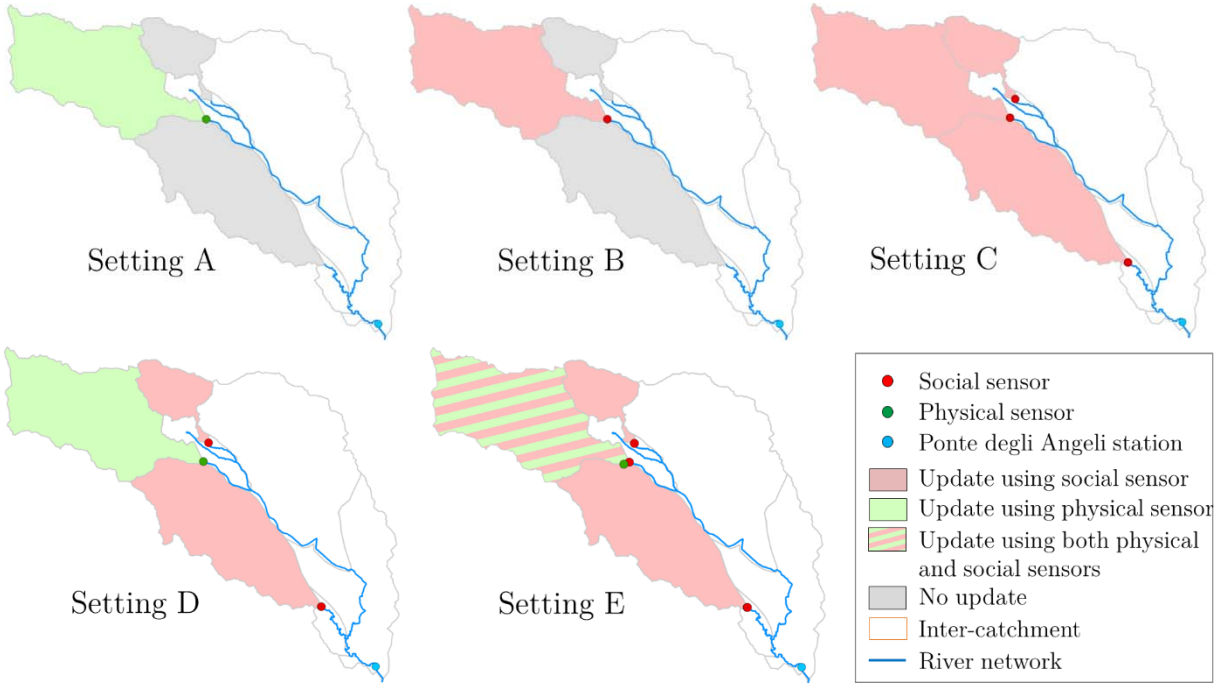

Figure 6. Different experimental settings implemented within the Bacchiglione catchment during Experiments 2

