University of Reading

Satellite-supported flood forecasting in river networks: a real case study

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Open Access

Garcia-Pintado, J., Mason, D., Dance, S. L. ORCID: https://orcid.org/0000-0003-1690-3338, Cloke, H. ORCID: https://orcid.org/0000-0002-1472-868X, Neal, J. C., Freer, J. and Bates, P. D. (2015) Satellite-supported flood forecasting in river networks: a real case study. Journal of Hydrology, 523. pp. 706-724. ISSN 0022-1694 doi: https://doi.org/10.1016/j.jhydrol.2015.01.084 Available at https://centaur.reading.ac.uk/39388/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.jhydrol.2015.01.084

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Journal of Hydrology 523 (2015) 706–724

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Satellite-supported flood forecasting in river networks: A real case study

Javier García-Pintado [a,b,*], David C. Mason [c], Sarah L. Dance [a,b], Hannah L. Cloke [a,c], Jeff C. Neal [d], Jim Freer [d], Paul D. Bates [d]

[a] School of Mathematical and Physical Sciences, University of Reading, UK
[b] National Centre for Earth Observation, University of Reading, Reading, UK
[c] School of Archaeology, Geography and Environmental Science, University of Reading, UK
[d] School of Geographical Sciences, University of Bristol, Bristol, UK

ARTICLE INFO

Article history:
Received 1 October 2014
Received in revised form 27 January 2015
Accepted 31 January 2015
Available online 9 February 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Yu Zhang, Associate Editor

Keywords:
Data assimilation
Earth Observation
Synthetic aperture radar
Flood forecast
Ensemble Kalman filter
Observation localization

SUMMARY

Satellite-based (e.g., Synthetic Aperture Radar [SAR]) water level observations (WLOs) of the floodplain can be sequentially assimilated into a hydrodynamic model to decrease forecast uncertainty. This has the potential to keep the forecast on track, so providing an Earth Observation (EO) based flood forecast system. However, the operational applicability of such a system for floods developed over river networks requires further testing. One of the promising techniques for assimilation in this field is the family of ensemble Kalman (EnKF) filters. These filters use a limited-size ensemble representation of the forecast error covariance matrix. This representation tends to develop spurious correlations as the forecast-assimilation cycle proceeds, which is a further complication for dealing with floods in either urban areas or river junctions in rural environments. Here we evaluate the assimilation of WLOs obtained from a sequence of real SAR overpasses (the X-band COSMO-Skymed constellation) in a case study. We show that a direct application of a global Ensemble Transform Kalman Filter (ETKF) suffers from filter divergence caused by spurious correlations. However, a spatially-based filter localization provides a substantial moderation in the development of the forecast error covariance matrix, directly improving the forecast and also making it possible to further benefit from a simultaneous online inflow error estimation and correction. Additionally, we propose and evaluate a novel along-network metric for filter localization, which is physically-meaningful for the flood over a network problem. Using this metric, we further evaluate the simultaneous estimation of channel friction and spatially-variable channel bathymetry, for which the filter seems able to converge simultaneously to sensible values. Results also indicate that friction is a second order effect in flood inundation models applied to gradually varied flow in large rivers. The study is not conclusive regarding whether in an operational situation the simultaneous estimation of friction and bathymetry helps the current forecast. Overall, the results indicate the feasibility of stand-alone EO-based operational flood forecasting.
© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

While there are recent advances in low-cost telemetered networks for long-life flood monitoring and warning applications, oriented to be deployed over large areas (e.g., Marín-Pérez et al., 2012), the actual number of operational gauges is actually declining in the world (Vörösmarty et al., 2001). On the other hand, in recent times, the technology of earth observation (EO) has begun to be adopted to improve flood visualization and reduce flood modeling uncertainties (e.g., Raclot, 2006, Schumann et al., 2007, 2011, Mason et al., 2010a). EO techniques for flood detection include, for example, high resolution Synthetic Aperture Radar (SAR) (such as TerraSAR-X), altimetry (such as RA-2 on Envisat, or Poseidon 3 on Jason-2), though the footprints are such that they are limited to level measurements in rivers >1 km wide, or even gravimetry (GRACE) for very large flood events. Specifically, in real-time mode, the assimilation of water level observations (WLOs) derived from EO may serve to keep forecasts obtained from flood simulations on track and, in hindcast mode, to obtain better estimates of the dynamic footprints of past flood events. The forecast mode may be used by civil protection services and industry for operational uses, while the post-flood mode may be used in damage assessment and flood defence design studies (Mason et al.,

* Corresponding author at: ESSC, Harry Pitt Building, 3 Earley Gate, University of Reading, Whiteknights Campus, RG6 6AH, Reading, United Kingdom. Tel.: +44 (0) 118 378 7722.
E-mail address: j.garcia-pintado@reading.ac.uk (J. García-Pintado).

http://dx.doi.org/10.1016/j.jhydrol.2015.01.084
0022-1694/© 2015 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

2014). In both modes, the key variable to forecast is the water level (and hence flood extent) (e.g., Hostache et al., 2010; Biancamaria et al., 2011). However, the assimilation of EO-based WLOs for the water level estimation problem, mostly in a forecast situation, is plagued by problems derived from errors in (a) inflows (discharge) into the modeled domain, and (b) model parameters (mostly friction) and bathymetry. Thus, the estimation and correction of (a) and (b), in a data assimilation (DA) context, have become themselves a major or secondary objective in recent studies. In an alternative scenario, with a focus on large rivers and longer timescales, discharge estimation, in itself, is a important objective for land surface water budget analyses (e.g., Andreadis et al., 2007; Balsamo et al., 2013). Overall, the estimation of water levels, water discharge and model parameter estimation are all inter-related in the abovementioned contexts. For example, Neal et al. (2009) conducted a case study of simultaneous water level and inflow estimation from EO-based WLOs, in a 10-km river reach. Also, Matgen et al. (2010) and Giustarini et al. (2011), in a synthetic case and a real case respectively, showed that the simultaneous correction of inflow errors led to improved water level forecast in a 19-km river reach. García-Pintado et al. (2013) conducted a synthetic sensitivity analysis of the flood forecast skills at different time horizons to satellite-based SAR acquisitions (first visit and revisit times) in order to support the scheduling of satellite imaging for operational uses. There is not real dataset yet available to conduct a similar kind of study.

Regarding bathymetry estimation, Durand et al. (2008), in a proof-of-concept study, indicated that bathymetry is a significant source of uncertainty in estimating discharge. They conducted two synthetic experiments to assimilate EO-based WLOs, assuming known inflows, in order to estimate a mean bathymetric slope and bathymetric depth at 5 specific locations in a 250-km reach in the Amazon river. Specifically, they used synthetic observations of the proposed Surface Water and Ocean Topography (SWOT) mission. The experiments were successful, and they concluded that model errors will likely dominate over SWOT-like WLO errors. Roux and Dartus (2008) succeeded in calibrating the mean bathymetry in a channel reach using real satellite-based flood extent data, and Gessese et al. (2011) directly obtained an explicit partial differential equation (PDE) for the 1-D inverse problem, and successfully estimated the depth of a rectangular horizontal channel with a bump, with known inflow and known downstream slope. Later, Durand et al. (2014) used real SAR-based WLOs to simultaneously estimate bathymetry and lateral inflows, along with channel roughness, for a major out-of-bank flood event in a river. They treated the river in term of reaches, so that they estimated mean values for three transects along the river, and assumed the major upstream inflow was known. Their results suggest that it should be possible to estimate river discharge via EO. In closely related work, the estimation of bathymetry in river estuaries through the assimilation of SAR-based waterlines with morphodynamic models has also been evaluated (Thornhill et al., 2012; Smith et al., 2013).

On the other hand, the chosen DA method itself may have intrinsic problems. Specifically, in recent years there has been a growing interest in DA ensemble schemes for flood studies. From these, the various methods derived from the Kalman filter are generically known as Ensemble Kalman Filters (EnKFs) (Evensen, 1994). These filters use a limited-size ensemble representation of the forecast error covariance matrix. This is updated as each set of observations is assimilated. The ensemble-based error covariances tend to underestimate the forecast error variance and develop physically unrealistic or spurious correlations. This may lead to ensemble collapse and filter divergence. Filter localization is often used to reduce the problem of spurious correlations. This increases the degrees of freedom available to fit nearby observations in the analysis by decreasing the weight given to observations far from the physical location of the estimated state variable (Hamill et al., 2001).

While previous experiments are encouraging regarding EO capabilities for flood and river flow monitoring and forecasting, they focus on specific single river reaches, albeit ones that are sometimes very large or subject to secondary lateral inflows. Our new contribution is to carry out a case study assimilating real EO data for the sequential monitoring and forecast of a flood developing on a river network with tributaries. In our case, uncertain forecasts from upstream rainfall-runoff models provide the discharge at seven catchments contributing to the flood. The case is based on possibly the best example of sequential monitoring of a flood extent by high-resolution Synthetic Aperture Radar (SAR) images currently available in the world. This is a 7-image set from the COSMO-Skymed constellation, which was acquired during a flood that occurred in November 2012 around the confluence of the Severn and Avon catchments in the western UK.

The assimilation is conducted via a Local Ensemble Transform Kalman Filter (LETKF) (Hunt et al., 2007) applied to a 2-D flood model. Our objective is to evaluate a number of strategies for real-time flood forecasting by assimilating high-resolution EO-based WLOs with the flood simulations assuming uncertain model parameters. That is, we want (a) to evaluate if localization is strictly required for avoiding problems arising from spurious correlations; (b) to evaluate if the flood forecast improves by jointly estimating inflow boundary condition errors simultaneously with the water level; and (c) to evaluate if, with an imminent flood situation, it is better to focus on state estimation (water levels), joint state-inflow estimation, or joint state-parameter estimation, where at the same time uncertain friction and bathymetry are estimated.

Regarding objective (a), it is noted that filter localization requires a distance metric for moderating the weights given to the observations. In this paper we introduce an along-network distance metric for filter localization, which is new to the DA literature. The proposed metric is physically meaningful for the flood over a network problem and, accordingly, we evaluate its influence on the forecast. Regarding objective (b), it is noted that the online correction of inflow errors into the flood model domain does not affect the hydrologic simulations of the upstream catchment-scale rainfall-runoff processes.

2. Methods

2.1. Study domain

This study focuses on an area of the lower Severn and Avon rivers in the South West United Kingdom, over a 30.6×49.8 km (1524 km^2) domain. Fig. 1 depicts the study area for the flood model. Four our investigation, we used a real case based on an event that occurred in November 2012. We used a previous event in July 2007 in the same location as a calibration scenario. In the calibration event, the two major rivers suffered a substantial degree of overbank flooding, and a maximum water depth of 5.90 m was recorded at the Saxon's Lode gauge near Tewkesbury. The event of 23 November–4 December 2012 recorded a maximum water depth of 5.21 m at the Saxon's Lode. Also both the Severn and Avon were in flood in this event. Tewkesbury lies at the confluence of the Severn, flowing from the Northwest, and the Avon, flowing from the Northeast.

2.2. Rainfall-runoff model and inflow generation

In the experimental setup we emulated a real forecast scenario, in which the precipitation data came from a network of tipping-

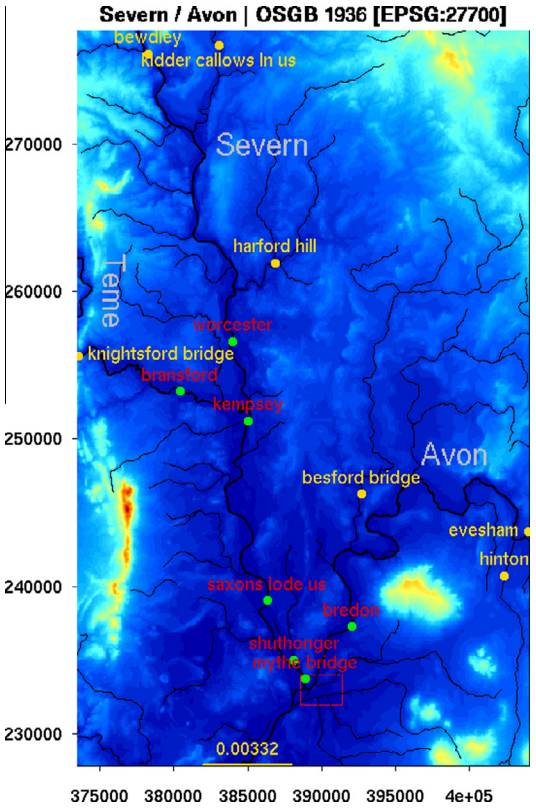

Fig. 1. Flood model domain. OSGB 1936 British National Grid projection; coordinates in meters. Gray labels indicate the three larger rivers (thick black lines). The red polygon surrounds the Tewkesbury urban area. Orange labels/dots refer to the 7 inflow boundary conditions, some of them on smaller tributaries (thin black lines). The yellow line to the South indicates a free-surface boundary condition, with the label indicating the *prior* mean bed slope. Red labels/green dots show locations with available stage observations, just used for validation in the forecast mode. The background is the 75 m resolution DEM used for the model, obtained by upscaling the 5 m NEXTMAP British digital terrain model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bucket gauges sparsely distributed in a 190 km W–E × 120 km S–N rectangular area covering 7 catchments discharging as inflow boundary conditions into the flood domain (i.e., we did not use a Numerical Weather Prediction model). In both the calibration and the forecast scenarios we estimated catchment-average precipitation in each of the seven catchments by Ordinary Kriging (OK) (Deutsch and Journel, 1998) the tipping bucket data over an 1 km resolution grid, at hourly timesteps, and integrating the distributed rainfall maps on the corresponding catchment areas. We treated the OK quantitative precipitation estimates as certain, so that all uncertainty in the discharge ensemble came from the parameters in the hydrologic models.

We used a lumped catchment-scale rainfall-runoff hydrologic model for each catchment, the Hydrologic Simulation Program-Fortran (HSPF) model (Donigian et al., 1995). HSPF is a US EPA program for simulation of watershed hydrology, with a long history and that is still the subject of active research (e.g., Ryu, 2009; Schultz et al., 2013; Kim and Ryu, 2014). Each of the seven major catchments used an independently calibrated HSPF model. Remaining minor point flows and lateral discharge were neglected in both the calibration and forecast stages.

