**hess-2015-54: Authors' response (in *blue italics*) to comments by Anonymous Referee #1 (in black). In green we indicate line numbers within the latexdiff version of the manuscript (it can be found at the end of this pdf document).**

General Comments:

This manuscript titled "Operational aspects of asynchronous filtering for hydrological forecasting" compares the traditional Ensemble Kalman fitler (EnKF) with its modified version, asynchronous EnKF (AEnKF). The main difference is in the update step, with the traditional EnKF using observation data from a single time step whereas the AEnKF uses observation data from multiple time steps in the past.

The paper is well written and structured in a systematic manner. The results clearly support the findings and make an important contribution to the data assimilation literature.

*Authors' response: We appreciate that Referee #1 sees our manuscript as well written, systematically structured and making an important contribution to the DA literature. We discuss and answer her/his comments in detail below.*

However, my main concern is how to identify the specific contribution from the AEnKF approach. That is, are the improvements in the discharge forecast purely from using multiple data points in the past at the update stage?

*Authors' response: Figure 7 shows the forecast accuracy of three different numbers of data points in the past (W=0, 5 and 11). This figure shows clearly that the differences in the forecast improvement of these various setups are purely due to using multiple data points in the past at the analysis step. We will stress this more clearly in the revised manuscript.* ***[see LINES 532-535, also 377-379 + 424-426]***

This question arises from the observation (in Figure 7) that the forecast improvements in the AEnKF are very systematic for most lead times. Given that the forecast improvements are systematic, I think it is safe to say that the improvements at the update stage are also systematic. If this is true, that means that there is a linear relationship between the updates made with a single observation data and those made with multiple data points in the past. Practically, this is unlikely because you expect different levels of updates especially for transition periods like dry to wet and vice-versa (or low to peak flows and vice-versa).

*Authors' response: Results in Figure 7 show the average behavior (over many forecasts) of an improved initial state on the forecast accuracy of the different filters. We agree that updates may vary in time. This is a valid approach, since data assimilation is expected to improve forecasts in an average sense. What we see is in line with many other DA studies that show a similar kind of behavior (not only in hydrology). That the improvements in forecast accuracy decay with lead time in a systematic fashion is to be expected. It would be very strange if they wouldn't, as the model slowly returns to the model climatology each time a forecast is made (no observations, so no updating). This is not a sign that a constant bias is present, but a sign that an update of the initial condition lasts for a while.* ***[see LINES 517-522]***

On a related note, the level of improvement from EnKF to AEnKF seemed to be relatively constant irrespective of the lead time (Figure 7). For me this is a worrying sign because it looks as if you are able to identify and quantify this constant difference (maybe call it bias) then you can get a better forecast. In other words, the AEnKF seemed to have a better treatment of bias. It will be interesting to know how the AEnKF will perform for low flows, normal flows, and peak flows independently. That is, will updating the model with multiple data points in the past always have a positive impact on forecasts made during low and peak flows?

*Authors' response: We will state in the revised manuscript that for the high flows, the AEnKF with a longer time window W is able to make corrections that last longer on average (this is not a constant bias) with respect to the shorter time window W. Additionally, we will change the title to "Operational aspects of asynchronous filtering for flood forecasting"* ***[see TITLE]****, to emphasize that the presented study investigates high flows only. We agree with the reviewer that characterization of the statistical properties of the temporal flow dynamics (i.e. typical time scales of flood peaks as compared to low flows) is a relevant issue. The length of the time window W has to be seen relative to the time scale of the river flow dynamics. We assume that for low flow conditions, the improved skill of longer W with respect to shorter W will become negligible, as low flows exhibit less temporal dynamics than high flows.* ***[see LINES 650-660]***

I think these are important questions the authors need to address, even if they do not have results to support it at least a clarification is needed about which conditions their methodology will mostly apply.

In Figure 6, the no-update looked pretty accurate and almost comparable to the assimilation, but in Figure 7 the evaluation measures for no-update is very poor. This is a stunning difference, please clarify.

*Authors' response: We agree that the no-update scenario matches the magnitude of the major peak quite well, although it has quite a large spread. Additionally, when we consider the ensemble mean of the no-update scenario with respect to the DA scenarios, the accuracy deteriorates. This is shown in Figure 7 for the set of eight flood peaks, while Figure 6 shows only one individual peak.* ***[see LINES 483-487]***

---

## hess-2015-54: Authors' response (in *blue italics*) to comments by Dr. Ming Pan (in black). In green we indicate line numbers within the latexdiff version of the manuscript (it can be found at the end of this pdf document).

The work titled "Operational aspects of asynchronous filtering for hydrological forecasting" by O. Rakovec et al. presented a data assimilation study for river discharge simulations using the Asynchronous Ensemble Kalman Filter (AEnKF). The experiments are mainly focused on testing the effects of two procedures: lumped filter updates against observations from multiple time steps and partial updating of the model states. The study is very carefully designed and carried out and the paper is well organized and well written. In general, the study and presentation is of fairly good scientific quality. I recommend its publication after minor revisions.

*Authors' response: We appreciate that Referee Dr. Ming Pan finds our manuscript well written, carefully designed and well organized. We discuss and answer his comments in detail below.*

Here are my main concerns:

First of all, the AEnKF procedure needs some more justification and clarification. The authors described the AEnKF as a "state augmentation". It is not exactly the case because it only updates the current state $x_k$ and none of the previous ones from $x_{k-1}$ to $x_{k-W}$ . (Of course, that probably shouldn't matter much if we only care about the forecasts.) Sakov et al. 2010 claimed that AEnKF is "formally equivalent to EnKS solution." This statement is only true if the dynamic system is strictly linear (see Equation 17 in Sakov et al. 2010). If the dynamic system is not linear, the step-by-step updates will make a big difference w.r.t. the lumped updates because the nonlinearity errors will accumulate during continuous and unconstrained model integration. Is the river routing scheme (kinematic wave) linear? Is the hydrological model linear? I guess not because otherwise the ensemble method wouldn't be used.

*Authors' response: We describe the AEnKF approach as an augmentation of the state matrix with past forecasted observations* ***[see LINES 254-261 + Eq 13]****. The Referee is correct that $x_{k-1}$ to $x_{k-W}$ are not included in the augmented state matrix, but we do not state that in the manuscript. If past model states would have been included in the analysis, we should have spoken rather about a smoother than a filter. Additionally, we will acknowledge in the revised manuscript that both the kinematic wave model for the routing, and the hydrological model exhibit nonlinear behavior. Additionally, we already mention on* ***[LINES 98-102]*** *that although "Kalman-type assimilation methods were developed for an idealized modeling framework with perfectly linear problems with Gaussian statistics, it has been demonstrated to work well for a large number of different nonlinear dynamical models (Evensen, 2009)". We also add a note on the non-linearity of the hydrological as well as the routing scheme* ***[see LINES 214-216].***

Given that, we can say the longer the update window W is, the more nonlinearity errors to accumulate. However, longer windows will bring more information to the updates. If the nonlinearity is not a problem, then the window should be as long as necessary. Pan and Wood 2013 experimented a river discharge assimilation approach that resorts to a full and explicit state augmentation over the longest necessary window, i.e., across the maximum streamflow travel time of the river basin involved. (Their study only works with a fully linear river routing scheme thus is free of nonlinearity errors.)