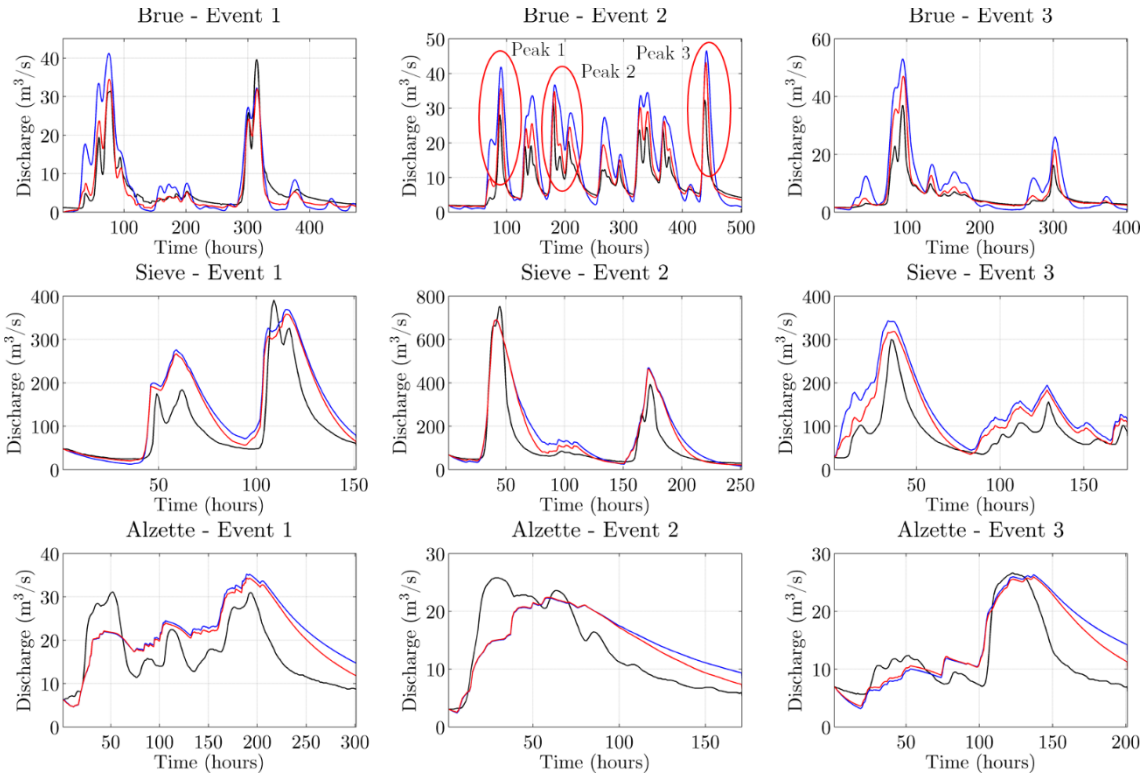

Figure 7. The observed and simulated hydrographs, with and without assimilation, for the nine considered flood events occurred in the Brue, Sieve and Alzette catchments

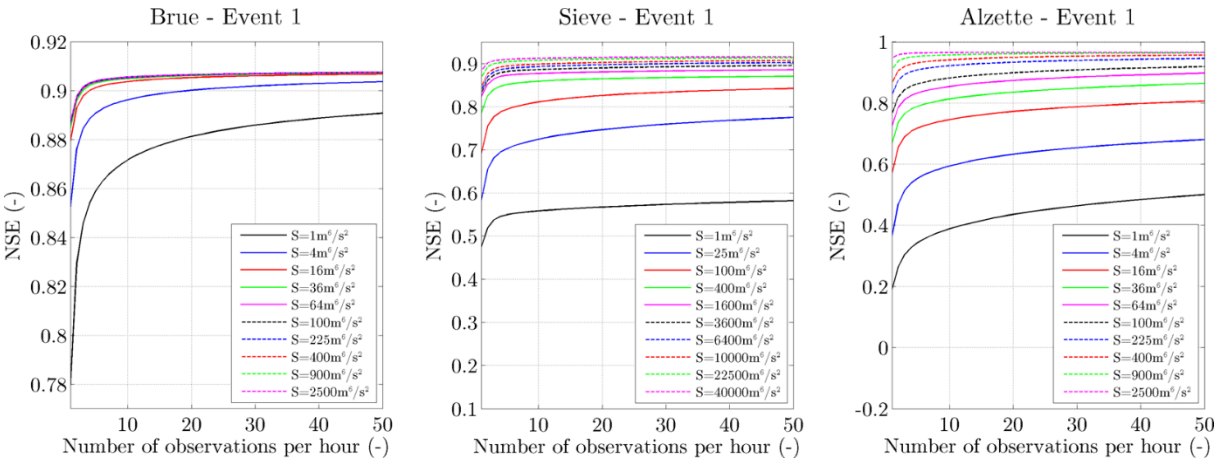

Figure 8. Model improvement in terms of NSE during flood event 1 of each case study, in case of different values of model error matrix **S** and 24-hours lead time, assimilating streamflow observations according to scenario 1