2.3. Hydrodynamic model

We used the flood simulation model LISFLOOD-FP, a coupled 1D/2D hydraulic model based on a raster grid (Bates and De Roo, 2000) with 75 m grid-spacing (664 × 408 pixels). After each assimilation step, the model is re-initialized with the updated state vector (water stage). LISFLOOD-FP has several formulations. Here, we apply the so-called "sub-grid" approach described by Neal et al. (2012), which uses a finite difference numerical scheme adapted from the reach scale inundation model of Bates et al. (2010), and utilises gridded river network data, assuming a rectangular channel geometry. García-Pintado et al. (2013) used the same domain, and provide some more details and a previous application of this model in a synthetic scenario.

2.4. Satellite observations

Satellite SAR observations of the event were acquired by the COSMO-SkyMed (CSK) constellation. The 4-satellite polar orbiting C-band CSK constellation was tasked by the authors. A sequence of 7 CSK Stripmap images giving good synoptic views of the flooding was acquired on a roughly daily basis covering the period 27 November–4 December 2012 (Fig. 2). Unfortunately the rising edge of the hydrograph was not imaged, though the first image in the sequence was acquired just before the flood peak in the Severn (see Fig. 3). Although the river went back in bank on 30 November, we continued the imaging as a substantial amount of water remained on the floodplain. It is worth noting that water levels on the floodplain at the end of the event were much higher than those in the channel. All CSK images were HH polarization, providing good discrimination between flooded and non-flooded regions. Details of the overpasses are given in Table 1. The look direction can be determined knowing the SAR is right-looking and the pass direction (descending/ascending). The latency (time from the CSK SAR image acquisition to the product availability at the User Ground Segment [UGS] plus the delivery time to us) was around 15 h as an average.

Processing level was 1C-GEC, which meant that the images were geo-corrected to ~100 m. It was necessary to register the images to British National Grid coordinates using ground control points and a digital map, when a registration accuracy of better than 2 pixels (of size 2.5 m) was obtained.

In the absence of significant surface water turbulence due to wind, rain or currents, flood water generally appears dark in a SAR image because the water acts as a specular reflector, reflecting backscatter away from the satellite. Detection of the flood extent in each image was performed using the segmentation technique described in Mason et al. (2012a), which groups the very large numbers of pixels in the scene into homogeneous regions. As there was no flooding of urban areas, only the rural flood detection algorithm was used. The scale parameters for the segmentation were the same as those used in Mason et al. (2012a), and also in a number of SAR images of other floods around the world, from several different high resolution SAR sensors. A critical step is the automatic determination of a threshold on the region mean SAR backscatter, such that regions having mean backscatter below the threshold are classified as flooded, and others as un-flooded. The threshold determined was checked manually and corrected if necessary.

The initial rural flood classification was improved by refining it in a number of ways. For example, emergent vegetation adjacent to the flood such as hedgerows may produce a high rather than low SAR backscatter even though they are flooded. This is due to double scattering, whereby radar rays transmitted from the sensor to the water are reflected first to the hedgerow then back to the sensor (or vice versa). Regions of high backscatter that are long, thin, fairly straight and adjacent to flooding were thus reclassified as flooded. The backscatter threshold was also raised to include in the flood category regions of flooding adjacent to the flood class that had slightly higher mean backscatter than the original threshold (e.g., due to wind ruffling the water surface in more exposed parts of

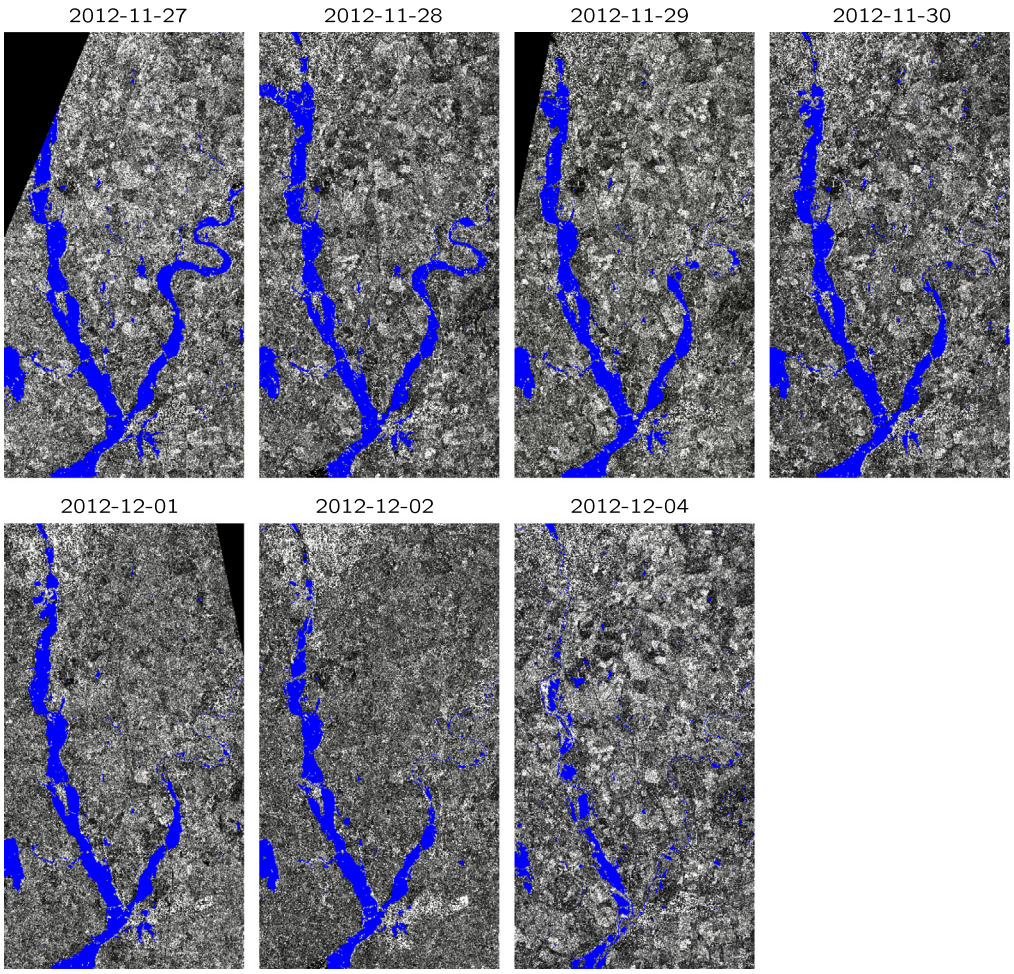

Fig. 2. Flood extents (blue) for the forecast event (November 2012), overlain on SAR in flood model domain. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

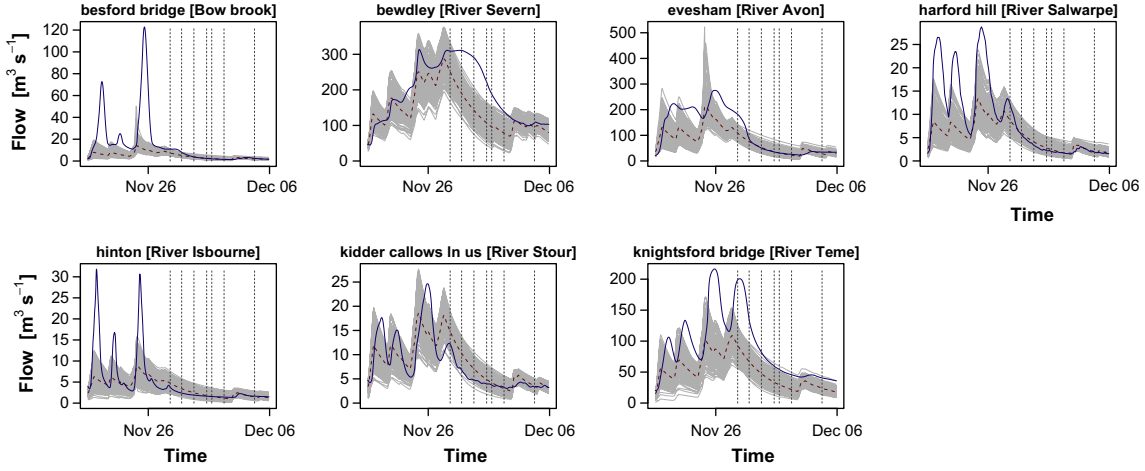

Fig. 3. Inflows into the flood domain for the forecast event, in November 2012. As a reference, blue lines are inflows as measured by standard gauges (not used as data input here). Gray lines are the 150-member forecast ensemble from the hydrologic models, used as input by the flood model. Dashed red lines are the ensemble means. Vertical dashed lines show COSMO-SkyMed overpass times. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the floodplain). Flood regions were also deleted if their mean height was above 14 m above ordnance datum (AOD).

The accuracy of the flood inundation maps in rural areas using the algorithm has been determined to be about 90% (Mason et al., 2012a). This is similar to the accuracies achieved by other researchers (e.g., Martinis et al., 2011). Note that the more accurate a flood inundation map is, the more accurate will be the WLOs derived from it.

Heights were obtained from an image constructed from 24 2×2 km UK Environment Agency (EA) LiDAR tiles covering the hydrodynamic model domain (Fig. 1). Fig. 2 shows the flood extents detected in the images, overlain on the SAR data in the

Table 1
Details of COSMO-SkyMed overpasses.

Timestamp (UTC)	Pass	Incidence angle (°)
27/11/12 19:20	Descending	49
28/11/12 18:01	Descending	51
29/11/12 18:20	Descending	32
30/11/12 19:32	Descending	53
01/12/12 05:38	Ascending	26
02/12/12 05:56	Ascending	48
04/12/12 18:14	Descending	40

hydrodynamic model domain. The sequence shows the flood wave moving down the river, and the flood at Tewkesbury gradually dying away, starting on the Avon. In general terms, regarding the spatial coverage of the images, the Severn was imaged up to the Latitude of Worcester, the Teme up to the Longitude of Bransford, and the Avon up to the Longitude of Besford Bridge. Also, the first image (2012-11-27) just covered up to ~2 km downstream from Kempsey.

WLOs were extracted from the flood extent waterlines by intersecting the extents with high resolution floodplain topography (airborne LiDAR of 1 m or 2 m pixel size) using the method described in Mason et al. (2012b). The method selects candidate waterline points in flooded rural areas having low slope and vegetation, so that small errors in waterline position have little effect on waterline level. The waterline levels and positions are corrected for the effects of double reflections between the water surface and emergent vegetation at the flood edge, wherever possible. At each pixel on the flood edge, the direction perpendicular to the edge moving away from the flood is calculated. A transect of backscatter values is constructed along this direction, traversing from inside the flood, across the waterline and across the region in which emergent vegetation might be expected. Along this transect, backscatter generally increases from low to high then decreases to a stable value between these where the vegetation emerges from the water. If found, the point where this occurs is taken as the waterline position, and the height of this point is the corrected water level. Further details are given in Mason et al. (2012b). The resulting points were not thinned to reduce spatial autocorrelation at this stage as in Mason et al. (2012b), this step being held over until the Quality Control (QC) stage of the assimilation is done.

The standard deviation for the thinned set of WLOs was 0.25 m (Mason et al., 2012b). This value was estimated by comparing waterline levels to those from nearby gauges, and also checking waterline positions against those determined manually from contemporaneous aerial photography. Satellite and aerial photography used to extract WLOs for the calibration event of July 2007 have been described previously in Mason et al. (2010b). The distribution of the WLOs in space and time was consistent with evolution of the flood.

2.5. River cross sections and bathymetry estimation

The EA provided field surveyed river cross-sections for the study area, with a total of 35 along the River Severn and 127 along the River Avon. The available cross-sections had a complete coverage and were evenly distributed along the Avon within the study area. However, the cross-sections for the Severn, while evenly distributed, just covered the transect from the junction with the Avon (near Mythe Bridge) to ~10 km upstream from the Saxons Lode US gauge.

We transformed the surveyed cross-sections into rectangular equivalents, as required by the chosen flood model to simulate channel bathymetry, by

$$\arg\min_{w_r,d_r}\left[\epsilon_w^{-1}\frac{1}{2}(w_c-w_r)^2+\epsilon_d^{-1}\frac{1}{2}(d_c-d_r)^2\right], \tag{1}$$

subject to the constraint $w_r d_r = A_c$, where A_c is the area of the surveyed cross-sectional area (precalculated by integration); w_c and d_c are the maximum of the (within bank) surveyed channel width and depth, respectively; w_r and d_r are width and depth of the rectangular equivalent cross-section; and ϵ_w and ϵ_d are the error variances in w_c and d_c, respectively, assumed proportional to the channel dimensions as

$$\epsilon_w = \frac{w_c}{w_c + d_c}, \tag{2}$$

$$\epsilon_d = \frac{d_c}{w_c + d_c}. \tag{3}$$

With the given constraint ($w_r d_r = A_c$), the substitution $d_r = A_c/w_r$ can be done in (1), which becomes a minimization of a single variable (w_r), which we solved by Newton–Raphson iteration. Then, for both the Severn and the Avon, we interpolated w_r and d_r along the river chainages. For the Severn, further than the area covered by the cross sections (i.e., downstream from Mythe Bridge and from ~10 km downstream from Kempsey up to Bewdley) we extrapolated the width and depth of the cross-sections at the corresponding extremes. For the other rivers in the network we assumed a constant rectangular-equivalent width and depth, with widths obtained from previous studies, and depths obtained from the power law relationship $d = \lambda w^{\gamma}$ between the channel width (w) and depth (d), where we used the parameters $\lambda = 0.30$, and $\gamma = 0.78$ [further details about these may be found in García-Pintado et al. (2013)]. Overall, the aim of model bathymetry based on simple rectangular (spatially-variable) cross-sections is to speed up processing times while attempting to simulate the behavior that higher resolution fully detailed channel cross-sections would have. Our *prior* rectangular bathymetry estimates are assumed subject to a reasonable degree of uncertainty, and local biases are plausible. Hence, we decided to conduct the experiment with these *prior* estimates, and test the forecast under the filter configurations described in Section 2.10 to evaluate how the filter could cope with this potential problem in a real-time flood forecast situation. This decision, in contrast to attempting a previous simulation-based off-line calibration of bathymetry, is also the motivation for recent studies concerning bathymetry estimation (e.g., Durand et al., 2014), which acknowledge that detailed river bathymetry is unknown across many rivers in the world.