*Authors' response: We thank the Referee for mentioning the interesting study by Pan and Wood (2013). We will include a note on their results in the revised version of our manuscript.* ***[see LINES 660-663]***

It is not clear whether the discharge observations at one gauge station are used to update all the grid cells in the entire basin or just those within the subcatchment that flows down to that gauge station. This is an important issue because the discharge from one gauge does not contain any information about the grid cells outside (i.e. downstream) of its own drainage area. Also, are the discharge data from all 6 river gauges assimilated altogether simultaneously, or one gauge at a time? The discharge from 6 gauges contain information of different lag times with respect to different grid cells. See Pan and Wood 2013 for a fully explicit handling of such lags in time and space.

*Authors' response: We assimilated all discharge gauges simultaneously in the case study where four gauges (1,3,5,6) were assimilated. This is the same as was presented by Rakovec et al (2012). This means that we did not apply any localization method here. We will make this clearer in the revised manuscript.* ***[see LINES 341-343]*** *The difference between forecast and analysis of individual model states is shown in a spatial manner in Figure 5 for one time instant (see also Figure 4). These results illustrate the effect of W and spatially distributed discharge observations on spatially distributed innovations. Note that two other gauges (2 and 4) were left out from the analysis for validation.*

Another major concern I have is the very short length of the study period. All we can see is just one winter event. It is really too short. We can't even see a robust model validation. We can't see how the DA behaves under other conditions (like low flows). This really limits the significance and robustness of any conclusion you can draw here. If extension is impossible, the conclusions have too be very carefully constrained.

*Authors' response: Although the exemplary Figures 6 and 8 present results for a single flood event, the overall results presented in Figure 7 include all eight flood events presented in Table 1, which is also written on* ***[LINES 512-514]*** *and in the caption of Figure 7. The Referee is right however that our analysis can be generalized only for high flows and not directly for low flow conditions. Therefore, we will include a note in the revised manuscript stating that analysis of low flow conditions was beyond the scope of the presented study.* ***[see LINES 653-660 + 125+634+636]*** *Additionally, we will change the title as follows: Operational aspects of asynchronous filtering for flood forecasting.* ***[see TITLE]***

I suggest the authors calculate the auto-correlation function of the innovation time series. That's the best way answer Referee 1's concerns on systematic errors. The EnKF types of methods are supposed to correct dynamic errors (i.e. time-random), persisting biases are considered static (time-invariant) errors, and they should be corrected using static methods.

*Authors' response: We thank the Referee Dr. Ming Pan for his suggestions, but we do not consider calculating the innovation auto-correlations. However, Figure 7 shows the RMSE (so +/- corrections contribute similarly) of many forecasts over many events. It is impossible to deduce on the basis of Figure 7 that there is a constant bias. That the forecast accuracy is decaying with leadtime is common to all DA studies, as the model slowly returns to the model climatology.*

Figure 6: I can't distinguish between the lines of different shades of red (different lead times). I can't even count how many lines there are on the plots. This has to be redone.

*Authors' response: We agree, we will improve quality and readability of Figure 6 in the revised manuscript. Color scheme and the size of the figure are adjusted.* ***[see FIGURE 6]***

*References:*


*Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer, doi:10.1007/978-3-642-03711-5, 2009.*

*Pan, M. and Wood, E. F.: Inverse streamflow routing, Hydrol. Earth Syst. Sci., 17, 4577-4588, doi:10.5194/hess-17-4577-2013, 2013.*

*Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435-3449, doi:10.5194/hess-16-3435-2012, 2012.*


---

## hess-2015-54: Authors' response (in *blue italics*) to comments by Anonymous Referee #3 (in black). In green we indicate line numbers within the latexdiff version of the manuscript (it can be found at the end of this pdf document).

General evaluation:

This paper is interesting especially because it tests a methodology for data assimilation, the asynchronous Ensemble Kalman Filter, which was not tested yet in hydrology. The comparison of this method with EnKF warrants publication in my opinion, but one critical point needs to be resolved.

*Authors' response: We appreciate that Referee #3 finds that the presented analysis warrants publication. We discuss and answer her/his comments in detail below.*

It was unclear to me whether in the experiments with EnKF and AEnKF the same amount of observations was assimilated. It is logical that assimilating more data would give a better result. Can the authors clarify this and also stress this more in the paper.

*Authors' response: The amount of observations being assimilated into the model depends on the magnitude of W. We tried to avoid statements suggesting that we did a comparison of EnKF and AEnKF, because we rather evaluated AEnKF for different lengths of W. We will state more clearly in the revised manuscript that the amount of information differs for different values of W.* ***[see LINES 375-377 + 424-426]***

There are other points which make that the assimilation of discharge data in this rainfall- runoff model shows significant deficiencies. These are the considered uncertainty sources in the experimental set-up, the magnitude of the observation error and the normality assumption for the observations, and considering a time lag for updating model states with help of discharge data. These limitations should be acknowledged in the paper in the abstract and the conclusions. I agree that the main new point of the paper is the comparison of AEnKF and EnKF, and that we can live with these limitations then, but they should be acknowledged.

*Authors' response: We will acknowledge these specific limitations (as further explained below) more clearly in the abstract and the conclusions.*

Altogether I believe moderate revision is needed.

Main points:

Section2.1: We will see later that rainfall is assumed deterministic, whereas all uncertainty is attributed to soil moisture. A normal procedure would be to assume rainfall uncertain as it is the most uncertain component for predictions with rainfall-runoff simulations. In spite of what the authors say, in several studies uncertainty in rainfall is considered in these studies. I think that this assumption should be more critically evaluated in the manuscript and in the discussion and it would be good if the authors discuss its implications. It would be good that this decision is also directly visible in the abstract and conclusions, as it is important information for the experiment.

*Authors' response: We agree with the reviewer that input uncertainty is a very important issue. In previous contributions we investigated the effect of more complex ways of perturbing the rainfall and its effect on forecast accuracy (see Rakovec et al., 2012a,b, Noh et al., 2014). The noise on the soil moisture used in this study more or less resulted in similar open loop simulations of the discharge as in the studies by Rakovec et al mentioned above. The focus of this manuscript is on the filter itself rather than on the effects of the applied noise. We will make this clearer in the revised manuscript.* ***[see LINES 8 + 354-360 + 666-670]***

Page 3173, line 16-17: Why was inverse distance weighting used and not kriging? Maybe add a short statement.

*Authors' response: We agree with the Referee that there are other ways of interpolating rain gauge observations, such as kriging, which might be more appropriate. However, evaluating the benefits of different rainfall interpolation techniques was deemed beyond the scope of the study. We used a method used in operational practice as this study is also oriented towards operational benefits of AEnKF over EnKF.* ***[see LINES 142-146]***

Eq.10-12: It is unclear how discharge is treated here, this is not discussed. But it seems that although discharge is typically not normally distributed, this is neglected here. At least additional discussion would be important here.

*Authors' response: We don't 100% understand the remark made by the Referee. Most probably the reviewer feels that we should make a statement about the non-normality of the discharge pdf. If that is the case, we already mention on* ***[LINES 98-102]*** *that although "Kalman-type assimilation methods were developed for an idealized modeling framework with perfectly linear problems with Gaussian statistics, it has been demonstrated to work well for a large number of different nonlinear dynamical models (Evensen, 2009)".*

Page 3180, Line 23: This is a simple error model. Why not uncertain precipitation and parameters? This decision warrants more discussion, as already indicated above.