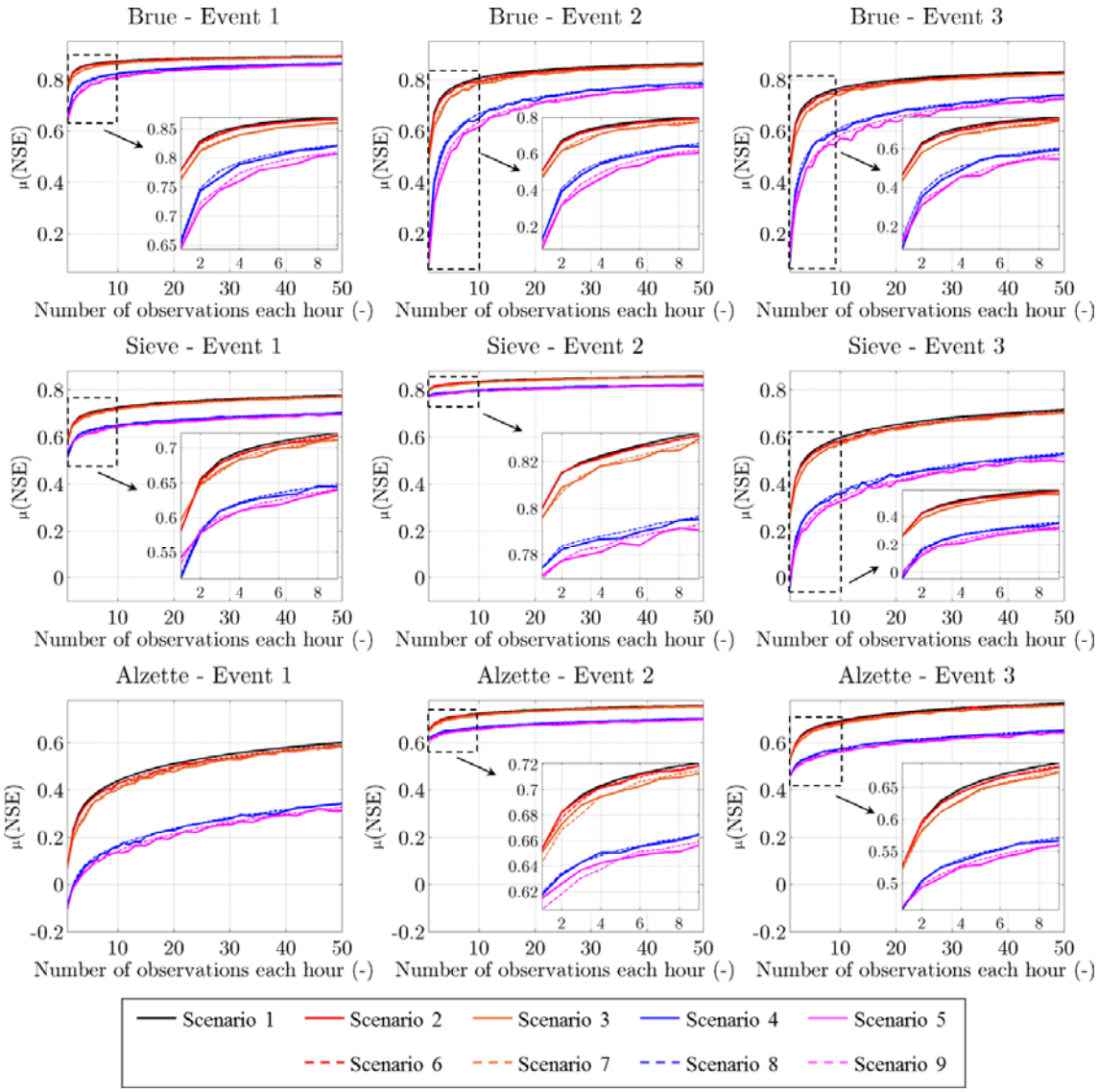

Figure 9. Dependency of μ(NSE) on the number of observations, for the scenarios 2, 3, 4, 5, 6, 7, 8 and 9 for the five considered flood events