2.6. The ensemble filter

We conducted synchronous assimilation of the observations. That is, the flood model simulations were sequentially interrupted as assimilation was conducted at the time of the corresponding CSK overpasses. Whenever we simultaneously estimated uncertain friction, bathymetry, or errors in inflow boundary conditions at the time of the assimilation, we did so as part of the data assimilation by using state space augmentation (Friedland, 1969). From the point of view of the assimilation, here we refer generically to these three as "parameters", with which the state vector is augmented. As the model state is augmented with model parameters, the assimilation scheme is able to take into account correlations between the errors in the parameters and the errors in the model variables. In data assimilation schemes using such an approach, the analysis updates an augmented state vector $\boldsymbol{\psi} \in \Re^n$,

$$\boldsymbol{\psi} = \begin{bmatrix} \mathbf{z} \\ \boldsymbol{\beta} \end{bmatrix}, \tag{4}$$

where $\mathbf{z}$ is the n_s-dimensional model state (water levels in our case) and $\boldsymbol{\beta}$ is a generic n_β-dimensional vector of parameters (including instantaneous inflow errors for the cases where these are simultaneously estimated). Thus, $n = n_s + n_\beta$.

The Ensemble Kalman Filter (EnKF), introduced by Evensen (1994), is characterized by a two step feedback loop: a prediction and an observation-based state update. In each step, an ensemble of augmented state vectors is interpreted as a statistical sample of the forecast or analysis uncertainty, respectively. Thus if $\{\boldsymbol{\psi}_i\}(i=1,\ldots,m)$ is an m-member ensemble, then the ensemble mean is the n-vector defined by

$$\overline{\boldsymbol{\psi}} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\psi}_i. \tag{5}$$

The ensemble perturbation matrix for the augmented state is the $n \times m$ matrix with columns defined by the ensemble perturbations from the mean as

$$\boldsymbol{\Psi} = \frac{1}{\sqrt{m-1}}\left[\boldsymbol{\psi}_1 - \overline{\boldsymbol{\psi}} | \boldsymbol{\psi}_2 - \overline{\boldsymbol{\psi}} | \ldots | \boldsymbol{\psi}_m - \overline{\boldsymbol{\psi}}\right]. \tag{6}$$

Then the ensemble error covariance matrix is given by

$$\mathbf{P} = \boldsymbol{\Psi}\boldsymbol{\Psi}^{\mathrm{T}} = \frac{1}{m-1}\sum_{i=1}^{m}(\boldsymbol{\psi}_i - \overline{\boldsymbol{\psi}})(\boldsymbol{\psi}_i - \overline{\boldsymbol{\psi}})^{\mathrm{T}}. \tag{7}$$

In the prediction phase, each individual ensemble member is evolved forward in time by the forecast model until the time of an observation. In our case this means that the model states (water-levels) are forecast by the hydrodynamic model with appropriate forecast boundary conditions. The dynamical model for friction and bathymetry is that they are constant in time. The treatment of these two and the inflows is described in more detail in Section 2.10.

At the time of an observation, an ensemble approximation of the Kalman filter equations (Kalman, 1960) to used to update the ensemble. The update of the ensemble mean is chosen to satisfy the following constraint:

$$\overline{\boldsymbol{\psi}^a} = \overline{\boldsymbol{\psi}^f} + \mathbf{K}\left(\mathbf{y} - \mathbf{H}\overline{\boldsymbol{\psi}^f}\right), \tag{8}$$

where the forecast (prior) and analysis (posterior) quantities are denoted by the superscripts f and a, respectively. The vector of observations is given by $\mathbf{y} \in \Re^p$. Note that the observations may be indirect or not located at model grid points, so the $p \times n$ matrix $\mathbf{H}$, known as the observation operator (or "forward" operator) is required to map the state vector to the observation space. The Kalman gain,

$$\mathbf{K} = \boldsymbol{\Psi}^f(\mathbf{H}\boldsymbol{\Psi}^f)^{\mathrm{T}}\left(\mathbf{H}\boldsymbol{\Psi}^f(\mathbf{H}\boldsymbol{\Psi}^f)^{\mathrm{T}} + \mathbf{R}\right)^{-1}, \tag{9}$$

is an $n \times p$ matrix, where the superscript "T" denotes matrix transposition, and $\mathbf{R}$ is the $p \times p$ observation error covariance matrix. This update may be thought of as a linear combination of the forecast and the observations, weighted by the uncertainty in the given model and observation data. The term $\delta\mathbf{y} = \mathbf{y} - \mathbf{H}\boldsymbol{\psi}^f$ is usually called the vector of "innovations", indicating the difference between the observations and the forecast; the Kalman gain, $\mathbf{K}$, contains the weights given to the innovations to update the system; $\delta\boldsymbol{\psi} = \overline{\boldsymbol{\psi}^a} - \overline{\boldsymbol{\psi}^f}$, is called the vector of "increments" and is the difference between the analysis and the forecast.

As well as updating the ensemble mean, we must also update the ensemble perturbations, giving an ensemble estimate of the analysis error covariance as

$$\mathbf{P}^a = \left(\boldsymbol{\Psi}^f - \mathbf{K}\mathbf{H}\boldsymbol{\Psi}^f\right)\left(\boldsymbol{\Psi}^f\right)^{\mathrm{T}}. \tag{10}$$

There are a number of possible computational implementations for updating the ensemble. In this work, we used an Ensemble Transform Kalman Filter (ETKF) in an unbiased formulation with a symmetric square-root (Ott et al., 2004; Wang et al., 2004; Hunt et al., 2007; Livings et al., 2008). Regarding notation in this study, being $\mathbf{A}$ a generic matrix, $\mathbf{A}_{i,:}$ means the ith row of $\mathbf{A}$, and $\mathbf{A}_{i,j}$ is the element at the ith row and the jth column.

2.7. Filter localization

Ensemble size is an important issue in any ensemble Kalman filter. We can see from Eqs. 8 and 9 that the increments lie in the subspace spanned by the ensemble perturbations. Another way to express this is that the increments are linear combinations of the columns of the forecast error covariance matrix. Typically the number of ensemble members is much smaller than the dimension of the state vector, leading to under-sampling. This often manifests itself as spurious (unphysical) forecast error correlations and for example, increments updating the state at locations further from the location of the observation than is plausible. Localization techniques are often used to ameliorate the problem.

There are a large number of variations on the localization technique with the principal difference being whether localization is applied directly to the forecast error covariance matrix (known as covariance localization, Houtekamer and Mitchell, 1998, 2001) or by a more indirect method known as domain localization (Ott et al., 2004; Hunt et al., 2007). In the latter, the assimilation is applied to independently to a series of disjoint local domains in physical space. For each local assimilation, only observations within some defined cut-off radius are considered. In this paper we use observation localization (OL) (Hunt et al., 2007; Nerger and Gregg, 2007), which allows us to obtain a similar effect to covariance localization while using domain localization. In OL, we construct a taper matrix from chosen correlation functions of compact support. The Schur or Hadamard (elementwise) product of the taper matrix and the inverse of the observation-error covariance matrix corresponding to a local analysis domain is computed. Thus, the weight of observations is reduced as a function of their distance from the local analysis domain by increasing their assumed error variance. Here we used a fifth-order polynomial (Gaspari and Cohn, 1999, Eq. (4.10)) for weighting the observations. The study of localization techniques and parameters is an active area of research (e.g., Kirchgessner et al., 2014). Our new contribution is to develop a novel distance metric based on a channel network distance, which allows us to distinguish between flows in adjacent channels that may be close together in a Euclidean sense.

For floods developed around a channel network (e.g., a river network, as in this case study), one could expect the physical connectivity of flows to influence the development of the forecast error covariance. Thus a localization taking into account the along-network distance would not only be more physically meaningful than an "as-the-crow-flies"-based localization, but also should lead to an improved forecast skill. To this end, assuming that the flood is developed around a pre-existing (river/urban) network, the channel network can be vectorised and the chainage of the network used for calculating along-network distances. Let X be the set of points of interest for localization, and we denote as X^c the minimum-distance mapping of X into the channel network. With this, let us define our along-network metric for localization, $d_n(x_i, x_j)$, as

$$d_n(x_i, x_j) = \max\left\{d_e(x_i, x_j), d_s(x_i^c, x_j^c)\right\}, \tag{11}$$

where $d_e(\cdot,\cdot)$ denotes the "as-the-crow-flies" Euclidean distance, which is evaluated upon X, and $d_s(\cdot,\cdot)$ is the distance evaluated along the chainage provided by the vectorised channel network, which is evaluated upon X^c. The rationale for including $d_e(\cdot,\cdot)$ in the definition of $d_n(\cdot,\cdot)$ is to provide a minimum distance threshold for nearby couple of points (x_i, x_j), which might for example fall at both sides of a flooded channel, and whose d_s distance is neglected

in their corresponding projection (x_i^c, x_j^c). One might argue that this can artificially decrease the covariance at some points near junctions. However, this should be a minor effect, and also one might well argue that it is sensible to use this metric at those points, just because they are in junctions, influenced by flows with different dynamics. The output of this metric is then used by Eq. (4.10) of Gaspari and Cohn (1999) in the same way that Euclidean distances would be. Note that this is not a translation invariant metric.

2.8. Ensemble inflation

Another mechanism to preserve the ensemble variance and help preventing ensemble collapse is the inflation of the ensemble perturbations (Anderson, 2001). We used a simple automatic multiplicative inflation approach applied to the posterior perturbations. For this, we obtained an inflated ensemble $\mathbf{\Psi}^{a,+}$ by multiplying each row in the pre-inflated analysed ensemble matrix $\mathbf{\Psi}^{a,-}$ times the corresponding element of an inflation vector $\boldsymbol{\lambda} \in \Re^n$ (i.e., $\mathbf{\Psi}^{a,+}_{i,:} = \lambda_i \mathbf{\Psi}^{a,-}_{i,:}$), where for any specific variable in the state vector the inflation factor λ_i was obtained as

$$\lambda_i = (1 - \alpha_p)\frac{\sigma_i^f}{\sigma_i^a} + \alpha_p, \tag{12}$$

where α_p is an input inflation parameter, with $\alpha_p \in [0, 1]$, and σ_i^f and σ_i^a are the background and updated sample (ensemble) standard deviations, respectively, for the assimilation step. Thus, $\alpha_p = 0$ would recover the variance prior to the assimilation, and $\alpha_p = 1$ would not apply any inflation. As indicated in Table 2, α_p can hold different values for the various types of elements in the augmented state vector.

2.9. Observation quality control

Online data assimilation techniques such as ensemble Kalman filters and particle filters tend to lose accuracy dramatically when presented with an unlikely observation. Such an observation may be caused by an unusually large measurement error or reflect a rare fluctuation in the dynamics of the system. In general, while the observations with large measurement errors should be screened out, those of existing rare fluctuations should be assimilated. For example, in the case of a flood event in a river network, one may consider the, unusual, failure of a small earth dam in an agricultural plot or terrace, with the purpose for impounding water. These retention structures can intercept important amount of water in agricultural plots, and a sudden failure would release an additional, relatively small, wave to river network. The anomaly would be higher and more local just after the failure, with the wave later diffusing into the general flood. In any case, Kalman filters and particle filters tend to fail when the system undergoes occasional transitions revealed by an observation that is inconsistent with the predictive distribution of the systems state (Vanden-Eijnden and Weare, 2012). Thus we follow here the practical approach of applying and outlier analysis to remove any unlikely observation not accounted for in the model dynamics. This is applicable as floods are a case of data assimilation in a low noise regime and observational noise of SAR-based WLOs is normally small. An alternative strategy could be to use specific assimilation schemes based on rare event simulation tools, as proposed by Vanden-Eijnden and Weare (2012).

Accordingly, we conducted an online Quality Control (QC) of the observations as follows, based on the innovation values. The innovations (see Section 2.6) are obtained for the complete set of WLOs (see Section 2.4). An experimental semivariogram is then obtained with the innovations, and the range of a corresponding adjusted exponential variogram model is used as the cut-off radius of a spatially moving window, which is sequentially centered at the location of each WLO to evaluate if its corresponding innovation is an outlier according to its neighborhood. An innovation is then considered an outlier if its value is out of the interval with 95% confidence level obtained from the population of innovations in the local window. Once the outliers are rejected according to the QC protocol, a thinning analysis is conducted on the surviving WLOs, as described by Mason et al. (2012b). The thinned dataset is finally used by the DA analysis for its assimilation. Our previous tests (not shown) indicated that QC of the observations, as for example the one included here, is a necessary step to avoid rare fluctuations, as indicated above. QC is a standard part of operational assimilation-forecast systems in other applications such as numerical weather prediction (e.g., Kalnay, 2002, Section 5.8). In our SAR-based WLOs, no significant bias between the SAR-derived and gauge levels could be detected (Mason et al., 2012b). For alternative processing chains, if WLO bias was found significant, and observation bias-aware DA scheme could potentially deal with this problem (e.g., Dee, 2005; Pauwels et al., 2013).