*Authors' response: As mentioned above, in previous contributions we investigated the effect of more complex ways of perturbing the rainfall and its effect on forecast accuracy (see Rakovec et al., 2012a,b, Noh et al., 2014). For this study it was important that the error model produced reasonable results in the open-loop and did not lead to any numerical instability. In the end perturbing the soil moisture has a similar effect to perturbing the precipitation.* ***[LINES 354-360]***

Page 3180, Line 26: I suggest showing results from these calculations as for the moment the paper is not very large and the number of figures not too high. It is interesting to learn how many ensemble members are needed for which type of model. For example, it is typically found that for distributed models the number of ensemble members has to be larger, especially if parameter estimation is also involved.

*Authors' response: We did not include a figure on the effect of the ensemble size, because really negligible differences were observed. Nevertheless, we will include a note that such a small ensemble size as presented in the manuscript would not be possible if parameter estimation would be involved or if more complex error models would be employed. More detailed analysis is left for further research beyond the scope of this study.* ***[LINES 364-367]***

Page 3180, Line 27: A discussion on the magnitude of the observation error is needed. Literature on observation errors for discharge measurements suggests in general a much higher measurement error.

*Authors' response: We agree that the magnitude of the observation error could be even larger, but we followed the approach of Clark et al (2008) and Weerts and El Serafy, (2006). The latter explicitly state that a standard deviation of 0.1 represents a large error. However, if an even larger observation error would be employed, we expect marginal differences in the model performance among individual scenarios (in a relative sense). Of course the forecast performance would deteriorate for all, as the weight of the observations would become smaller.*

Page 3182, Line 10-16: Sorry if I missed something, probably did not get it right. You compare EnKF and AEnKF, where AEnKF assimilates the current observation and ten observations for the past. Did you apply EnKF then for this time period at each time step when data became available? This would mean, if you applied AEnKF with W=10, did you apply EnKF with W=0 eleven times for this period, so that both methods ingested the same amount of data. This is needed for a fair comparison, but it is not clear to me if this has been done. Please clarify.

*Authors' response: The amount of observations being assimilated into the model depends on the magnitude of W. We tried to avoid statements suggesting that we did a comparison of EnKF and AEnKF, because we rather evaluated AEnKF for different lengths of W. We will state more clearly in the revised manuscript that the amount of information differs for different values of W.* ***[see LINES 375-377 + 424-426]***

Page 3185, Line 1-11: It is unclear how discharge is related to past soil moisture or upper zone storage states. If it is used to update current soil moisture the procedure is suboptimal I think as discharge will have a higher correlation with past UZ/SM-conditions. If the time lag is not considered some of the conclusion (i.e., updating soil moisture not important) might be related to this specific set-up. In this case, it would be good to add some relaxing statement in the discussion.

*Authors' response: This is done through the Kalman gain. When the correlation is lower the update will be smaller. AEnKF exploits the correlation between the present state and the state not only at the previous time step, but also further in the past. Therefore knowing the present state is sufficient to determine forecast. It may be possible to use the correlation between discharge at the present time and UZ/SM in the past for data assimilation, however, this is beyond the scope of this study. Nevertheless, we speculate that this will only be useful in a smoothing context (i.e. present discharge may bring information on UZ/SM in the past), not in a filtering context as in the present study.* ***[see LINES 553-564]***

Page 3188, Line 2: "a rainfall-runoff model" instead of "hydrological model in operational settings".

*Authors' response: We thank the Referee for suggestion and we will include this suggestion.* ***[see LINE 685]***

Figure 6. Extend caption to again mention the different scenarios that are displayed here.

*Authors' response: We will extend the caption of Figure 6 such that it becomes self-explanatory. Additionally, to improve readability of the figure, the color scheme and the figure size were adjusted* ***[see FIGURE 6]***

Figure 7. Not so clear to me as there are a bit too many lines. Maybe you can find a better solution.

*Authors' response: We agree that the current figure is rather small, which is due to limited margins of the HESSD layout. We will make sure that the Figure 7 is large enough in the final HESS layout, where more space for Figures is provided.*

Editorial:

P3175, L 1: change to: "(..) as for the EnKF".

*Authors' response: We thank the Referee for the suggestion and we will include this suggestion in the revised manuscript.* ***[see LINE 195]***

P3176, L9: change to: "(..) the model states of the ensemble member are updated as follows:"

*Authors' response: We will change this to "(..) the model states of the $i^{th}$ ensemble member are updated as follows:".* ***[see LINES 235]***

*References:*

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

*Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer, doi:10.1007/978-3-642-03711-5, 2009.*

*Noh, S. J., Rakovec, O., Weerts, A. H., and Tachikawa, Y.: On noise specification in data assimilation schemes for improved flood forecasting using distributed hydrological models, J. Hydrol., 519, 2707–2721, doi:10.1016/j.jhydrol.2014.07.049, 2014.*

*Rakovec, O., Hazenberg, P., Torfs, P. J. J. F., Weerts, A. H., and Uijlenhoet, R.: Generating spatial precipitation ensembles: impact of temporal correlation structure, Hydrol. Earth Syst. Sci. Discuss., 9, 3087-3127, doi:10.5194/hessd-9-3087-2012, 2012a.*

*Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435-3449, doi:10.5194/hess-16-3435-2012, 2012b.*

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

Manuscript prepared for Hydrol. Earth Syst. Sci.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 5 June 2015

# Operational aspects of asynchronous filtering for ~~hydrological~~ flood forecasting

**O. Rakovec**[1,3], **A. H. Weerts**[1,2], **J. Sumihar**[2], **and R. Uijlenhoet**[1]

[1]Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University, the Netherlands
[2]Deltares, P.O. Box 177, 2600 MH Delft, the Netherlands
[3]UFZ – Helmholtz Centre for Environmental Research, Leipzig, Germany

*Correspondence to:* O. Rakovec (oldrich.rakovec@ufz.de)

**Abstract.** This study investigates the suitability of the Asynchronous Ensemble Kalman Filter (AEnKF) and a partitioned updating scheme for hydrological forecasting. The AEnKF requires forward integration of the model for the analysis and enables assimilation of current and past observations simultaneously at a single analysis step. The results of discharge assimilation into a grid-based hydrological model (using a soil moisture error model) for the Upper Ourthe catchment in the Belgian Ardennes show that including past predictions and observations in the data assimilation method improves the model forecasts. Additionally, we show that elimination of the strongly non-linear relation between the soil moisture storage and assimilated discharge observations from the model update becomes beneficial for improved operational forecasting, which is evaluated using several validation measures.

## 1 Introduction

Understanding the behaviour of extreme hydrological events and the ability of hydrological modellers to improve the forecast skill are distinct challenges of applied hydrology. Hydrological forecasts can be made more reliable and less uncertain by recursively improving initial conditions. A common way of improving the initial conditions is to make use of data assimilation (DA), a feedback mechanism or update methodology which merges model estimates with available real world observations (e.g., Evensen, 1994, 2009; Liu and Gupta, 2007; Reichle, 2008; Liu et al., 2012).

Data assimilation methods can be classified from different perspectives. Traditionally, we distinguish between sequential and variational methods. The sequential methods are used to correct model state estimates by assimilating observations, when they become available. Examples of sequential methods are the popular Kalman and particle filters (e.g., Moradkhani et al., 2005a, b; Weerts and El Serafy, 2006; Zhou et al., 2006). The variational methods on the other hand minimize a cost function over a simulation period, which incorporates the mismatch between the model and observations (e.g., Liu and Gupta, 2007).