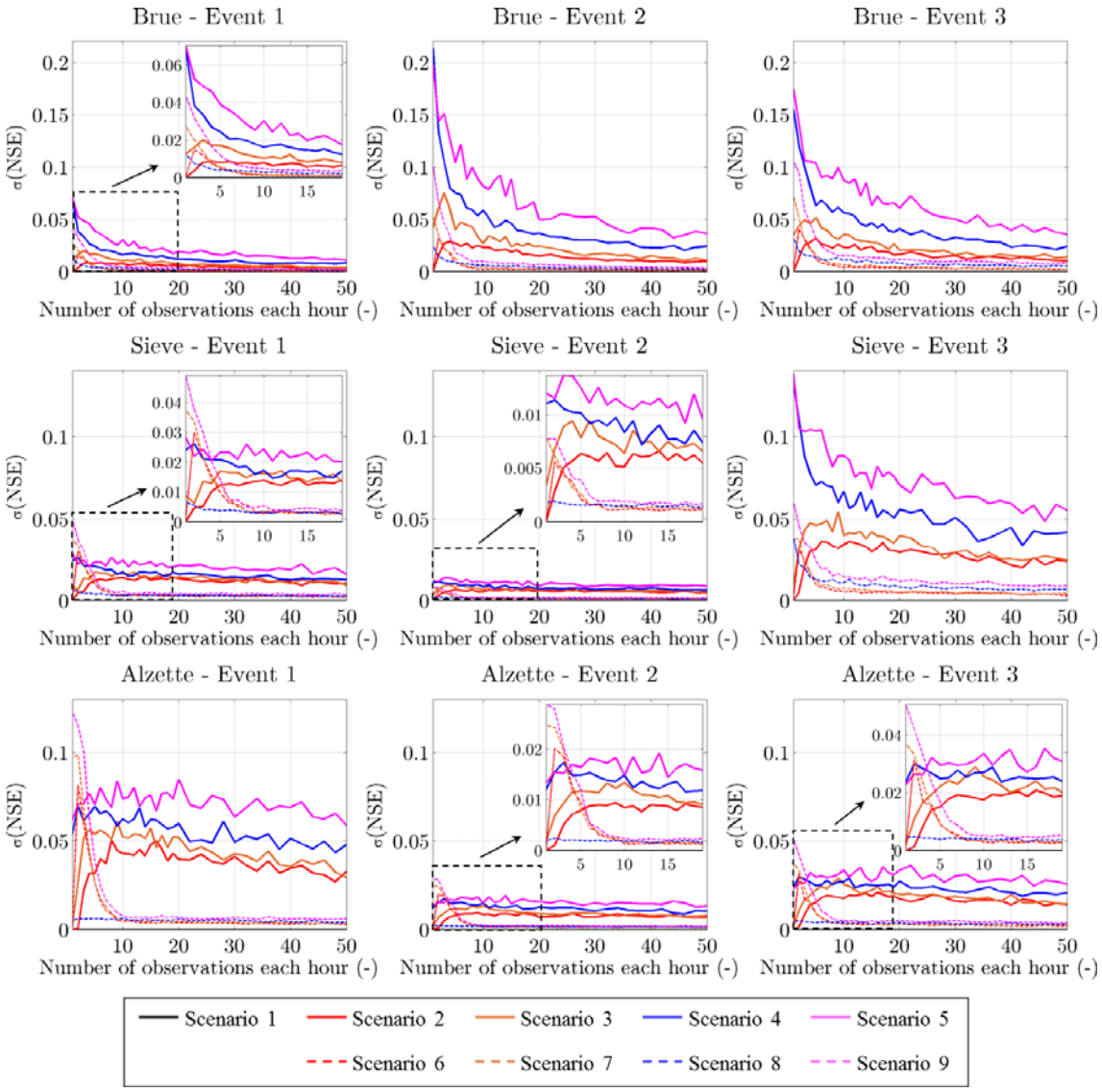

Figure 10. Dependency of σ(NSE) on the number of observations, for the scenarios 2, 3, 4, 5, 6, 7, 8 and 9 for the five considered flood events

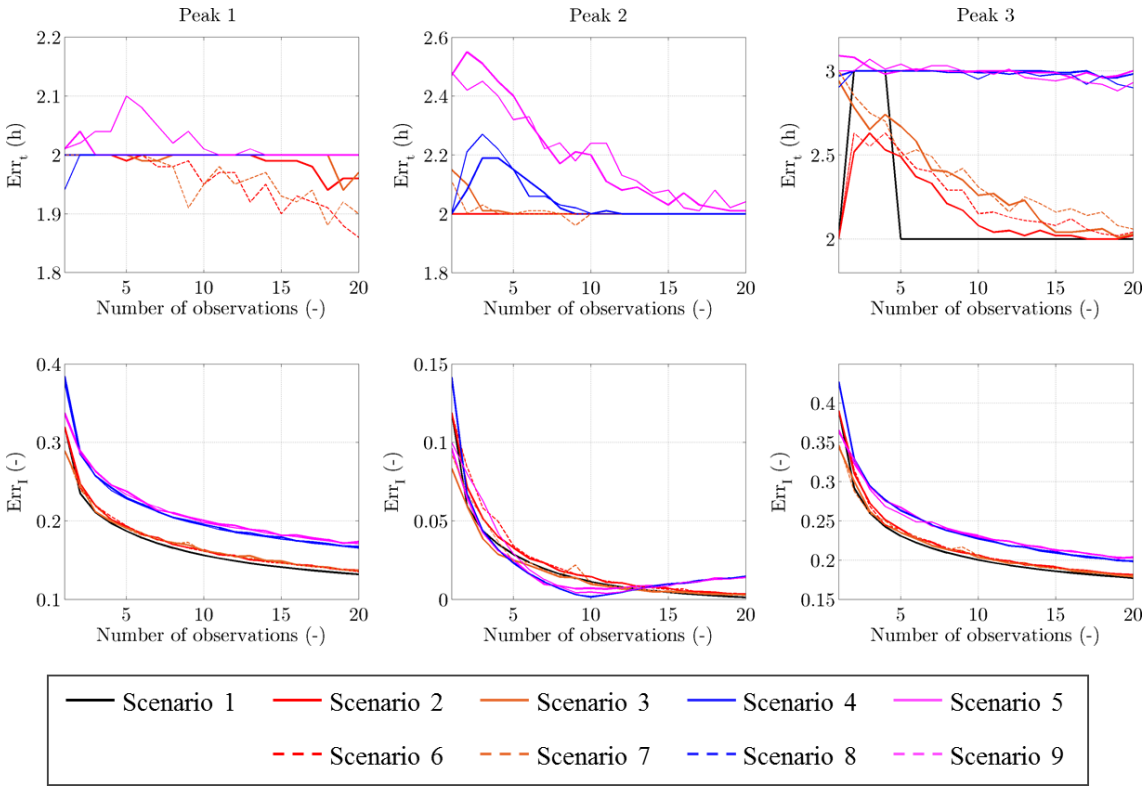

Figure 11. Representation of the Err_t and Err_I as function of number of crowdsourced observations and experimental scenarios for three different flood peaks occurred during flood event 2 in Brue catcment