Table 2
Summary of filter configurations for assimilation.[a]

code	*h*			*q*			*dsl*			η_c			*bat*		
	u[b]	i[c]	l[d]	u	i	l	u	i	l	u	i	l	u	i	l
a	T	0.5	F	F	–	–	T	0.0	F	F	–	–	F	–	–
b	T	0.5	F	T	0.0	F	T	0.0	F	F	–	–	F	–	–
c	T	0.5	$d_e(20)$	F	–	–	T	0.0	$d_e(10)$	F	–	–	F	–	–
d	T	0.5	$d_e(20)$	T	0.0	$d_e(20)$	T	0.0	$d_e(10)$	F	–	–	F	–	–
e	T	0.5	$d_n(20)$	F	–	–	T	0.0	$d_n(10)$	F	–	–	F	–	–
f	T	0.5	$d_n(20)$	T	0.0	$d_n(20)$	T	0.0	$d_n(10)$	F	–	–	F	–	–
g	T	0.5	$d_n(20)$	F	–	–	T	0.0	$d_n(10)$	T	1.0	F	F	–	–
h	T	0.5	$d_n(20)$	F	–	–	T	0.0	$d_n(10)$	F	–	–	T	1.0	$d_n(10)$
i	T	0.5	$d_n(20)$	F	–	–	T	0.0	$d_n(10)$	T	1.0	F	T	1.0	$d_n(10)$
j	T	0.5	$d_n(20)$	T	0.0	$d_n(20)$	T	0.0	$d_n(10)$	T	1.0	F	F	–	–
k	T	0.5	$d_n(20)$	T	0.0	$d_n(20)$	T	0.0	$d_n(10)$	F	–	–	T	1.0	$d_n(10)$
l	T	0.5	$d_n(20)$	T	0.0	$d_n(20)$	T	0.0	$d_n(10)$	T	1.0	F	T	1.0	$d_n(10)$
m	T	0.5	$d_n(20)$	T	0.0	$d_n(20)$	T	0.0	$d_n(10)$	T	1.0	F	T	1.0	$d_n(5)$

[a] *code* is filter configuration, *h* is distributed water level; *q* are the 7 inflow boundary conditions; *dsl* is the slope of the downstream free surface boundary condition; η_c refers to the 2 global Manning coefficients for channels; *bat* refers to the distributed bathymetry. T is TRUE (applied). F is FALSE (not applied). "–" if meaningless.
[b] u refers to whether the variable/parameter is being updated in the assimilation.
[c] i refers to the inflation parameter α_p (see Section 2.8).
[d] l is FALSE for no localization, $d_e(r)$ for isotropic "as-the-crow-flies" distance, and $d_n(r)$ for along-network distance, being r the localization radius in km (see Section 2.7).

2.10. Calibration and experimental design

We used a previous event in July 2007 in the same location as a calibration scenario. With precipitation and potential evapotranspiration as input data and flow at the outlet of each catchment as calibration data (yellow named points in Fig. 1), we calibrated the HSPF lumped catchment-scale rainfall-runoff hydrologic model for each of the 7 catchments. Calibration was done with a simple Monte Carlo run, sampling from prior uniform distributions with common ranges as indicated in the HSPF model documentation. From an initial set of 500 samples, we selected the 150 best set of parameters for each catchment according to the Nash coefficient. Off-line, we used the calibrated ensemble (size $m = 150$) discharge of the hydrologic models as input into the flood model domain, and calibrated the latter using a similar approach that employed available time-series of water levels at a number of gauges as calibration data (green points with red labels in Fig. 1). For model spin-up, in both the calibration event and the evaluation one, the rainfall-runoff simulations started 1 month before the 2-D flood simulations. As indicated, the ensemble size was $m = 150$.

With the cascaded calibrated hydrologic-flood models we conducted the assimilation with a number of filter configurations for the November 2012 event in forecast mode. Table 2 summarizes the assimilation conditions for each filter configuration. The first six configurations are focused on evaluating the effect of localization and inflow estimation on the assimilation and forecast. These configurations have fixed values for both friction and bathymetry: (a) a global ETKF, (b) as (a) with inflow error estimation and correction, (c) a local ETKF (LETKF), (d) as (c) with inflow error estimation and correction, and (e) and (f), which are respectively as (c) and (d) but with an along-network metric for localization (described in 2.7). Another six configurations support the evaluation of the effect of friction and bathymetry estimation, and all of them use an along-network metric for localization: (g), (h) and (i) do not estimate and correct inflow errors, so that (g) estimates friction, (h) estimates bathymetry, and (i) estimates both friction and bathymetry. Then, (j), (k) and (l), are respectively as (g), (h) and (i) but with simultaneous inflow error estimation and correction. A further configuration m) is as (l) but with a smaller localization radius for bathymetry, to provide a minimal analysis of the influence of the localization radius on the distributed bathymetry estimation and the general estimation process.

Regarding the parameter spatial support, friction was considered as two global parameters described by the Manning's coefficient: one scalar value for the three larger rivers (Severn, Avon and Teme) [with *prior* $\eta_{c1} \sim \mathcal{N}(0.035, 1.00e-06)$], and another scalar value for the other smaller rivers in the domain [with *prior* $\eta_{c2} \sim \mathcal{N}(0.040, 1.00e-06)$]. This is a parsimonious friction parameter set, while the priors still acknowledge that smaller rivers tend to have more vegetated riverbanks, leading to higher friction in high flow conditions. The prior means seem reasonable given our previous experience on the study area (e.g., García-Pintado et al., 2013). Bathymetry was considered as local and spatially distributed along the channel network. That is, bathymetry was just uncertain for channels and every channel pixel had an independently updated bathymetry, without any further constraint. The distinction between "local" and "global" here refers to whether the parameters or variables to be estimated may be assigned a point location, and so be directly subject to localization in the filter.

Several methods have been proposed for correcting the errors in river discharge predictions, as these degrade the quality of the downstream flood forecasts (e.g., Bogner and Pappenberger, 2011). For example, Andreadis et al. (2007) used a first order autoregressive error forecast model for a 3-month case synthetic experiment in the context of discharge estimation based on assimilating EO-based water levels. However, for the context of storm-flood events (our focus here), the sparse satellite overpasses in relation with the duration of the event make it impractical (if possible) to estimate the parameters of error models more complex than a simple stationary model (i.e., to assume a constant bias between satellite overpasses). Thus whenever the inflow errors were estimated and updated (in the seven inflow boundary conditions), we assumed a constant bias forecast error model for the inflows, as previous studies in the field (Matgen et al., 2010; García-Pintado et al., 2013). In addition, in all the filter configurations, we included simultaneous estimation of the downstream free surface slope boundary condition, as a local parameter where the *prior* distribution was $dsl \sim \mathcal{N}(3.32e-03, 2.25e-10)$ (see mean as yellow label in Fig. 1), the mean value being obtained from the calibration for the July 2007 event, and the variance representing the reasonable degree of uncertainty we had on the *prior* mean. See Section 2.5 for a description of the *prior* estimates of bathymetry, and rationale for the distributed online updating approach.

In the filter configurations with parameter updating [(c), and (d)], we added a spatially correlated error to the *prior* bathymetry. For this, we selected an unconditional Gaussian error simulation to generate the ensemble of 2D random error fields, which we masked with the channel network. The selection of the parameters for the Gaussian simulation, and the assumption that all errors come from bathymetry (i.e. perfect hydraulic structures) is somehow arbitrary. Specifically, we assumed the bathymetry errors ϵ_b were proportional to the prior estimates of channel depth, with c.v. = 0.15, and having a spatial correlation length of 5000 m, defined as the range parameter of an exponential variogram model.

2.11. Validation approach

We used gauge water level time series from the network shown in Fig. 1 (red labels). For validation we conducted a visual examination of the ensemble time series along the event, and RMSE as general statistics. All level gauges indicated in Fig. 1 are in the major three rivers, except for Shuthonger, which is in a small tributary, close to its junction with the Severn and likely affected by backwater effects.

3. Results and discussion

3.1. Structure of results and discussion

We structure the results and discussion around three major topics: (i) influence of localization on the system updating and flood forecast, (ii) capability of inflow estimation and its influence on the flood forecast, and (iii) capability of model parameter estimation (friction and bathymetry) and its influence of the flood forecast. This structure is to ease discussion. We will see however that the three topics are strongly inter-related, and we include cross-references as needed.

3.2. Influence of the metric for localization

In this study we propose an along-network metric for localization, based on the understanding that it should be sound for the flood forecast problem in river/channel networks. In accordance with Section 2.7, here we test and discuss the use of a global ETKF vs. filtering schemes with localization (LETKF), with either a standard "as-the-crow-flies" Euclidean distance (d_e) or an along-network distance (d_n). Note we have arbitrarily chosen localization parameters which seem relevant to the problem, so that it is unlikely we have selected the optimum in either filter configuration.

We will use Fig. 4, as a base example for discussion. This depicts a symbolic representation (square symbols) of the elements $\mathbf{K}_{i,:}$ in the Kalman gain for several filter configurations. Here, i is the row corresponding to the updating of the water level at Bredon, in the river Avon catchment. The symbols are plotted at locations corresponding to the jth element of the observation vector, with the WLOs obtained from the 7th CSK overpass. While it is important to remember that the increment is the sum of the matrix–vector product between the Kalman gain and the vector of innovations, a consideration of the individual elements of the Kalman gain provides information about the relative influence of different observation locations on the analysis.

In Fig. 4 the updated forecast error covariance between level at Bredon and levels elsewhere is also shown as the background for each configuration. Note that both the color scales and the symbolic representation of $\mathbf{K}_{i,:}$ are independent for each plot to ease visualization. For this figure we have selected filters with inflow updating, as these tend to behave better than similar ones without inflow estimation (see Section 3.3 below). Also, we focus here on the situation pertaining to the assimilation of WLOs from the last CSK overpass, as this summarizes the cumulative feedback of the different filters along the sequential assimilation in the event.

Filter (b), the global filter, leads to sparsely distributed non-negligible $\mathbf{K}_{i,j}$ values throughout the domain, with many distant observations influencing the updating at Bredon. Not only are there many significant gain values along the whole Severn, but also the highest value ($\max\{\mathbf{K}_{i,:}\} = 0.023$) at this last overpass ($t = 7$) is in a tributary of the Severn, a significant distance away from Bre-

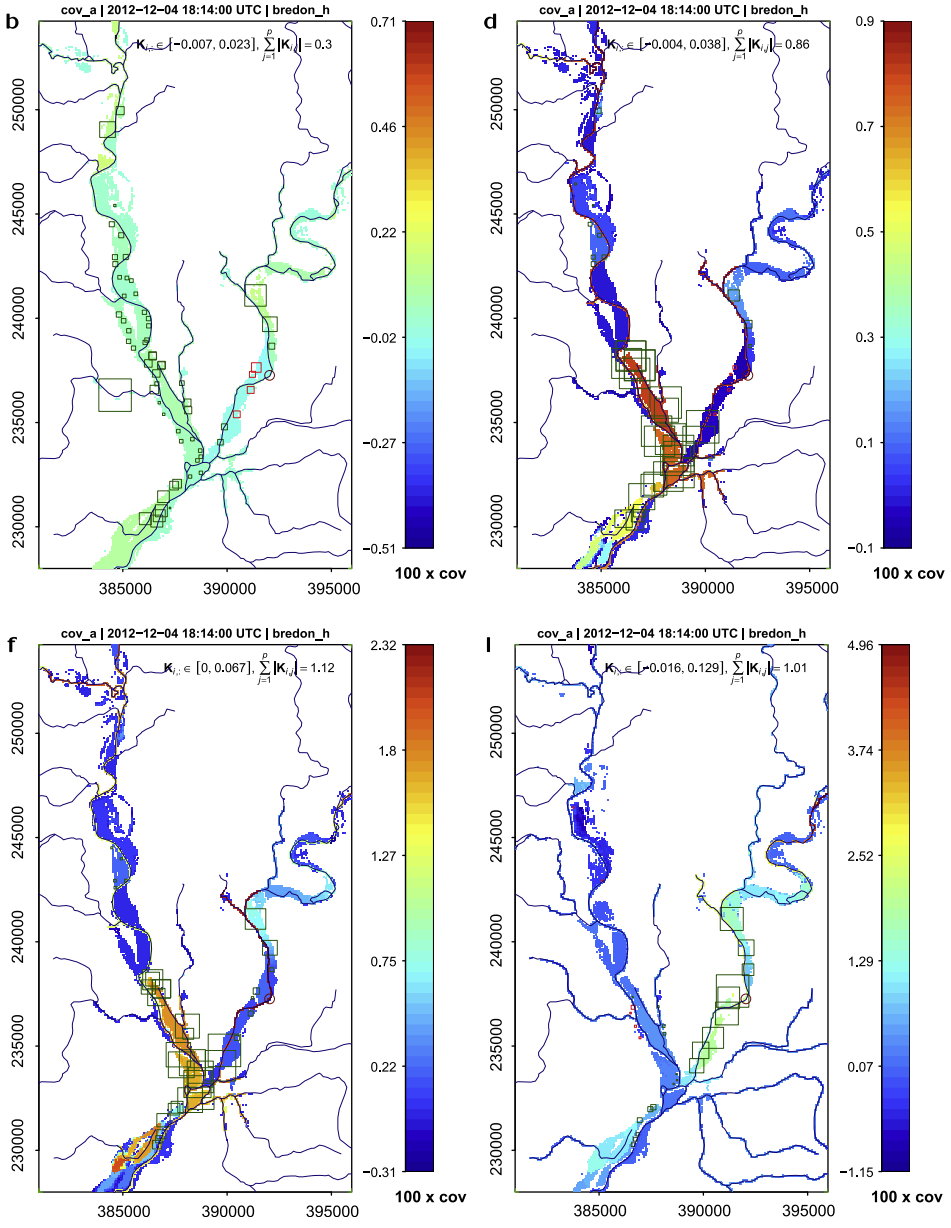

Fig. 4. Updated error covariance between the state variable (water level) at Bredon and the state vector (water level elsewhere) at the last assimilation step (7th CSK overpass). Plot labels (b, d, f, and l) refer to the corresponding filters (see Table 2). The red circle indicates the location of Bredon. The set of squares, with each one centered at the corresponding observation location, is a symbolic representation of $\mathbf{K}_{i,:}$ (being i the state vector index corresponding to water level Bredon) at the corresponding assimilation step and filter. The side length of each square is proportional to the corresponding $\mathbf{K}_{i,j}$ value, where the biggest square in each plot relates to $\max\{\mathbf{K}_{i,:}\}$ (e.g., 0.023 in filter [b]). The sum of the absolute Kalman gain values in the row is indicated by $\sum_{j=1}^{p} |\mathbf{K}_{i,j}|$. Green/red squares are positive/negative gain values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

don. Also, the gain values have a skewed distribution with a just a few WLOs having high gain values and then many other observations gathering around low, albeit non-negligible, values. In general, the situation is far from what one would expect from a properly constructed assimilation system, and it is also likely to hamper the robustness of the filter to possible anomalous (outlying) innovations. This situation arises from the assimilation sequence throughout the event, in which spurious correlations are not properly damped. Thus the global filter ends up with a system in which these spurious correlations have a dominant effect, leading to relationships which are unlikely to happen in the real physical system. Moreover, at this stage in the event the water levels at observation locations surrounding Bredon have a negative correlation with the level at Bredon, leading to negative gain values (red squares), which do not have any physical reason to happen. A strongly-related problem is the collapse of the variance, as the development of spurious correlations leads in turn to too much weight being put onto the observations in the early stages of the event and promotes variance collapse. As the assimilation proceeds, the global filter leads to a general underestimation of the variance and filter divergence, here exemplified via the over-flattened map of the forecast error covariance with Bredon, and its low sum of absolute values of the corresponding Kalman gain values ($\sum_{j=1}^{p}|\mathbf{K}_{i,j}| = 0.3$).