A next distinction can be made between synchronous and asynchronous methods. Synchronous methods, also called three-dimensional (3-D), assimilate observations which correspond to the time of update. The Ensemble Kalman Filter (EnKF, e.g., Evensen, 2003) is a popular synchronous approach, which propagates an ensemble of model realizations over time and estimates the background error covariance matrix from the ensemble statistics. Asynchronous methods, also called four dimensional (4-D), refer to an updating methodology, in which observations being assimilated into the model originate from times different to the time of update (Evensen, 1994, 2009; Sakov et al., 2010). The Ensemble Kalman Smoother (EnKS) is a common example of an asynchronous method (e.g. Evensen and Van Leeuwen, 2000; Dunne and Entekhabi, 2006; Crow and Ryu, 2009; Li et al., 2013). The EnKS extends the EnKF by introducing additional information by propagating the contribution of future measurements backward in time. The EnKS reduces the error variance as compared to the EnKF for the past (Evensen, 2009). EnKS and EnKF are identical for forecasting (including nowcasting).

The essential difference between a smoother and a filter is that a smoother assimilates "future observations", while a filter assimilates "past observations". This implies that for operational forecasting purposes, we need a filter rather than a smoother. A smoother can help improve the model accuracy in the past (e.g. for re-analysis), but it does not help improve forecast accuracy (Evensen, 2009). Therefore, Sakov et al. (2010) introduced the Asynchronous Ensemble Kalman Filter (AEnKF), which requires forward integration of the model to obtain simulated results necessary for the analysis and model updating at the analysis step using past observations over a time window. The difference among the EnKF, EnKS and AEnKF is schematized in Fig. 1.

Sakov et al. (2010) showed that the formulation of the EnKS provides a method for asynchronous filtering, i.e. assimilating past data at once, and that the AEnKF is a generalization of the ensemble-based data assimilation technique. Moreover, unlike the 4-D variational assimilation methods, the AEnKF does not require any adjoint model (Sakov et al., 2010). The AEnKF is particularly attractive from an operational forecasting perspective as more observations can be used with hardly any extra additional computational time. Additionally, such an approach can potentially account for a better representation of the time-lag between the internal model states and the catchment response in terms of the discharge.

Discharge represents a widely used observation for assimilation into hydrological models, because it provides integrated catchment wetness estimates and is often available at high temporal resolution (Pauwels and De Lannoy, 2006; Teuling et al., 2010). Therefore, discharge is a popular variable in data assimilation studies used for model state updating (e.g., Weerts and El Serafy, 2006; Vrugt and Robinson, 2007; Blöschl et al., 2008; Clark et al., 2008; Komma et al., 2008; Pauwels and De Lannoy, 2009; Noh et al., 2011a; Pauwels et al., 2013) or dual state-parameter updating (e.g. Moradkhani et al., 2005b; Salamon and Feyen, 2009; Noh et al., 2011b).

The Kalman-type of assimilation methods was developed for an idealized modelling framework with perfect linear problems with Gaussian statistics, however, it has been demonstrated to work well for a large number of different nonlinear dynamical models (Evensen, 2009). It remains interesting to evaluate whether elimination of the non-linear nature from the model updating can be beneficial. For example, Xie and Zhang (2013) introduced the idea of a partitioned update scheme to reduce the degrees of freedom of the high-dimensional state-parameter estimation of a distributed hydrological model. In their study, the partitioned update scheme enabled to better capture covariances between states and parameters, which prevented spurious correlations of the non-linear relations in the catchment response. Similarly, decreasing the number of model states being perturbed and updated was suggested by McMillan et al. (2013) to increase the efficiency of the filtering algorithm while conserving the forecast quality. Such an approach was proposed especially to states with small innovations, which in their case was mainly the soil moisture storage.

In this study we present a follow-up of the work of Rakovec et al. (2012b), in which discharge observations were assimilated into a grid-based hydrological model for the Upper Ourthe catchment in the Belgian Ardennes by using the EnKF. Here we scrutinize the applicability of the AEnKF using the same updating frequency (i.e. the same computational costs) as in the previous study. To our knowledge this is the first application of the AEnKF in a ~~hydrological~~ flood forecasting context. Firstly, the effect of assimilating past asynchronous observations on the forecast accuracy is analyzed. Secondly, the effect of a partitioned updating scheme is scrutinized.

## 2 Material and methods

### 2.1 Data and hydrological model

We carried out the analyses for the Upper Ourthe catchment upstream of Tabreux (area $\sim 1600\,\mathrm{km}^2$, Fig. 2), which is located in the hilly region of the Belgian Ardennes, Western Europe (Driessen et al., 2010). We employed a grid-based spatially distributed HBV-96 model (Hydrologiska Byråns Vattenbalansavdelning; Lindström et al., 1997), with spatial resolution of $1\,\mathrm{km} \times 1\,\mathrm{km}$ and hourly temporal resolution. The model is forced using deterministic spatially distributed rainfall fields, which were obtained by inverse distance interpolation from about 40 rain gauges measuring at hourly time step. Evaluation of the benefits of different rainfall interpolation techniques was deemed beyond the scope of the study. We used a method used in operational practice as this study is also oriented towards operational benefits of asynchronous filtering. Additionally, there are six discharge gauges (hourly time step) situated within the catchment, from which some are used for discharge assimilation and some for independent validation.

For a more detailed description of the catchment and model structure and definition of the hydrological states and fluxes, we refer to Rakovec et al. (2012b) and to Fig. 3. Briefly, for each grid cell the model considers the following model states: (1) snow (SN), (2) soil moisture (SM), (3) upper zone storage (UZ) and (4) lower zone storage (LZ). The dynamics of the model states are governed by the following model fluxes: rainfall, snowfall, snowmelt, actual evaporation, seepage, capillary rise, direct runoff, percolation, quick flow and base flow. The latter two fluxes force the kinematic wave model (Chow et al., 1988; PCRaster, 2014). This routing scheme calculates the overland flow using two additional model states, the water level ($H$) and discharge ($Q$) accumulation over the drainage network. Model parameterization is based on the work of Booij (2002) and van Deursen (2004).

In contrast to Rakovec et al. (2012b), in the current study we employed the HBV-96 model built within a recently developed open source modelling environment OpenStreams (2014), which is suitable for integrated hydrological modelling based on the Python programming language with the PCRaster spatial processing engine (Karssenberg et al., 2009; PCRaster, 2014). The advantage of using OpenStreams (2014) is that it enables direct communication with OpenDA (2014), an open source data assimilation toolbox. OpenDA (2014) provides a number of algorithms for model calibration and assimilation and is suitable to be connected to any kind of environmental model (e.g., Ridler et al., 2014).

The import and export of hydrological and meteorological data to the system is done using Delft Flood Early Warning System (Delft-FEWS, Werner et al., 2013), an open shell system for managing forecasting processes and/or handling time series data. Delft-FEWS is a modular and highly configurable system, which is used by the Dutch authorities for the flood forecasting for the River Meuse basin (called RWsOS Rivers), in which the Upper Ourthe is located. The current configuration is a stand-alone version of RWsOS Rivers, however, it can be easily switched into a configuration with real-time data import.