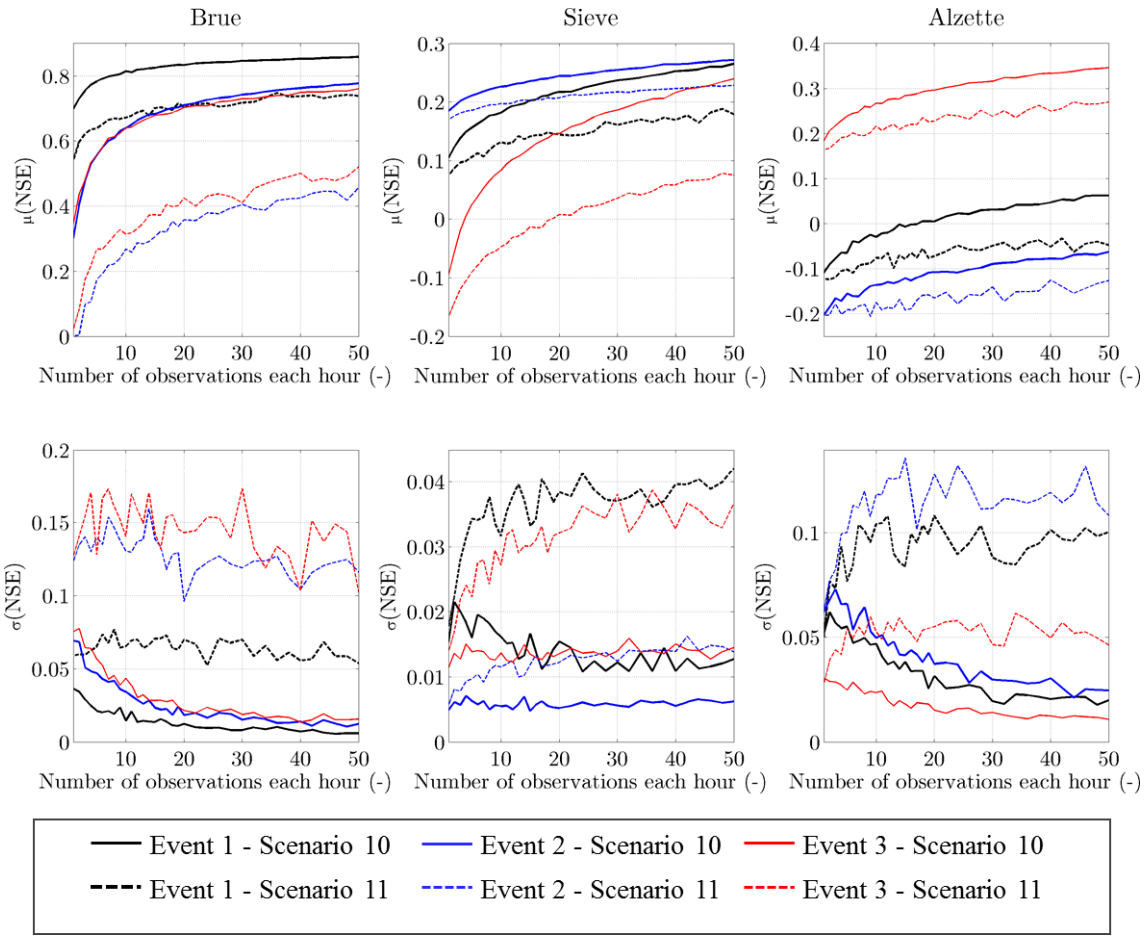

Figure 12. Dependency of the μ(NSE) and σ(NSE) on the number of observations, for the scenarios 10 and 11 in case of the five flood events

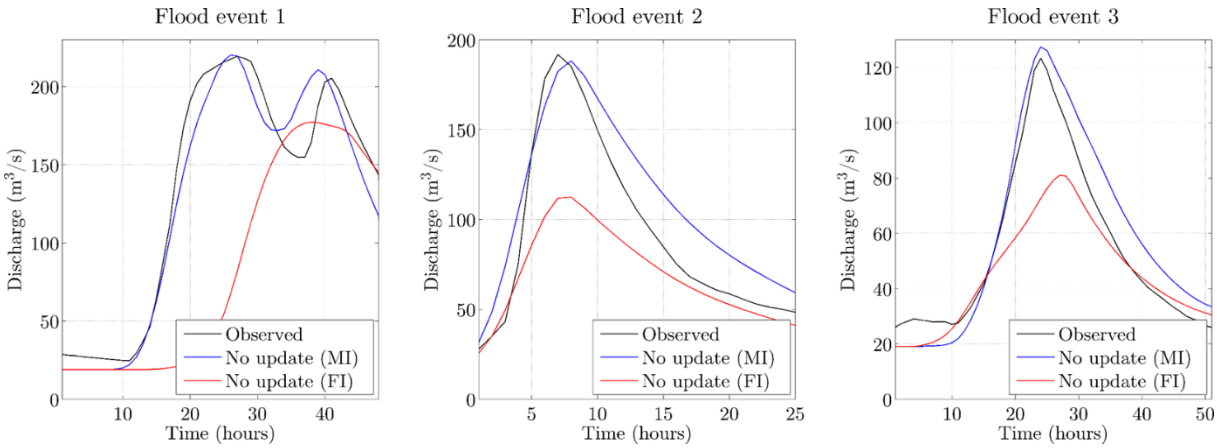

Figure 13. The observed and simulated hydrographs, without update, using measured input (MI) and forecasted input (FI), for the three considered flood events occurred in 2013 (event 1), 2014 (event 2) and 2016 (event 3) on the Bacchiglione catchment