On the other hand, filter (d), with d_e-metric localization, shows a very different situation, with a much higher variance and higher spatial variability in the model error covariance with Bredon water levels. Significant values in the corresponding $\mathbf{K}_{i,:}$ row in are much closer to Bredon, and more observations share a fair contribution to the updating, making the filter more robust to outliers in the observations. Overall, the spurious correlations have been quite effectively filtered out, and this seems a much better situation than that from filter (b). Also the situation that the most influencial gain values are clustered downstream from Bredon has some physical basis, as this situation corresponds to the recession of the flood event, and the bulk of the flood wave has now passed Bredon. Thus the variance of water levels is higher downstream and lower upstream, while the standard deviation of the WLOs is set to 0.25 m everywhere. This naturally maps into higher/lower Kalman gain values for those observations with higher/lower variance in the corresponding background water levels. However, one would still expect higher Kalman gain values to be closer to Bredon, and to have some influence, even if minor, from upstream observations in the updating.

Filter (f), with d_n-metric localization seems an improved version of (d). It does not contain any negative gain values, and there is also a good number of observations with roughly equally high gain values leading to a robust situation with respect to outliers existence in the WLOs. The distant WLOs in the Severn now have lower gain values, and the highest gain values are closer to Bredon. Also WLOs upstream from Bredon now have a relatively higher gain values. Overall, this seems the best filter of the three, among which the only difference is the localization approach.

The influence of simultaneous parameter estimation is discussed in Section 3.4. However, we include filter (l), which is similar to (f) but with simultaneous estimation of global channel friction and distributed bathymetry, to summarize that it leads to a situation which seems even more physically sound than that from (f) from the point of view of the spatial distribution and share of the Kalman gain values. The higher gain values are now very well distributed around Bredon, and with slightly higher weights given to those WLOs downstream from Bredon. The situation seems very close to ideal, with properly developed forecast error covariances, with respect to what one could expect at this stage of the event. There are a few distant minor negative gain values in the row, but given their relative values these become insignificant in the updating. Other aspects of the case are discussed in Section 3.4.

3.3. Inflow estimation

As indicated in Section 2.4, the satellite sequence was not covering the boundary condition locations for the major inflows to the flood model domain. For the three major rivers, the coverage was up to ~20 km downstream from Bewdley (inflow to the Severn), ~10 km downstream from Evesham (inflow to the Avon), and ~8 km downstream from Knightsford bridge (inflow to the Teme). In fact, Besford bridge (inflow to the Bow brook) was the only inflow location within the SAR coverage. Thus this case is an extrapolation situation regarding the online estimation and correction of inflow errors. Table 2 shows a number of filter configurations which performed simultaneous estimation of inflow errors at the assimilation times, and used these error estimates to correct the inflows from the hydrologic models, assuming a constant bias as error forecast model. Table 3 indicates the root mean squared error (RMSE) for these filter configurations, where the RMSE is evaluated against the gauged inflows (blue lines in Fig. 3) for the time between the 1st CSK overpass (2012-11-27 19:20:00 UTC) and about one day after the last overpass (2012-12-05 23:00:00 UTC). In parentheses, Table 3 indicates the increment in the RMSE between the updated inflows, along the event, and the background (the open loop hydrologic forecast) inflow errors. Thus a positive/negative increment in RMSE indicates the updated inflows are further from/closer to the gauged inflows than the background.

The filters, in any case, just calculate the innovations (WLOs minus the forecast levels at the time of the overpass and on WLO locations) and use the background and observation error covariances to map these innovations into increments in the state vector (water levels and, possibly, parameters, depending on the filter configuration). So, positive/negative innovations, along with the generally positive correlations between observation errors and water level errors, are basically stating that more/less water should have entered into the system by the time of the assimilation and correcting this by updating the water levels and inflow errors. In the setup in this study, apart from the 7 inflow boundary conditions, we are not making any provision for lateral inflows along the river network, inflows from smaller tributaries, nor for groundwater infiltration/exfiltration. These unaccounted inflows/outflows in the flood model may well lead to increased/decreased local innovations and the corresponding mapping into the distributed increments in the water level as a result of the assimilation. Also if inflow errors at the time of the assimilation are correlated with water levels at downstream observation points at that time, the abovementioned unaccounted flows may map, through the innovations, into an over- or underestimate of the specific errors at the inflow boundary conditions (over/under-shooting). However, this over/under-shooting is not necessarily a bad thing for the actual (next) forecast step. The DA is just attempting to estimate overall inflow errors and assigning them to the only provision we have made for that, i.e. to the specific inflow boundary conditions. Thus the RMSEs in Table 3 are just an indication of how close the gauged inflows (which, themselves may be subject to biases because of errors or unaccounted hysteresis in the rating curves, etc.) are to the assimilation-based estimates of total inflow to the system.

Summarising Table 3, the assimilation in all filter configurations generally moves the prior inflows away from the gauged inflows. This is indicated by the positive values in parenthesis in the RMS row. As indicated, this should not be interpreted as whether the assimilation is doing a bad job. On the contrary, this may well be the case that the satellite-based WLOs are providing information,

Table 3
RMSE of inflows for filters with inflow updating.[a,b]

	b	d	f	j	k	l
besford_bridge_q	1.84(0.24)	1.94(0.35)	1.79(0.20)	1.74(0.15)	1.82(0.23)	1.75(0.16)
bewdley_q	82.87(−2.95)	122.26(36.44)	108.86(23.04)	107.34(21.52)	97.63(11.81)	95.92(10.10)
evesham_q	33.43(12.79)	32.07(11.43)	23.82(3.18)	23.49(2.85)	23.62(2.98)	23.93(3.30)
harford_hill_q	1.12(0.37)	0.96(0.20)	1.32(0.57)	1.21(0.46)	0.96(0.20)	0.95(0.19)
hinton_q	0.65(0.10)	0.48(−0.07)	0.49(−0.06)	0.49(−0.06)	0.53(−0.02)	0.52(−0.03)
kidder_callows_ln_us_q	1.24(−0.60)	1.20(−0.64)	1.42(−0.42)	1.41(−0.42)	1.50(−0.33)	1.54(−0.30)
knightsford_bridge_q	48.38(4.66)	49.56(5.84)	59.34(15.61)	58.08(14.35)	51.13(7.41)	49.00(5.27)
RMS[c]	38.42(5.27)	51.33(14.61)	47.73(10.59)	46.99(9.84)	42.61(5.39)	41.72(4.49)

[a] [$m^3 s^{-1}$]. RMSE measured against gauged inflows within [2012-11-27 19:20:00 UTC, 2012-12-05 23:00:00 UTC].
[b] In parentheses is the RMSE minus the RMSE of the prior inflows (forecast of the hydrologic models).
[c] RMS of the values for the corresponding column.

to improve the estimation of total inflows into the system, not contained either in the gauged inflow boundary condition nor in the prior inflows (i.e., the forecast from the catchment-scale hydrologic models). Without further information, to evaluate whether the assimilation is then performing well in estimating inflow errors and whether this online inflow error estimation/correction is an useful operational strategy, one needs to evaluate how the flood forecast behaves downstream and whether the estimated inflow errors are properly allocated to the corresponding sources. In the context studied, this refers to a loose relationship between flooded areas and corresponding assumed point inflow boundary conditions.

For example, Table 3 shows two opposed configurations with similar RMS [filter (b) and filter (l)]. Filter (b) is in fact the only configuration which brings the updated inflow from Bewdley closer to the gauged inflows (δRMSE $= -2.95$ $m^3 s^{-1}$). Fig. 5 shows the evolution of the updated inflows at Bewdley for these two filters; i.e. the global filter (b), and the filter (l), with d_n-metric localization and simultaneous friction and bathymetry estimation. Filter (b) behaves rather erratically, in agreement with the discussion about the lack of robustness of the filter in Section 3.2. For example, the assimilation of the WLOs from the 2nd and 6th overpasses creates positive increments, which are interspersed with the negative increments related to the 3rd and the 7th overpasses. On the other hand, filter (l) has a small increment at the 1st overpass, and then onwards, the increments become negligible. To provide some insight into the reasons leading to these different situations, let us focus now on the forecast error covariances after the 1st assimilation step between the flows at Bewdley (inflow to the Severn) and the water levels elsewhere. This is depicted by Fig. 6 for the same filters as Fig. 4. Filter (b) shows a strong component of the updating is due to spurious correlations, not only from smaller tributaries downstream, but also even from a set of negative Kalman gain values assigned to WLOs too distant in the Avon. The evolution of the spatial distribution of the Kalman gain values in filter (b) is highly erratic along the event, with the highest gain values continually displacing from one location to another between sequential assimilation steps (not shown; available on request), and leading to a degenerate situation by the 7th overpass, where a highly skewed distribution of the gain values (in the row) and the growth of spurious correlation with WLOs in smaller tributaries is very similar to that of Fig. 4 for the same filter.

Filter (d), however, adequately takes into account the most upstream observations in the Severn to update the inflows. Still there are non negligible spurious gain values in tributaries downstream. Filters (f) and (l) are both similar to (d) but more effective at damping the spurious correlation with water levels at downstream tributaries. The evolution of the distribution of Kalman gain values in the sequential assimilation is then very similar for (d), (f), and (l) (not shown). For these, the spatial distribution of gain values is much more stable in time, and the filters are effective in removing the spurious correlations. There is still a general trend to put more weight into a few observations over time. However, this is partially due to the fact that less WLOs are available to assimilate as the flood recedes.

Table 4 indicates the RMSE of the water levels through the event evaluated at the seven available water level gauges. Note that it is also possible that the gauges have some bias. According to Table 4, a pairwise comparison of filter configurations with similar configuration but without/with simultaneous inflow estimation indicates that the online inflow updating lead to improved forecasts if localization is applied [e.g., (c) vs. (d), or (e) vs. (f)]. This improvement also applies if friction is simultaneously estimated [(g) vs. (j)], but the statistics are similar for those configurations with simultaneous bathymetry estimation [(h) vs. (k), or (i) vs. (l)].

On the other hand, in the global filters [(a) vs. (b)] the simultaneous inflow updating further promotes ensemble collapse and divergence. This is reflected in the larger RMSE in (b) with respect to (a), and can be seen, e.g., in the water level time series plots for configuration (b) in Worcester in the Severn (Fig. 7), Mythe Bridge downstream in the Severn close to the junction with the Avon (Fig. 8), or Bredon in the Avon (Fig. 9). Thus, just the filters with localization, with improved accounting of the forecast error covariances, are able to better exploit the added freedom of inflow

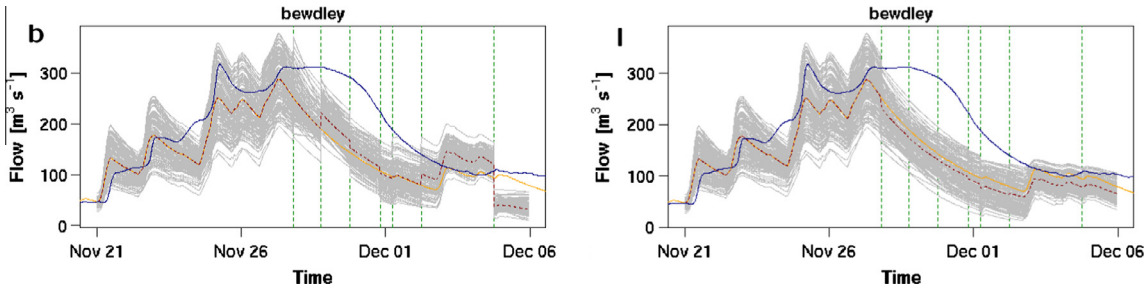

Fig. 5. Inflow estimation at Bewdley for filters with configuration (b) and (l), as specified in Table 2.

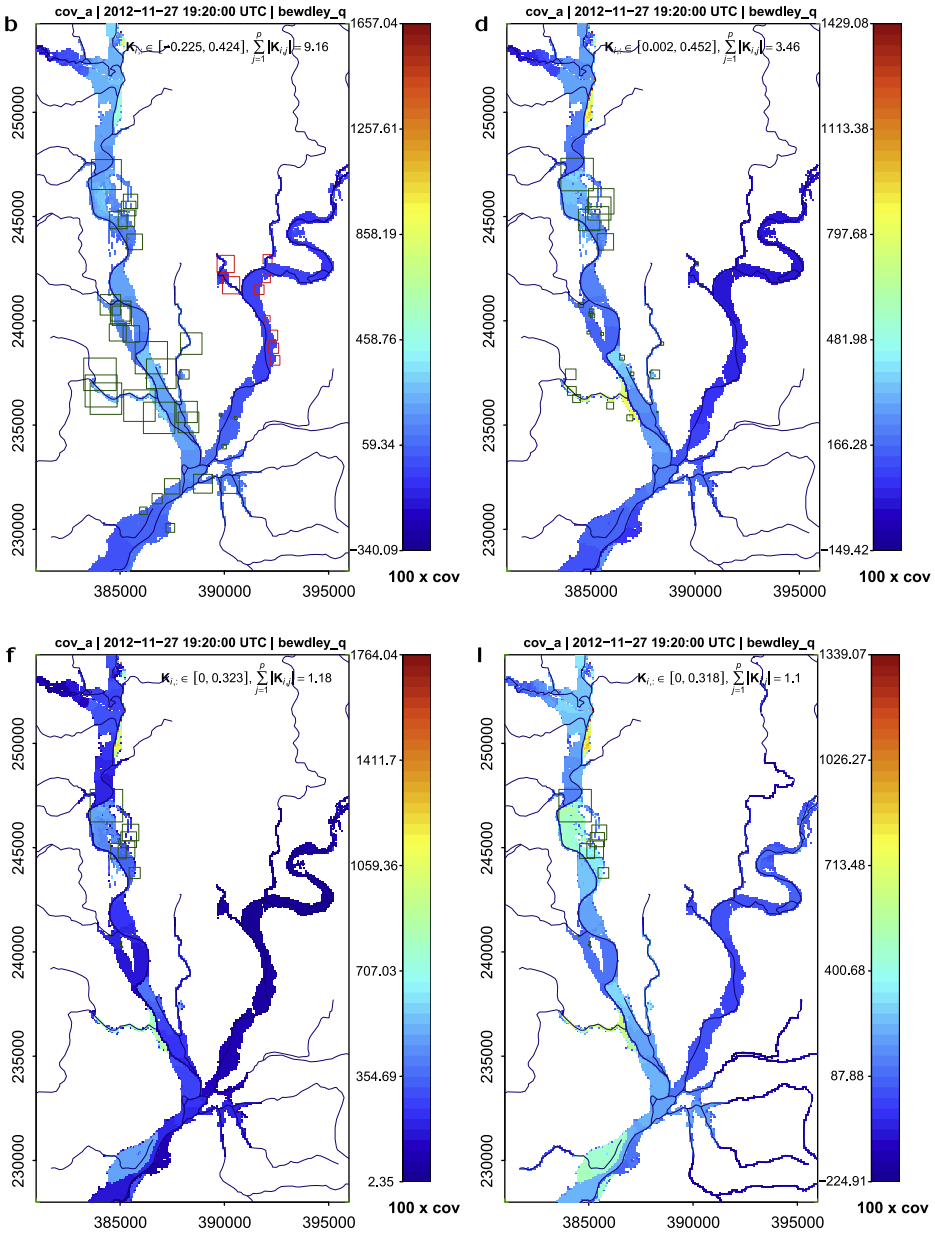

Fig. 6. Updated error covariance between the inflow boundary conditions at Bewdley and the state vector (water level elsewhere) at the first assimilation step (1th CSK overpass) for the same filter than Fig. 4. Plots focus on the satellite coverage area, thus Bewdley location is not shown. Description is as Fig. 4, being now *i* the state vector index corresponding to inflow errors at Bewdley.