### 2.2 Data assimilation for model initialization

As stated in the introduction, we investigate the potential added value of the Asynchronous EnKF (AEnKF) (Sakov et al., 2010) as compared to the traditional (synchronous) EnKF for operational flood forecasting. The derivation of the AEnKF (Sect. 2.2.2) is based on the equations using the same updating frequency (i.e., same computational costs, different number of observations) as for the EnKF (Sect. 2.2.1), as among others presented by Rakovec et al. (2012b).

#### 2.2.1 Ensemble Kalman Filter (EnKF)

First, we define a dynamic state space system as

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}, \mathbf{u}_{k-1}) + \boldsymbol{\omega}_k, \tag{1}$$

where $\boldsymbol{x}_k$ is a state vector at time $k$, $f$ is an operator (hydrological model) expressing the model state transition from time step $k-1$ to $k$ in response to the model input $\mathbf{u}_{k-1}$ and time-invariant model parameters $\boldsymbol{\theta}$. The noise term $\boldsymbol{\omega}_k$ is assumed to be Gaussian white noise (i.e., independent of time). It incorporates the overall uncertainties in model structure, parameters and model inputs.

Second, we define an observation process as

$$\boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{\nu}_k, \tag{2}$$

where $\boldsymbol{y}_k$ is an observation vector derived from the model state $\boldsymbol{x}_k$ and the model parameters through the $h$ operator (in our case the kinematic wave routing model generating discharge). The noise term $\boldsymbol{\nu}_k$ is additive observational Gaussian white noise, with zero mean and covariance $\mathbf{R}_k$. For independent measurement errors, $\mathbf{R}_k$ is diagonal. Note that both the kinematic wave routing model $h(.)$ as well as the hydrological model $f(.)$ exhibit nonlinear behavior.

After the model update at time $k-1$, the model is used to forecast model states at time $k$ (Eq. 1). The grid-based model states form a matrix, which consists of $N$ state vectors $\boldsymbol{x}_k$ corresponding to $N$ ensemble members:

$$\mathbf{X}_k = (\boldsymbol{x}_k^1, \boldsymbol{x}_k^2, \ldots, \boldsymbol{x}_k^N), \tag{3}$$

where

$$\boldsymbol{x}_k^i = (\mathrm{SN}_{1:m}^i, \mathrm{SM}_{1:m}^i, \mathrm{UZ}_{1:m}^i, \mathrm{LZ}_{1:m}^i, H_{1:m}^i, Q_{1:m}^i)_k^T, \tag{4}$$

$\mathrm{SN}^i$, $\mathrm{SM}^i$, $\mathrm{UZ}^i$, $\mathrm{LZ}^i$, $H^i$ and $Q^i$ are the HBV-96 model states of the $i$th ensemble member (Sect. 2.1), $m$ gives the number of grid cells and $T$ is the transpose operator. The ensemble mean

$$\overline{\boldsymbol{x}}_k = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_k^i \tag{5}$$

is used to approximate the forecast error for each ensemble member:

$$\mathbf{E}_k = (\boldsymbol{x}_k^1 - \overline{\boldsymbol{x}}_k, \boldsymbol{x}_k^2 - \overline{\boldsymbol{x}}_k, \ldots, \boldsymbol{x}_k^N - \overline{\boldsymbol{x}}_k). \tag{6}$$

The ensemble estimated model covariance matrix $\mathbf{P}_k$ is defined as

$$\mathbf{P}_k = \frac{1}{N-1}\mathbf{E}_k\,\mathbf{E}_k^T. \tag{7}$$

When observations become available, the model states of the $i$th ensemble member are updated as follows:

$$\boldsymbol{x}_k^{i,+} = \boldsymbol{x}_k^{i,-} + \mathbf{K}_k(\boldsymbol{y}_k - h(\boldsymbol{x}_k^{i,-}) + \boldsymbol{\nu}_k^i), \tag{8}$$

where $\boldsymbol{x}_k^{i,+}$ is the analysis (posterior, or update) model state matrix and $\boldsymbol{x}_k^{i,-}$ is the forecast (prior) model state matrix. $\mathbf{K}_k$ is the Kalman gain, a weighting factor of the errors in model and observations:

$$\mathbf{K}_k = \mathbf{P}_k\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T + \mathbf{R}_k)^{-1}, \tag{9}$$

where $\mathbf{P}_k\mathbf{H}_k^T$ is approximated by the forecasted covariance between the model states and the forecasted discharge at the observing locations, and $\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T$ is approximated by the covariance of forecasted discharge at the observing locations (Houtekamer and Mitchell, 2001):

$$\mathbf{P}_k\mathbf{H}_k^T = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_k^i - \overline{\boldsymbol{x}_k})(h(\boldsymbol{x}_k^i) - \overline{h(\boldsymbol{x}_k)})^T, \tag{10}$$

$$\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T = \frac{1}{N-1}\sum_{i=1}^{N}(h(\boldsymbol{x}_k^i) - \overline{h(\boldsymbol{x}_k)})(h(\boldsymbol{x}_k^i) - \overline{h(\boldsymbol{x}_k)})^T, \tag{11}$$

where

$$\overline{h(\boldsymbol{x}_k)} = \frac{1}{N}\sum_{i=1}^{N} h(\boldsymbol{x}_k^i). \tag{12}$$

### 2.2.2 Asynchronous Ensemble Kalman Filter (AEnKF)

The AEnKF should not be considered as a new method, but rather a simple modification of the (synchronous) EnKF (Sect. 2.2.1) using a state augmentation approach. This means that the $i$th vector of model states ($\boldsymbol{x}_k^i$) at time $k$ (see Eq. 4) is augmented with the past forecasted observations $h(\boldsymbol{x}_{k-1}^i),\ldots,h(\boldsymbol{x}_{k-W}^i)$ (i.e., model outputs corresponding to the observation locations) from $W$ previous time steps, which yields

$$\tilde{\boldsymbol{x}}_k^i = \begin{pmatrix} \boldsymbol{x}_k^i \\ h(\boldsymbol{x}_{k-1}^i) \\ h(\boldsymbol{x}_{k-2}^i) \\ \vdots \\ h(\boldsymbol{x}_{k-W}^i) \end{pmatrix}. \tag{13}$$

Remember that the size of $\boldsymbol{x}_k^i$ and $h(\boldsymbol{x}_{k-1}^i),\ldots,h(\boldsymbol{x}_{k-W}^i)$ can significantly differ: $\boldsymbol{x}_k^i$ contains the complete set of model states, while $h(\boldsymbol{x}_{k-1}^i),\ldots,h(\boldsymbol{x}_{k-W}^i)$ contains only the forecasted observations. Additionally, with the new state definition comes a new augmented observer operator $\tilde{\boldsymbol{h}}_k$ (in which $\boldsymbol{I}$, with the corresponding subscript, stands for identity elements on the diagonal matching the dimensions in Eq. 13), a new augmented observation vector $\tilde{\boldsymbol{y}}_k$ and its corresponding observation covariance matrix $\tilde{\mathbf{R}}_k$:

$$\tilde{\boldsymbol{h}}_k = \begin{pmatrix} \boldsymbol{h}_k & & & & \\ & \boldsymbol{I}_{k-1} & & 0 & \\ & & \boldsymbol{I}_{k-2} & & \\ & 0 & & \ddots & \\ & & & & \boldsymbol{I}_{k-W} \end{pmatrix}, \tag{14}$$

$$\tilde{\boldsymbol{y}}_k = \begin{pmatrix} \boldsymbol{y}_k \\ \boldsymbol{y}_{k-1} \\ \boldsymbol{y}_{k-2} \\ \vdots \\ \boldsymbol{y}_{k-W} \end{pmatrix}, \tag{15}$$

$$\tilde{\mathbf{R}}_k = \begin{pmatrix} \mathbf{R}_k & & & & \\ & \mathbf{R}_{k-1} & & 0 & \\ & & \mathbf{R}_{k-2} & & \\ & 0 & & \ddots & \\ & & & & \mathbf{R}_{k-W} \end{pmatrix}. \tag{16}$$