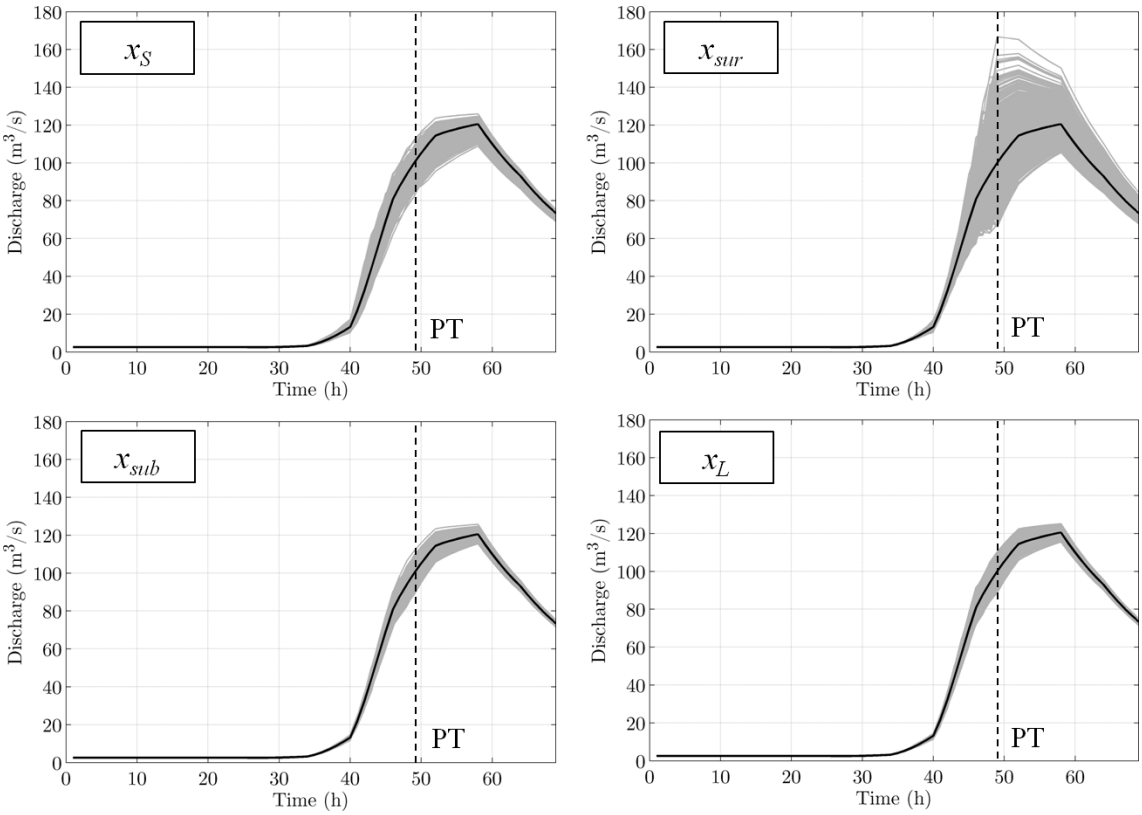

Figure 14. Effect of perturbing the model states on the model output, Bacchiglione case study. PT=Perturbation Time