Table 4
RMSE of water levels at gauged locations for the filters evaluated.[a]

	a	b	c	d	e	f	g	h	i	j	k	l	m
bransford_h	0.79	0.90	0.80	0.95	0.81	1.34	0.85	1.00	0.98	1.30	1.14	1.18	1.00
bredon_h	0.66	0.65	0.69	0.40	0.69	0.40	0.67	0.85	0.89	0.45	0.74	0.72	0.60
kempsey_h	1.22	1.43	1.26	0.57	1.27	0.60	1.17	1.22	1.28	0.65	1.16	1.18	1.06
mythe_bridge_h	0.69	0.79	0.73	0.50	0.73	0.46	0.72	0.86	0.79	0.51	0.76	0.76	0.65
saxons_lode_us_h	0.94	1.12	0.98	0.56	0.99	0.55	0.94	1.16	1.22	0.60	1.20	1.26	1.24
shuthonger_h	0.38	0.49	0.42	0.22	0.42	0.22	0.41	0.63	0.55	0.25	0.52	0.55	0.39
worcester_h	1.33	1.55	1.37	0.48	1.38	0.61	1.28	1.29	1.48	0.66	1.27	1.23	1.02
RMS	0.91	1.06	0.94	0.56	0.95	0.68	0.91	1.03	1.07	0.70	1.01	1.02	0.90

[a] [m]. RMSE measured against gauged water levels within [2012-11-27 19:20:00 UTC, 2012-12-05 23:00:00 UTC].

updating, behaving better throughout the event than the versions with prescribed inflows. The benefit of the simultaneous inflows estimation shown in Table 4 is also shown by a pairwise comparison of the filter with d_e-metric localization [(c) vs. (d)] or the filters with d_n-metric localization [(e) vs. (f)] in time series (Figs. 7–9).

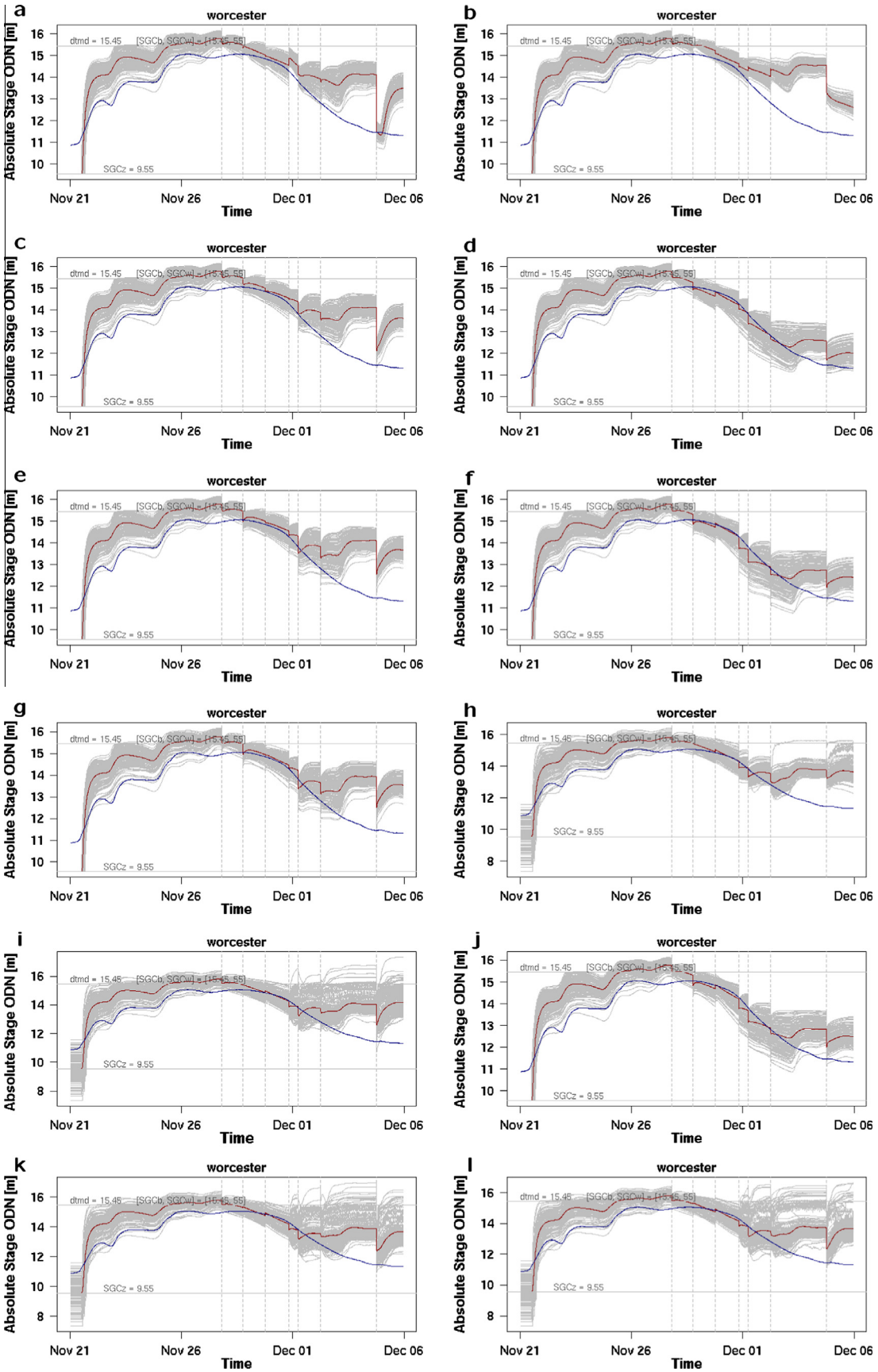

Fig. 7. Water level forecast at Worcester, whose major inflows come from Bewdley (river Severn), Kidder Callows Ln Us (river Stour), and Harford Hill (river Salwarpe). Plot labels refer to the corresponding filter configurations (Table 2). For each plot, gray lines are the forecast ensemble, the red line is the mean forecast and the blue line is the gauged water level, included as a reference. Vertical dashed lines indicate the times of the CSK overpasses/assimilation. Horizontal lines indicate the bank level (labelled as "dtmd"), and the *prior* mean channel bottom level (labelled as "SGCz"). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

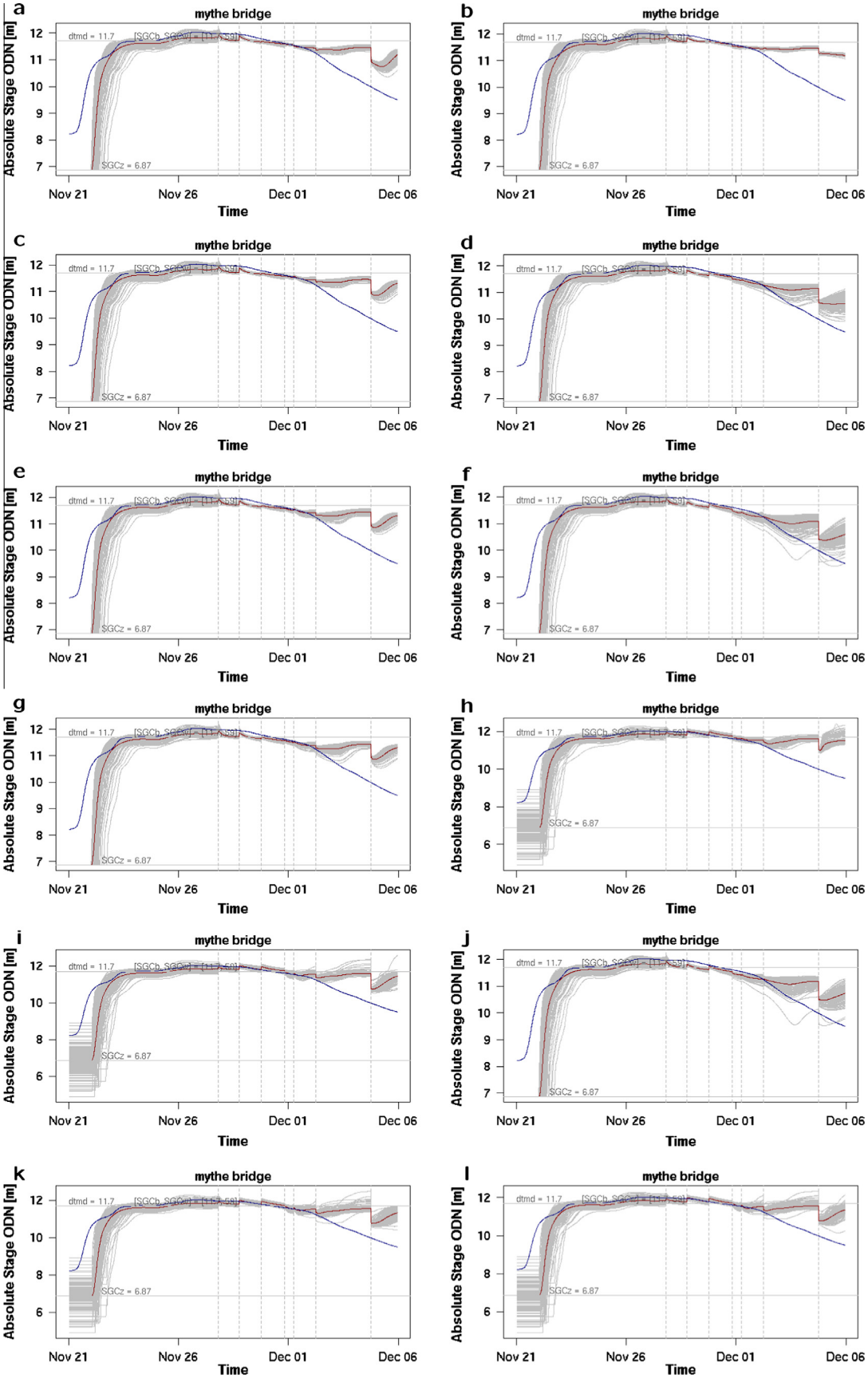

Fig. 8. Water level forecast at Mythe Bridge, in the Severn. Description as in Fig. 7.

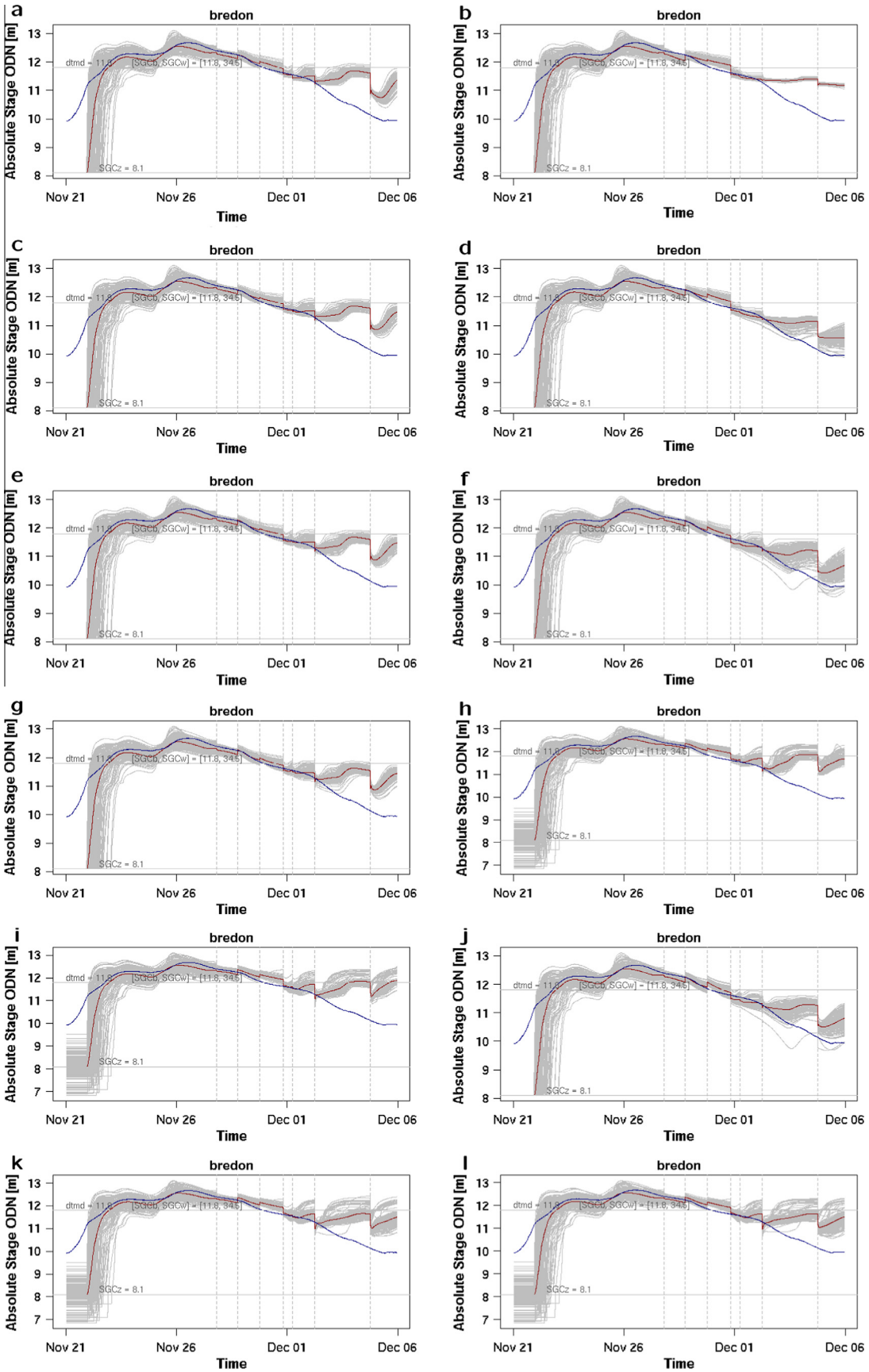

Fig. 9. Water level forecast at Bredon, in the Avon. Description as in Fig. 7.