Having these augmented equations for $\tilde{\boldsymbol{x}}_k^i$, $\tilde{\boldsymbol{h}}_k$, $\tilde{\boldsymbol{y}}_k$ and $\tilde{\mathbf{R}}_k$, it is straightforward to carry out the assimilation in the same manner as presented in Sect. 2.2.1. Note that although current and past observations are used to construct the augmented state vector in the Eq. (13), in practice Eq. (8) is solved only to the current state $\tilde{x}_k^i$ (i.e. the indices that correspond to $\boldsymbol{x}_k^i$) and the rest is ignored. The presence of past observation terms increases the dimension of $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{K}}_k$ (see Eqs. 7 and 9) in both directions (rows and columns). Each column of $\tilde{\mathbf{K}}_k$ corresponds to an observation. The extra column of $\tilde{\mathbf{K}}_k$ corresponds to the past observations. Hence, it is possible to simply solve the equations for the first rows, which correspond only to $\boldsymbol{x}_k^i$. Note that the first rows of $\tilde{\mathbf{K}}_k$ also contain the contributions of the past observations to the current state. These contributions arise from the off-diagonal terms of the augmented covariance $\tilde{\mathbf{P}}_k$. Finally, if the time window equals the current single time step, then $W = 0$ and the AEnKF problem reduces to the traditional EnKF.

From the operational point of view, it is preferable to have a longer assimilation window, because less frequent assimilation eliminates a disruption of the ensemble integration by an update and a restart. When assimilation is done more frequently, it will cause considerably higher calculation costs, which can often be a burden for real-time operational settings (Sakov et al., 2010). The AEnKF uses a longer assimilation window and assimilates all observations in a single update. This makes the AEnKF attractive to be used. The added value of a longer assimilation window will be a subject for investigation in this work. Especially, it can provide an improved representation of the time-lag between the internal model states and the catchment response in terms of the discharge. Such an idea was investigated for example by Li et al. (2013), who compared the effect of time-lag representation using the EnKF and EnKS.

## 2.3 Model uncertainty

In this study, we assume the source of model uncertainty to be the HBV soil moisture, which provides boundary conditions for surface runoff and represents interaction from interception, evapotranspiration, infiltration and input uncertainty by rainfall. The uncertainty is represented as a noise term $\boldsymbol{\omega}$ as in Eq. (1). Based on expert knowledge, the noise is modelled as an autoregressive process of order 1 with a de-correlation time length of 4 h. The noise process is further assumed spatially isotropic with a spatial de-correlation length of 30 km. The noise is assumed to have a spatially uniform SD of 1 mm. The 2-D noise fields with such statistics were generated by using OpenDA (2014) toolbox. This parameterization of the noise model ensures that the ensemble spread in the simulated discharge corresponds well with the control simulations as presented by Rakovec et al. (2012b) (not shown). Ideally, all sources of uncertainty should be accounted for in a DA scheme. However, this is not yet a common approach in operational hydrologic data assimilation. Moreover, as the objective of the current manuscript is to compare the operational benefits of application of the AEnKF, we kept the noise model relatively simple. For more work on the effect of noise specification on DA using complex spatially distributed hydrological models see, Noh et al. (2014).

### 2.4 Experimental setup

This section provides a configuration setup of the filtering methods (Sect. 2.2.1 and 2.2.2) to assimilate discharge observations into a spatially distributed hydrological model of the Upper Ourthe catchment. The objective is to improve the hydrological forecast at the catchment outlet (at Tabreux, gauge 1 in Fig. 2) by assimilating up to four discharge gauges, numbered as 1, 3, 5, 6 in Fig. 2. Note that discharge data from multiple gauges are assimilated simultaneously and no localization is employed in this study. Additionally, validation at an independent location is also performed. The discharge assimilation is done every 24 h, however, the forecasts are issued every 6 h, i.e. 4 times a day, with different independent starting points at 00:00, 06:00, 12:00, 18:00 UTC, which is the same implementation as used by Rakovec et al. (2012b). This study analyses the 8 largest floods peaks observed within the catchment since 1998. An overview is provided in Table 1.

The ensemble of uncertain model simulations is obtained by perturbing the soil moisture state (SM) with the spatio-temporally correlated error model (Sect. 2.3). With this approach we ensured that the error model produced reasonable results in the open-loop and did not lead to any numerical instability. More complex ways of perturbing the model and their effects on forecast accuracy were studied before (see Rakovec et al., 2012a; Noh et al., 2014) and were deemed beyond the scope of this manuscript. The ensemble size in this study was defined to be 36 realizations (for computational reasons). Note that increased ensemble sizes of 72 and 144 realizations did not influence the results (not shown). Nevertheless, such a small ensemble size as presented in the manuscript would not be possible if parameter estimation would be involved or if more complex error models would be employed. The error in the discharge observations is considered to be a normally distributed observation error with a variance of $(0.1\,Q_{\mathrm{obs}},k)^2$ (after e.g. Weerts and El Serafy, 2006; Clark et al., 2008).

The experimental setup scrutinizes the problem of asynchronous filtering from two perspectives. First, we investigate the effect of state augmentation using the past observations and assimilation of distributed observations on the state innovation (Sect. 3.1). Recall that the number of observations being assimilated into the model depends on the magnitude of W. Furthermore, the choice of which model states are included in the analysis step to be updated is analysed (Sects. 3.2, and 3.3). This means that besides updating all of the model states, we will test two other alternatives. The first alternative will leave out from the model analysis the soil moisture state (noSM), which is known to exhibit the most non-linear relation to Q. The second alternative will eliminate all the model states except for the two routing ones (HQ). The scenarios of the partitioned state updating schemes are shown in Table 2, including the control run without state updating (no update).

The performance of the data assimilation procedure regarding discharge forecasting is evaluated using the Ensemble Verification System (EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrological variables at discrete locations (Brown et al., 2010), which provides a number of probabilistic verification measures. In this study we used three popular measures: the root mean square error (RMSE), the relative operating characteristic (ROC) score and the Brier skill (BS) score. We refer to e.g. Wilks (2006); Brown et al. (2010); Brown and Seo (2013), and Verkade et al. (2013) for exact definitions of these measures. In summary, the perfect forecast in terms of the RMSE has a value of 0, while positive values indicate errors in the same units as the variable. The perfect forecast in terms of the ROC and BS scores has a value of 1 and values smaller than 1 indicate forecast deterioration.