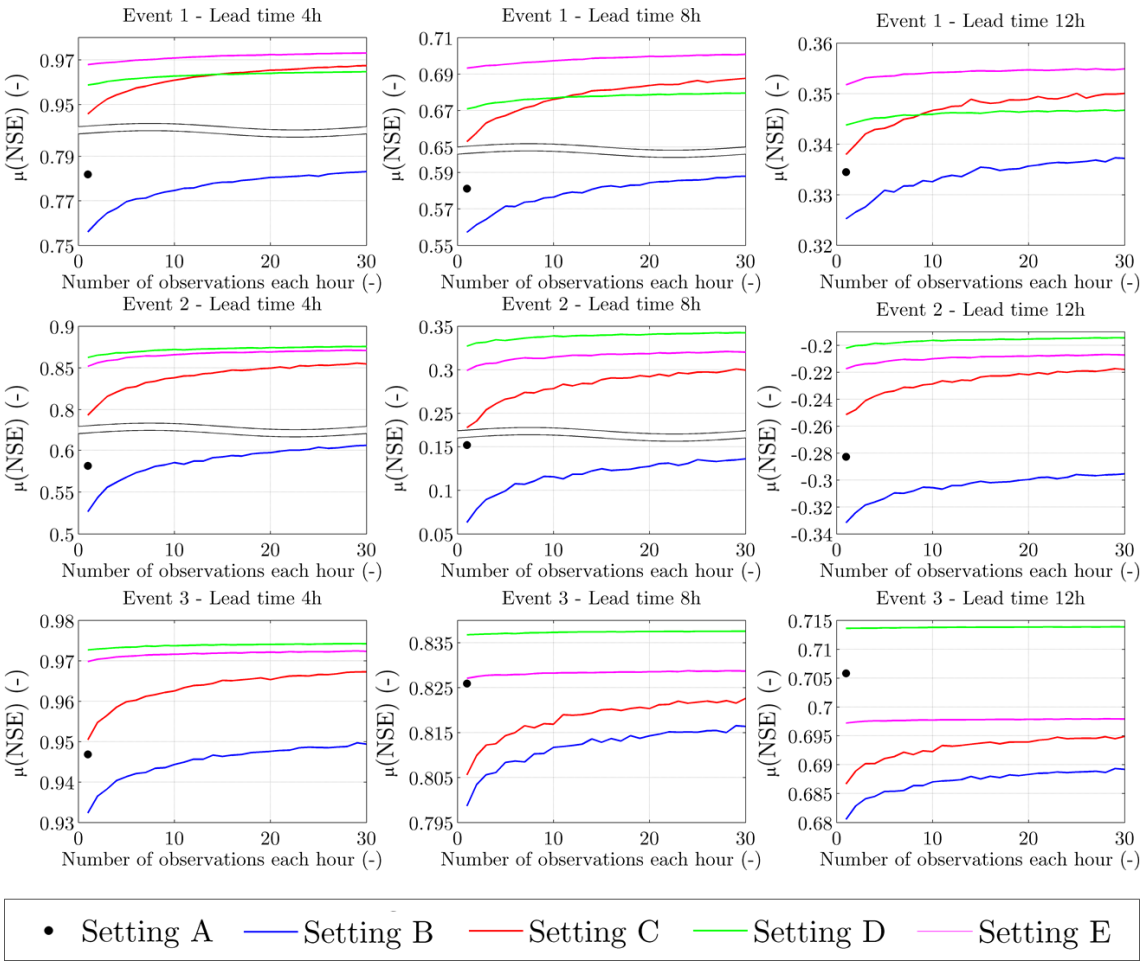

Figure 15. Model performance expressed as μ(NSE) – assimilating different number of crowdsourced observations during the three considered flood events, for the three lead time values, having characteristic of scenario 10