Overall, the two filters with better performance in the group without friction and/or bathymetry estimation (a–f) are the filters with localization and simultaneous inflow estimation. According to Table 4, these are filter (d) with d_e-metric localization (RMS = 0.56 m), and (f) with d_n-metric localization (RMS = 0.68 m). While the RMS is slightly better for filter (d), the evaluation of the forecast error covariance (for example, as shown in Figs. 4, and 6) indicates that the along-network-based localization is preferable as a forecast error covariance moderation process, and helps further to prevent the development of spurious correlations, which should be adequate for local parameter estimation. Also in the downstream areas, where most of the flood occurred (Mythe Bridge, Saxons Lode US, and Bredon) the RMSE is equal or better for filter (f).

Thus in the following section on simultaneous parameter estimation we focus the discussion on filter configurations with d_n-based localization.

3.4. Parameter estimation

In this section we focus on filter configurations with simultaneous friction and/or bathymetry estimation. We evaluate if these parameters can be simultaneously estimated, and also if this simultaneous estimation leads to an improvement in the flood forecast.

Let us first make a comment about the accuracy generally achieved in this type of flood forecast. Schumann et al. (2011) evaluated the accuracy of sequential aerial photography and SAR data for observing urban flood dynamics, with a case study of the 2007 floods in the same area as our case study. In fact, their 2007 event was the same one that we used for calibrating the model in our case. They reported that, from a number of SAR sensors, the best vertical accuracy was obtained for the high-resolution TerraSAR-X SAR data (which has an horizontal resolution similar to CSK), with a 0.56 m RMSE (evaluated against a high resolution hydraulic model simulations as a surrogate for the unknown true water levels). It is worth noting than in our case study, the agreement between the water levels recorded by gauges and those just updated by the assimilation (i.e. water levels at assimilation times), was generally higher than that indicated by Schumann et al. (2011). Fig. 8 is an indication of this. At that location, in the filter (f), which had the best RMSE, the difference between gauged levels and assimilated levels at the time of the CSK overpasses was kept very low (order of a few centimeters), except for the last overpass, where the difference was ~0.45 m. In fact, most of the RMSE shown in Table 4 relates to the capability of the filter to remain stable and even being able to forecast the levels in the recession of the flood. Let us note also that the estimation of the recession levels was more difficult as rivers were back in bank, whereas WLOs were on the floodplain at those times. With this, Table 4 indicates that, according to the RMS criterion, there is no benefit in simultaneous parameter estimation regarding the flood forecast during the short time span of a single event in this case study. Let us explore possible reasons.

The estimation of Manning's friction coefficient for the major channels appears to converge systematically across all the filter configurations. Fig. 10 shows this convergence for filter (j). With slightly different convergence rates, all filters with friction estimation ([g], [i], [j], and [l]) had a very similar trend, from an initial mean $\overline{\eta_{c1}} = 0.035$ to a final mean $\overline{\eta_{c1}} \approx 0.033$. This systematicity in time and across filters, independent from simultaneous inflow and bathymetry estimation, supports the confidence in the friction estimates. The fact that the convergence is gradual is intrinsic to the assimilation-based estimation, but also the gradual friction decay may well be physical. The washout effect of high water levels on channel beds, resulting in hysteresis of the rating curves (stage-discharge relationships), is known to happen for some flood events (e.g., García-Pintado et al., 2009). On the other hand, the second Manning's friction coefficient, affecting the minor tributaries, did not converge systematically across the filters. For example, from the initial $\overline{\eta_{c2}} = 0.040$ the final estimate was $\overline{\eta_{c2}} = 0.039$ in filter (j), but other filters saw final values slightly higher than the prior ones. With a smaller general influence on the water levels, and generally more distant from the WLOs, the estimation of $\overline{\eta_{c2}}$ seems to be influenced by spurious correlations and, as a summary, does not seem sound.

However, despite the likely adequate estimation of channel friction in the major channels (the ones with a higher influence on general water levels), the feedback of friction estimation on the flood forecast within this event seems negligible. For example, compare the RMS of (e) vs. (g), both without inflow estimation, or (f) vs. (j), both with inflow estimation, where in both comparisons the second filter includes friction estimation. The negligible difference in RMSEs can also be seen by the pairwise comparison of these filters in Figs. 7–9. As $\overline{\eta_{c1}}$ is consistently estimated, one could expect that this sensitivity of friction to the WLOs should be reflected in a sensitivity of the flood forecast to the (likely) improved friction estimates, so leading to a better forecast. However, the convergence being gradual, it seems the DA-forecast cycle does not have time to benefit from the updated friction.

Fig. 11 shows the evolution of bathymetry, along the event, for the rivers Severn and Avon, and filter configuration (k). The chainage 0 for the Avon refers to its junction with the Severn, very close to Mythe Bridge. All the filters including bathymetry estimation with identical localization radius for bathymetry, either with/without simultaneous friction estimation, or with/without simultaneous inflow bias correction (i.e. filters [h], [i], [k], and [l]) show a nearly identical convergence, supporting the robustness of the estimation shown in Fig. 11, independently from other factors. The sequential updating converges systematically toward a profile in which, after the event, the lower part of the Severn is nearly 2 m higher than the prior bathymetry, and the transect between Saxons Lode US and Kempsey gauges is lower than the prior (at some points reaching ~1.5 m of difference with respect to the prior). The highest increments in the updating are due to the assimilation of WLOs from the first overpass. Thereafter, the updating increments become gradually smaller along time. The updatings summarize the influence of the channel conveyance on the flood development. Globally, the SAR-WLOs seem to indicate that the

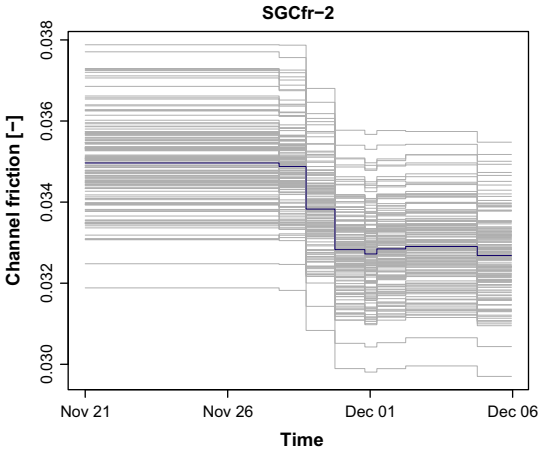

Fig. 10. Evolution of the estimate of the global Manning's coefficient along the sequential assimilation steps for the three major rivers (Severn, Avon, and Teme), and filter configuration (j).

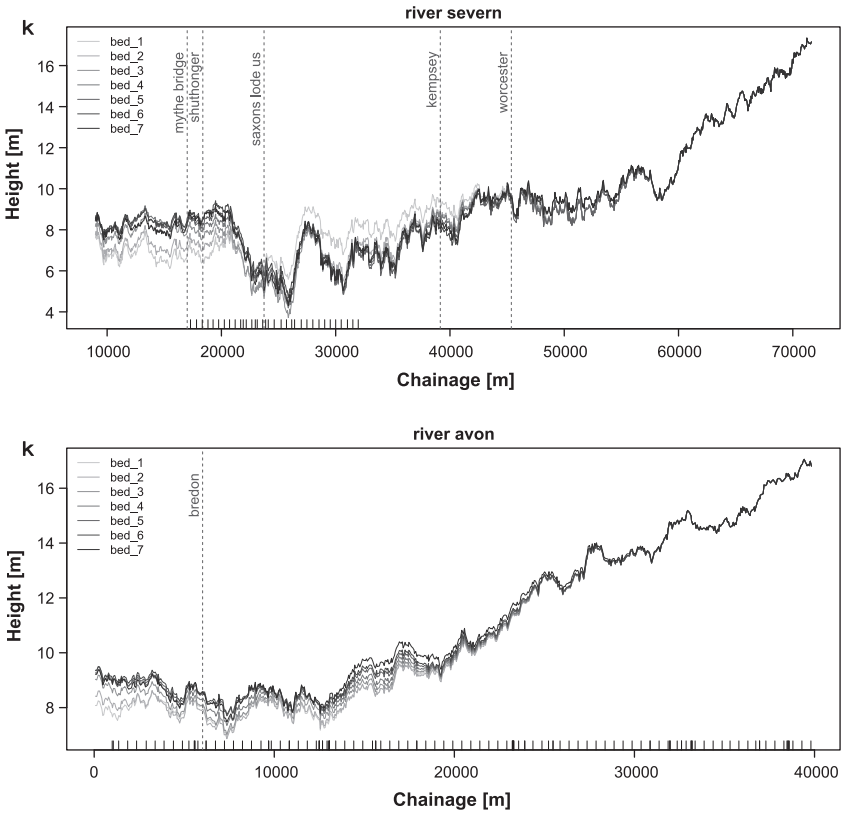

Fig. 11. Evolution of the estimate of bathymetry along the sequential assimilation steps for the river Severn (top), and the Avon (bottom), for the filter configuration (k). The ticks at the bottom indicate the location of the available cross sections. The vertical dashed lines and corresponding labels indicate the location of level gauges used for water level validation.

prior bathymetry was leading to a model which overestimated the release of water from the flooded domain during the early stages of the event. The sequential increments in the bathymetry along the Avon are also systematic, leading to a raised channel bed profile with respect to the prior. In both rivers, the effect of the localization is clearly visible. That is, moving upstream, the increments become gradually smaller as the bed locations move away from the observations (e.g., in the Severn the WLOs roughly generally covered up to the 40 km chainage coordinate, close to Kemspey). The consistent and systematic sequential increments indicate a physical basis for these, as happened with $\overline{\eta_{c1}}$. Note the prior bathymetry was based on an interpolation of the rectangular approximations of the real cross-sections. Thus these plots refer to the bottom of the rectangular channels approximating the real ones.

One could argue that a possible explanation is that since the time at which the surveys for the cross sections were done, the channel bed may have evolved, with some erosion happening upstream in the Severn (in the Saxons Lode US–Kempsey transect) and some sediment deposition happening downstream. We did not have available the metadata indicating the dates for the cross-section surveys. Thus it is not possible for us to check the possibility/degree of this factor. However, the surveys were likely collected before 2002 (UK Environment Agency, personal communication), and several flood events have occurred since that time.

As with friction, despite the consistency in bathymetry estimation, the flood forecast does not improve as a result of the updated bathymetry. Figs. 7–9 shows that the filters including bathymetry estimation had difficulties in forecasting the recession limb, this being a the major effect contributing to the worsened RMSE for these filters. The recession limb is related to the moment in which the bulk of the flood reaches the downstream boundary location and leaves the domain through the South boundary. Thus the updated bathymetry may have led to an improved estimate of the flood extent for the peak of the flood, but the assimilation may be overshooting the estimation of the most downstream bathymetry, preventing an accurate release of water from the domain in the last stages of the flood. This seems to be the case. That is, as the flood wave evolves the control from bathymetry on the flood development is gradually moving downstream. In the first assimilation steps, it is likely that bathymetry around Saxons Lode US and Shuthonger exerts a stronger control on the flood than the most downstream areas, and the simultaneous updating of these most downstream bathymetries is collaterally caused by their correlated errors with those from the bathymetry some kilometers upstream. An additional difficulty is that the last 5 km of the domain remained unobserved during the event (i.e., as a results of the multicriteria screening to obtain the WLOs to be assimilated, none of these was located in the last 5 km of the Severn within the domain). As in the experimental design we did not provide any inflation for bathymetry, the channel bed estimated variance is gradually reduced along the sequential assimilation, and by time the most downstream area (around Mythe Bridge and toward the South) has the strongest influence on the release of water from the domain, the bathymetry spread is too low to be properly updated (and the WLOs did not reach the last ~5 km of the South of the domain —see, e.g., Fig. 4—). A plot similar to Fig. 11, but regarding the evolution of the bathymetry ensemble spread along the sequential assimilation indicates that the standard deviation of bathymetry around Mythe Bridge decreases from an initial 0.8 m until a final 0.4 m (not shown; available on request). The chosen c.v. = 0.15 in the bathymetry error generation, reflects our confidence in the prior bathymetry estimates.