## 3 Results

### 3.1 The effect of state augmentation and distributed observations on state innovation

To investigate and understand the effect of augmented operators (Eqs. 13, 14, and 15) on the innovation of spatially distributed model states, we present the following example. Figure 4 shows discharge simulations and corresponding discharge observations at 4 locations within the catchment on 31 December 2002, 00:00 UTC. Note that the magnitude of the discharge observations is a function of the location within the catchment; for downstream gauges the magnitude is larger than for the more upstream gauges. The discharge observations are further distinguished according to the time window length of the state augmentation, which is set to $W = 0$ and $W = 11$. The first example represents the traditional EnKF algorithm, while the latter assimilates observations from a 12 h time window (i.e., 1 current observation and 11 past observations), which is arbitrarily defined as a half of the 24 h assimilation time window. For some cases alternative assimilation windows were tested, which did not lead to noticeable differences however (not shown). Note that the amount of information being assimilated into the model differs for different values of $W$.

The mean difference between the forecasted and updated model states for the whole ensemble is illustrated in Fig. 5 for four scenarios. These examples improve our understanding about the behaviour of the updated model states in relation to the information content of the observations from two perspectives: (1) the effect of assimilating also past observations in addition to observations at the current (analysis) time, and (2) the effect of assimilating spatially distributed observations into a grid-based hydrological model.

Let us first consider the traditional EnKF (i.e., no state augmentation with $W = 0$) to update all the grid-based model states by assimilating the observation at the catchment outlet (gauge 1). We observe that the single observation is measured approximately in the middle of the simulated ensemble (see the open circle for gauge 1 in Fig. 4). Therefore, there is hardly any difference between the forecasted and updated model states as we show in Fig. 5a. In the second scenario, we still assimilate only one gauge at the outlet, however, we use the augmented operators with $W = 11$. Because the mean of the ensemble simulations is predominantly underestimated as compared to the assimilated observations (see black dots in Fig. 4 for gauge 1), after the update more water is added spatially equally into the system, as shown in Fig. 5b. In the third scenario, we include all 4 gauges being assimilated into the model without any augmentation. Because the model simulations at the interior gauges are mostly overestimating the observations, water is removed from the catchment during the update. Moreover, since the model overestimation is largest at gauges 3 and 6, we can also observe in Fig. 5c how well the EnKF is capable of identifying corresponding regions in a spatial manner. In the fourth scenario (Fig. 5d) we still assimilate all 4 gauges, however we augment the state with $W = 11$. We can observe that the innovation of the model states gets even more spatially differentiated; the updated SM and UZ model states in the downstream part of the catchment increase the amount of water in the system, while the updated SM and UZ model states in the upstream part decrease the amount of water in the system.

The presented educational examples showed an update for several scenarios starting from the same initial conditions. This enabled a fair comparison between scenarios, however, the sensitivity of state augmentation needs to be further scrutinized in terms of its cumulative effect over time.

### 3.2 The effect of the four partitioned update schemes and asynchronous assimilation on forecast accuracy

We present a qualitative interpretation of the hydrological forecasts with a lead time of 48 h in Fig. 6 for different partitioned state updating schemes as defined in Table 2, including both a non-augmented state ($W = 0$) and an augmented state ($W = 11$). This analysis focuses on a characteristic winter flood event (December 2002–January 2003) being typical for a moderate temperate climate caused by a fast-moving frontal stratiform system (Hazenberg et al., 2011). We observe that the ensemble of the control runs (top panel of Fig. 6) simulates the major flood peak reasonably well, including the timing and the magnitude~~.~~, however, it has a larger spread with respect to the assimilation scenarios. Additionally, when we consider the ensemble mean of the no-update scenario with respect to the assimilation scenarios, the accuracy deteriorates. When discharge assimilation is employed, an overall reduction of the uncertainty in the forecasted ensemble is observed. Nevertheless, the forecasted flood peak becomes underestimated and the forecasted recession remains overestimated, which is acceptable because of the defined uncertainty in the observed discharge. This happens in particular for the scenario in which all states are updated; there are marginal differences between the non-augmented and augmented model states. Furthermore, when we leave out SM from the state update (noSM), we can observe that the major flood peak is forecasted more accurately, including the rising limb around 31 December 2002. Moreover, for the augmented state with $W = 11$, the ensemble spread becomes somewhat wider for lead times exceeding 12 h than for the non-augmented state. Nevertheless, the observations correspond approximately with the ensemble mean. Finally, we present the effect of the scenario in which only the two routing states are updated. The results suggest that the flood peak is captured most accurately of all scenarios, however with somewhat wider uncertainty bands. Therefore, it seems more appropriate to exclude the UZ storage (noSM scenario) in the model state updating, which represents water storage available for quick catchment response in the concept of the HBV model.

Besides a qualitative interpretation of the forecasted hydrographs presented in Fig. 6 for one particular event, we summarize these results in a more quantitative manner for the whole set of 8 flood events (see Table 1) ~~in terms of the root-mean-square-error (RMSE) vs. lead time. This is shown in Fig.~~ using three statistical measures with respect to the lead time. Figure 7 shows the average behavior (over many forecasts) of an improved initial state on the forecast accuracy for the different filter settings, although individual partial updates may vary in time. In general, the improvements in forecast accuracy decay with lead time in a systematic fashion as is to be expected.

Figure. 7a shows the root-mean-square-error (RMSE) as a function of lead time for different partitioned state updating schemes and for three scenarios for the state augmentation at the catchment outlet (Tabreux). The control model run with no update has a constant RMSE of about $32\,\mathrm{m^3\,s^{-1}}$ and an improved hydrological forecast has a RMSE lower than the control run. The results suggest that all assimilation scenarios improve the hydrological forecast, however with marked differences between the scenarios.

~~We can~~ Figure 7a also clearly shows that the differences in the forecast improvement of these various setups are purely due to using multiple data points in the past at the analysis step. We can further observe that updating all model states except for SM (noSM scenario) consistently leads to the most accurate forecasts across the whole range of lead times. Additionally, state augmentation using $W = 5$ and $W = 11$ indicates improvements compared to the case without augmentation ($W = 0$). However, for lead times longer than the travel time from the most upstream gauges to the outlet (i.e. exceeding 20 h), the difference between state augmentations $W = 5$ and $W = 11$ diminishes. Moreover, when only the two routing states (HQ scenario) are updated, the RMSE is lowered for short lead times, but the improved effect does not last as long as for the noSM scenario. The smallest improvement at shorter lead times is achieved when all model states are updated (scenario all). This is due to the strongly non-linear relation between the assimilated observations and the SM storage, which is further articulated by the time-lag between the state and the catchment response. Nevertheless, for longer lead times it seems slightly better to update all states rather than only the routing states. Discharge is related to the SM and UZ storages through the Kalman gain. When the correlation is lower the update will be smaller. AEnKF exploits the correlation between the present discharge state and the discharge state not only at the previous time step, but also further in the past. It may be possible to use the correlation between discharge at the present time and UZ/SM in the past for data assimilation, however, this is deemed beyond the scope of this study. Nevertheless, we speculate that this will only be useful in a smoothing context (i.e. the present discharge may bring information on UZ/SM in the past), not in a filtering context as in the present study.

Validation of the model setup in terms of the RMSE is presented in Fig. 7d for an independent evaluation of the forecasting results at Durbuy, an interior location, which was not used for assimilation. These results show that an improvement of discharge assimilation also occurs at the validation location and that the pattern corresponds well to the results presented in Fig. 7a. Such an analysis indicates that there is no spurious update of the model states.