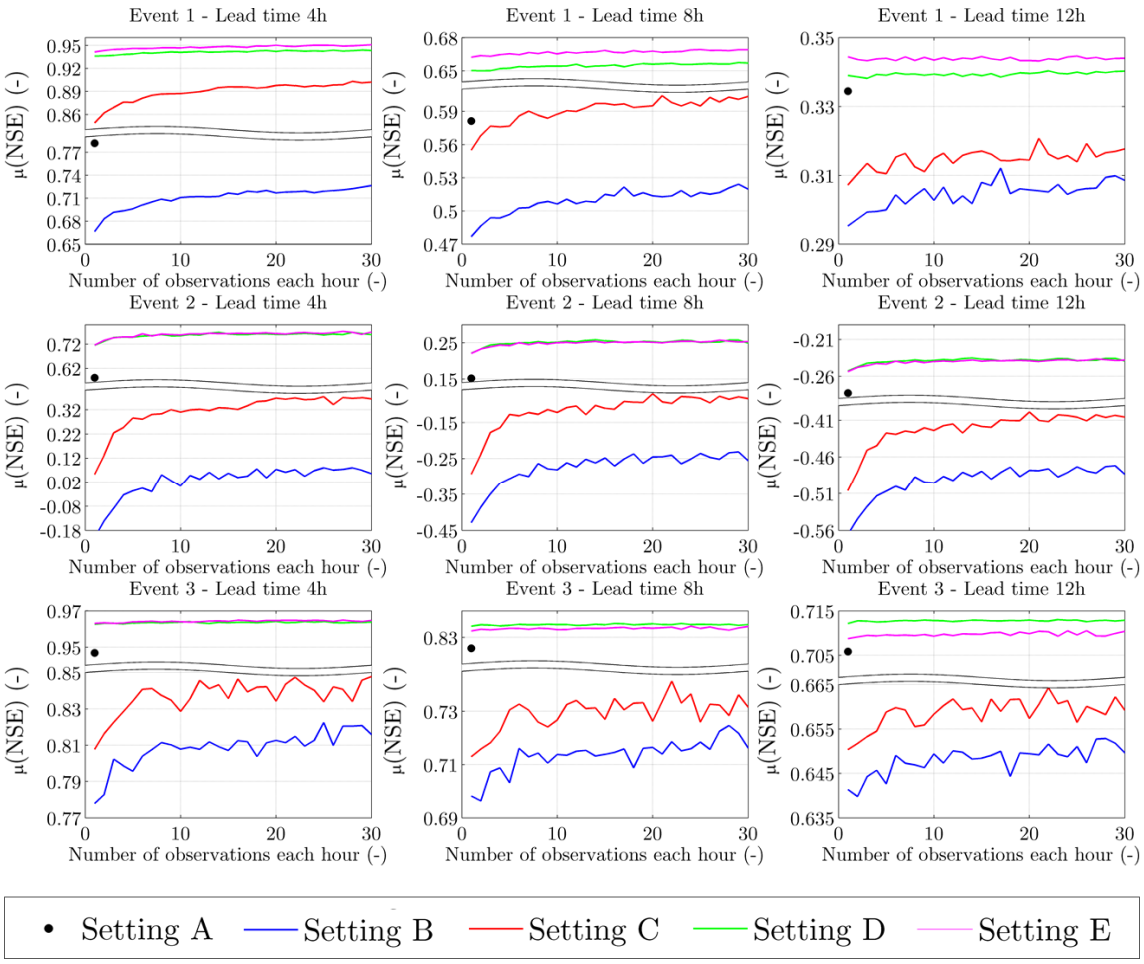

Figure 16. Model performance expressed as μ(NSE) – assimilating different number of crowdsourced observations during the three considered flood events, for the three lead time values, having characteristic of scenario 11

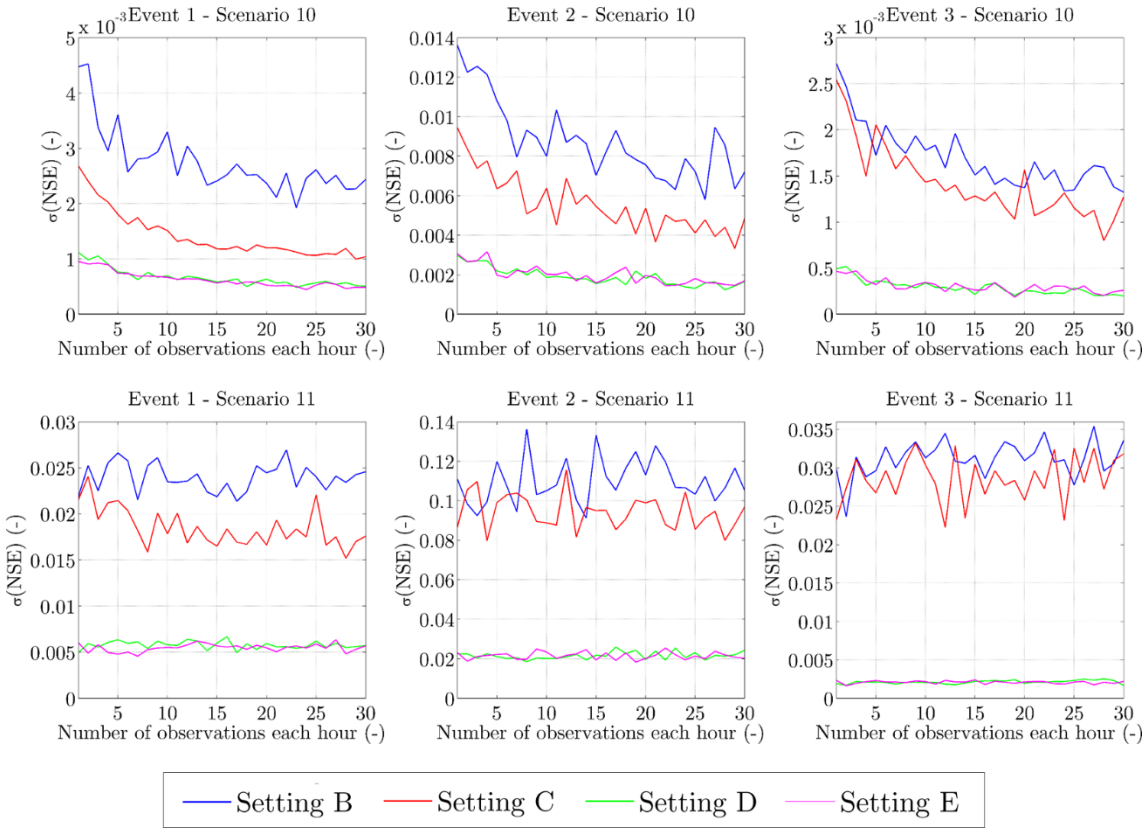

Figure 17. Variability of performance expressed as σ(NSE) – assimilating crowdsourced observations within setting A, B, C and D, assuming the lead time of 4h, for scenarios 10 and 11 during the three considered flood events