Overall, it seems that either the chosen 5000 m spatial correlation length in the stochastic generation of the bathymetry error was too high or the 10,000 m localization window for the d_n-based

localization for bathymetry estimation was too high (or both factors), leading to an overshooting of the downstream bathymetry increments, and subsequent problems. To test this point, we conducted a further simulation (filter configuration [m]) with 5000 m as localization window for bathymetry estimation. In effect, the general trends in the sequential bathymetry updating are similar to the previous experiments, but the increments gradually fade downstream (see supplementary material). This translates into a steeper recession limbs (closer to those of configuration [f]) and better statistics (see [m] vs. [l] in Table 4). Thus everything indicates that by tuning the localization radii and correlation length in the bathymetry error generation the simultaneous parameter/state estimation process could be further improved. However, as indicated in the experimental design, to provide a detailed exploration of the parameter space and localization parameters goes beyond the scope of the current study.

4. Conclusions

We have shown that under a relatively complex scenario with simultaneous uncertain inflows into a flooded domain, a satellite-based forecast of the flood with high accuracy is possible through the assimilation of the satellite-based WLOs into a flood forecast model. However, several aspects should be taken into account for a successful operational application of EnKF-based assimilation of EO-based WLOs and forecast. First, a moderation of the forecast error covariance based on spatial localization is necessary to avoid filter divergence. Second, inflow estimation also improves the forecast. This second point is only valid if localization is applied, otherwise the incorrect forecast error covariance development prevents the global filter from adequately locating inflow errors and the corresponding benefit in their online correction. Third, the implementation should consider the possible uncertainty in model parameters and their simultaneous online estimation.

Importantly, inflow estimation, for the evaluated filter configurations which do so in this study, simultaneously attempts to correct for missing inflows/outflows in the domain covered by the hydrodynamic model (e.g., precipitation or minor tributary inflows) between the satellite WLO locations and corresponding inflow boundary locations, which are unaccounted for by the upstream hydrologic model. Our results show that, for properly localized filters, this is not to be seen as a problem, but as a practical benefit for the actual flood forecast, as indicated above. Further, the possible strategy of attempting to adjust the state of the coupled upstream hydrologic model would just be sensible provided no additional inflow/outflow occurred between the outlet of the hydrologic model (i.e. inflow boundary condition location) and the WLOs contributing to the update of the corresponding inflow time series, which cannot be assured in the general case.

The study shows that if the physical connectivity of flows is considered in the form of the newly proposed along-network metric for the localization, the development of forecast error covariances is sounder than that resulting from the use of a standard as-the-crow-flies distance. The relevance of this regarding the forecast skill should depend on the geometry of the network in each specific case, and further studies would be needed to assess this relevance.

The study is not conclusive about how simultaneous parameter estimation (friction and bathymetry —considered as a parameter from the estimation point of view—) interacts with the flood forecast. There seems to be a benefit in the development of sound forecast error covariances, and also, the convergence of the parameters seems to be consistent. However, the simultaneous parameter estimation does not improve the on-going flood forecast skill in the studied case. In other cases (steeper rivers, faster flow, etc.) things might be different. The localization parameters used in the case study for bathymetry estimation seem to be far from optimal, and tuning these parameters could lead to a better estimates in the inverse problem (i.e. bathymetry estimation), with improved feedback on the flood forecast. The exploration of the localization parameters and localization method on the simultaneous state and parameter estimation problem warrants further investigation.

There should be also a more positive feedback from the simultaneous parameter estimation (provided parameters are properly estimated) for longer events, in floods developed over bigger areas. The estimated parameters resulting from the assimilation should also lead to improved forecasts in future events in the same domain. This is mostly the case for bathymetry estimation, to which the forecast sensitivity is higher that the sensitivity to channel friction, at least in this study.

Other factors influencing filter performance have not been explored here. For example, the sensitivity of the skill to the ensemble size, or the possibility of conducting transformations to diminish the negative effect of the nonlinearities in the filtering process.

The study described here has been a retrospective forecasting (hindcasting) rather than a real forecasting process. Current high resolution satellite SARs including the COSMO-SkyMed constellation and TerraSAR-X do not yet provide near real-time geo-registered imagery from which WLOs could be extracted and assimilated to provide a flood forecast (the only exception is RADARSAT-2). However, the European Space Agency (ESA) will shortly launch the Sentinel-1 two-satellite high resolution SAR constellation which will give almost daily coverage of floods at European latitudes. The first satellite of the pair was launched in April 2014, and the second is scheduled to follow in 2016. The system will allow processed multi-look geo-registered SAR images to be available to the user only an hour or so after download to the ground station. In addition, ESA have developed the Fully Automated Acqua Processing Service (FAAPS) to process SAR images in near real-time to create geo-registered rural flood extent maps and deliver them to the user via the Internet. It should therefore be possible in the near future to use the techniques developed here to help to provide flood forecasts in near real-time.

Overall, further possible improvements notwithstanding, the study indicates that a properly constructed stand-alone EO-based flood forecast is accurate enough for operational applications even for floods developed within relatively complex river networks.

Acknowledgements

This work was supported by NERC through the DEMON project (NE/I005242/1), within the NERC SRM (Storm Risk Mitigation) programme, and the SINATRA project (NE/K00896X/1), within the NERC FFIR (Flooding From Intense Rainfall) programme. Thanks are due to Moira Mason for assistance with the satellite image acquisition. We thank the EA for the precipitation and standard flow/level gauge datasets. We also thank two anonymous reviewers, whose insightful comments helped to improve this manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2015.01.084.

References

Anderson, J.L., 2001. An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev. 129, 2884–2903.

Andreadis, K.M., Clark, E.A., Lettenmaier, D.P., Alsdorf, D.E., 2007. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. Geophys. Res. Lett. 34, L10403–.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Cloke, H., Dee, D., Dutra, E., Muñoz Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., Vitart, F., 2013. Era-interim/land: a global land water resources dataset. Hydrol. Earth Syst. Sci. Discuss. 10, 14705–14745.

Bates, P., De Roo, A., 2000. A simple raster-based model for flood inundation simulation. J. Hydrol. 236, 54–77.

Bates, P., Horrit, M., Fewtrell, T., 2010. A simple inertial formulation of the shallow water equations for efficient two dimensional flood inundation modelling. J. Hydrol. 387, 33–45.

Biancamaria, S., Durand, M., Andreadis, K., Bates, P., Boone, A., Mognard, N., Rodríguez, E., Alsdorf, D., Lettenmaier, D., Clark, E., 2011. Assimilation of virtual wide swath altimetry to improve arctic river modeling. Remote Sens. Environ. 115, 373–381.

Bogner, K., Pappenberger, F., 2011. Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. Water Resour. Res. 47.

Dee, D.P., 2005. Bias and data assimilation. Q. J. R. Meteorol. Soc. 131, 3323–3343.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford UP, NY.

Donigian, A., Bicknell, B.R., Imhoff, J.C., 1995. Hydrological simulation program – fortran (hspf). In: Singh, V. (Ed.), Computer models of watershed hydrology. Water Resources Publications, Littleton, CO, pp. 395–442.

Durand, M., Andreadis, K.M., Alsdorf, D.E., Lettenmaier, D.P., Moller, D., Wilson, M., 2008. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. Geophys. Res. Lett. 35, L20401.

Durand, M., Neal, J., Rodríguez, E., Andreadis, K.M., Smith, L.C., Yoon, Y., 2014. Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope. J. Hydrol. 511, 92–104.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143–10162.

Friedland, B., 1969. Treatment of bias in recursive filtering. IEEE Trans. Autom. Control 14, 359–367.

García-Pintado, J., Barberá, G.G., Erena, M., Castillo, V.M., 2009. Calibration of structure in a distributed forecasting model for a semiarid flash flood: dynamic surface storage and channel roughness. J. Hydrol. 377, 165–184.

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

Gaspari, G., Cohn, S.E., 1999. Construction of correlation functions in two and three dimensions. Q. J. R. Meteorol. Soc. 125, 723–757.

Gessese, A.F., Sellier, M., Houten, E.V., Smart, G., 2011. Reconstruction of river bed topography from free surface data using a direct numerical approach in one-dimensional shallow water flow. Inverse Prob. 27, 025001.

Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V.R.N., De Lannoy, G.J.M., De Keyser, R., Pfister, L., Hoffmann, L., Savenije, H.H.G., 2011. Assimilating sar-derived water level data into a hydraulic model: a case study. Hydrol. Earth Syst. Sci. 15, 2349–2365.

Hamill, T.M., Whitaker, J.S., Snyder, C., 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon. Wea. Rev. 129, 2776–2790.

Hostache, R., Lai, X., Monnier, J., Puech, C., 2010. Assimilation of spatially distributed water levels into a shallow-water flood model. Part ii: Use of a remote sensing image of Mosel river. J. Hydrol. 390, 257–268.

Houtekamer, P.L., Mitchell, H.L., 1998. Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev. 126, 796–811.

Houtekamer, P.L., Mitchell, H.L., 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Wea. Rev. 129, 123–137.

Hunt, B.R., Kostelich, E.J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Physica D 230, 112–126.

Kalman, R., 1960. A new approach to linear filtering and prediction problems. Trans. ASME Ser. D: J. Basic Eng. 82, 35–45.

Kalnay, E., 2002. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press.

Kim, J., Ryu, J., 2014. Threshold of basin discretization levels for hspf simulations with nexrad inputs. J. Hydrol. Eng. 19, 1401–1412.

Kirchgessner, P., Nerger, L., Bunse-Gerstner, A., 2014. On the choice of an optimal localization radius in ensemble Kalman filter methods. Mon. Wea. Rev. 142, 2165–2175.

Livings, D.M., Dance, S.L., Nichols, N.K., 2008. Unbiased ensemble square root filters. Physica D 237, 1021–1028.

Marín-Pérez, R., García-Pintado, J., Skarmeta, A.G., 2012. A real-time measurement system for long-life flood monitoring and warning applications. Sensors 12, 4213–4236.

Martinis, S., Twele, A., Voigt, S., 2011. Unsupervised extraction of flood-induced backscatter changes in sar data using markov image modeling on irregular graphs. IEEE Trans. Geosci. Remote Sensing 49, 251–263.

Mason, D., Schumann, G.-P., Bates, P.D., 2010a. Data utilization in flood inundation modelling. In: Pender, G., Faulkner, H. (Eds.), Flood Risk Science and Management. Wiley-Blackwell, pp. 209–233.

Mason, D., Speck, R., Deveraux, B., Schumann, G., Neal, J., Bates, P., 2010b. Flood detection in urban areas using TerraSAR-x. IEEE Trans. Geosci. Remote Sensing 48, 882–894.

Mason, D., Davenport, I., Neal, J., Schumann, G., Bates, P., 2012a. Near real-time flood detection in urban areas and rural areas using high resolution synthetic aperture radar images. IEEE Trans. Geosci. Remote Sensing 50 (8), 3041–3052.

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

Mason, D., García-Pintado, J., Dance, S., 2014. Improving flood inundation monitoring and modelling using remotely sensed data. in: Civil Engineering Surveyor, pp. 34–37.

Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N., De Lannoy, G.J.M., De Keyser, R., Savenije, H.H.G., 2010. Towards the sequential assimilation of sar-derived water stages into hydraulic models using the particle filter: proof of concept. Hydrol. Earth Syst. Sci. 14, 1773–1785.

Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., Pappenberger, F., 2009. A data assimilation approach to discharge estimation from space. Hydrol. Process. 23, 3641–3649.

Neal, J., Schumann, G., Bates, P., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resour. Res. 48, n/a–n/a.

Nerger, L., Gregg, W.W., 2007. Assimilation of seawifs data into a global ocean-biogeochemical model using a local {SEIK} filter. J. Mar. Syst. 68, 237–254.

Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., Patil, D., Yorke, J., 2004. A local ensemble Kalman filter for atmospheric data assimilation. Tellus A 56.

Pauwels, V.R.N., De Lannoy, G.J.M., Hendricks Franssen, H.-J., Vereecken, H., 2013. Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid Kalman filter. Hydrol. Earth Syst. Sci. 17, 3499–3521.

Raclot, D., 2006. Remote sensing of water levels on floodplains: a spatial approach guided by hydraulic functioning. Int. J. Remote Sens. 27, 2553–2574.

Roux, H., Dartus, D., 2008. Sensitivity analysis and predictive uncertainty using inundation observations for parameter estimation in open-channel inverse problem. J. Hydraulic Eng. 134, 541–549.

Ryu, J., 2009. Application of hspf to the distributed model intercomparison project: case study. J. Hydrol. Eng. 14, 847–857.

Schultz, C., Ahmed, S., Mandel, R., Moltz, H., 2013. Improvement in hspf's low-flow predictions by implementation of a power law groundwater storage-discharge relationship. JAWRA J Am. Water Resour. Assoc., 1–19

Schumann, G., Hostache, R., Puech, C., Hoffmann, L., Matgen, P., Pappenberger, F., Pfister, L., 2007. High-resolution 3-d flood information from radar imagery for flood hazard management. IEEE Trans. Geosci. Remote Sensing 45, 1715–1725.

Schumann, G.J.-P., Neal, J.C., Mason, D.C., Bates, P.D., 2011. The accuracy of sequential aerial photography and sar data for observing urban flood dynamics, a case study of the uk summer 2007 floods. Remote Sens. Environ. 115, 2536–2546.

Smith, P.J., Thornhill, G.D., Dance, S.L., Lawless, A.S., Mason, D.C., Nichols, N.K., 2013. Data assimilation for state and parameter estimation: application to morphodynamic modelling. Quarterly J. R. Meteorol. Soc. 139, 314–327.

Thornhill, G., Mason, D., Dance, S., Lawless, A., Nichols, N., Forbes, H., 2012. Integration of a 3d variational data assimilation scheme with a coastal area morphodynamic model of morecambe bay. Coast. Eng. 69, 82–96.

Vanden-Eijnden, E., Weare, J., 2012. Data assimilation in the low noise regime with application to the Kuroshio. Mon. Wea. Rev. 141, 1822–1841.

Vörösmarty, C., Askew, A., Grabs, W., Barry, R.G., Birkett, C., Dell, P., Goodison, B., Hall, A., Jenne, R., Kitaev, .L., Landwehr, J., Keeler, M., Leavesley, G., Schaake, J., Strzepek, K., Sundarvel, S.S., Takeuchi, K., Webster, F., Group, T.A.H., 2001. Global water data: a newly endangered species. Eos Trans. AGU 82, 54–58.

Wang, X., Bishop, C.H., Julier, S.J., 2004. Which is better, an ensemble of positive–negative pairs or a centered spherical simplex ensemble? Mon. Wea. Rev. 132, 1590–1605.