To present the results in a more robust way, we also analyzed them (at Tabreux) in terms of other probabilistic verification measures: the relative operating characteristic (ROC) score and the Brier skill (BS) score (see Fig. 7b, c). Recall that values of 1 represent a perfect forecast, while values smaller than 1 indicate forecast deterioration. Similar to the RMSE results, updating only the two routing states (HQ) is most efficient for short lead times, but this skill disappears quickly for longer lead times. In terms of the ROC and BS scores, for a given augmentation size, there are marginal differences between the scenarios which update all states (all) and which leave the soil moisture out (noSM). However, it is notable that the state augmentation case ($W = 11$) improves the forecast performance as compared to the no augmentation case ($W = 0$). Note that the state augmentation of $W = 5$ was not carried out.

### 3.3 Temporal nature of model state innovations

To reveal the temporal nature of the model being updated using the AEnKF, using $W = 0$ and $W = 11$, we present in Fig. 8a and b time series of normalized differences between the ensemble means for the 3 partitioned update schemes and the ensemble mean for the no update scenario. The normalization is achieved by dividing the aforementioned difference by the no update scenario mean. In such a way we obtain the relative change in each of the model states. For the AEnKF using $W = 0$ (Fig. 8a), we can observe that for the scenario "all", which updates all the model states, the magnitude of the percentage change is approximately the same for all 4 model states and ranges up to 25 %. When all model states except for the SM are updated, no changes in the SM storage occur and the overall magnitude of the changes in the other states is slightly decreased and smoothed. Furthermore, when only the two routing states are updated (HQ), the SM and UZ storages remain constant over time and we observe a different temporal behaviour of the routing states in comparison with the previous cases. For the HQ scenario, the updated time series have a clear zigzag shape, which indicates that the effect of updating diminishes faster, because only the river channel is updated. In contrast, the routing states for the other cases show a more stable behaviour over time, illustrated by the stepwise shape. These more persistent results correspond to the updates in the UZ storage, which is used for a quick catchment response and has an impact for a longer time. The benefits of including the UZ storage in the update and leaving the SM storage out was already presented from a different point of view in Fig. 7a for longer lead times.

For the AEnKF using $W = 11$ (Fig. 8b), we can observe that the overall pattern of the temporal changes in the model states is similar as for $W = 0$, but the behaviour of using $W = 11$ shows somewhat larger variability. By assimilating more observations ($W = 11$), we expect even a larger update, assuming that more observations contain more information about the unknown truth. Assuming the underlying forecast model has a significant error, by assimilating more observations the Kalman filter will pull the model even closer to the truth, yielding a larger abrupt update.

## 4 Conclusions

We applied the Asynchronous Ensemble Kalman Filter (AEnKF) (Sakov et al., 2010) and identified the effect of augmenting the state vector with past simulations and observations. To our knowledge this is the first application of the AEnKF in ~~hydrological~~ flood forecasting. We showed that the effect of an augmented assimilation vector improves the ~~hydrological~~ flood forecasts, but the contribution gets smaller for longer lead times. Overall, the AEnKF can be considered as an effective method for model state updating taking into account more (e.g. all) observations at hardly any additional computational burden. This makes it very suitable for operational hydrological forecasting. When compared to standard EnKF, the AEnKF allows the choice of a certain assimilation window length, which adds a degree of freedom to the data assimilation scheme. The optimal window is very likely related to the catchment size (i.e. concentration time). It was noted (not shown) that for the smaller upstream catchments the optimal window was smaller than for the complete Upper Ourthe catchment, although there was no negative effect of a longer assimilation window ($W = 5$ vs. $W = 11$). For the high flows analysed in this study, the AEnKF with a longer time window $W$ is able to make corrections that last longer on average than with the shorter time window $W$. Characterization of the statistical properties of the temporal flow dynamics (i.e., typical time scales of flood peaks as compared to low flows) is however a relevant issue. The length of the time window $W$ has to be seen relative to the time scale of the river flow dynamics. We assume that for low flow conditions, the improved skill of longer $W$ with respect to shorter $W$ will become negligible, as low flows exhibit less temporal dynamics than high flows. We refer to Pan and Wood (2013) for an analysis about explicit handling of lags in space and time, which uses a state augmentation approach for a linear inverse streamflow routing model. Note that it was not the objective of this study to determine the optimal assimilation window for the AEnKF ~~.~~ given various river flow dynamics. Another limitation of this study is the relatively simple error model for perturbing only soil moisture states. More complex ways of perturbing the model and their effects on forecast accuracy deserve more attention in future studies.

We investigated the effect of a partitioned update scheme recently suggested by Xie and Zhang (2013). We showed that for the Upper Ourthe catchment reducing the number of model states of a grid-based HBV model using AEnKF can lead to better forecasts of the discharge. In terms of the root-mean-square-error, the largest improvements in the forecast accuracy were observed for the scenario where the soil moisture was left out from the analysis (similar to the PDM updating scheme presented by Moore, 2007). This indicates that elimination of the strongly non-linear relation between the soil moisture storage (SM) and assimilated discharge observations can become beneficial for an improved forecast when soil moisture observations are not considered. On the other hand, it was recently demonstrated that ~~hydrological model in operational settings~~ a rainfall-runoff model can be improved when constrained by remotely sensed soil moisture (e.g., Alvarez-Garreton et al., 2014; Wanders et al., 2014a, b) or in-situ soil moisture (e.g., Lee et al., 2011). Moreover, we showed that keeping the quick catchment response storage (upper zone; UZ) in the model analysis is important, especially for longer lead times, when compared to the scenario in which only two routing storages were updated. The UZ seems to compensate the effect of SM on discharge. The fact that excluding SM extends the improvements suggests that in our case the discharge forecasts with a lead time of two days (and for major flood events) are less dependent on SM. A possible alternative to excluding the SM storage from the analysis, would be to investigate the use of other algorithms, for example the Maximum Likelihood Ensemble Filter (MLEF) (Zupanski, 2005; Rafieeinasab et al., 2014), which is more suited for use with highly non-linear observation operators.

*Acknowledgements.* We would like to thank Arno Kockx and Martin Verlaan for their help with the OpenDA configuration and Jaap Schellekens for help with the OpenStreams configuration (all from Deltares). We thank Paul Torfs (Wageningen University)~~and~~, Seong Jin Noh (KICT, Korea), Ming Pan (Princeton University), two anonymous reviewers and the Editor for their comments on the manuscript. This project is financially supported by the Flood Control 2015 program (http://www.floodcontrol2015.nl), which is gratefully acknowledged. We thank the Hydrological Service of the Walloon Region of Belgium (MET-SETHY) and the Royal Meteorological Institute of Belgium (KMI) for providing the hydrological and meteorological data.

## References

Alvarez-Garreton, C., Ryu, D., Western, A., Crow, W., and Robertson, D.: The impacts of assimilating satellite soil moisture into a rainfall–runoff model in a semi-arid catchment, J. Hydrol., 519, 2763–2774, doi:10.1016/j.jhydrol.2014.07.041, 2014.

Blöschl, G., Reszler, C. R., and Komma, J.: A spatially distributed flash flood forecasting model, Environ. Modell. Softw., 23, 464–478, 2008.

Booij, M.: Appropriate modelling of climate change impacts on river flooding, Ph.D. thesis, University of Twente, Enschede, the Netherlands, 2002.

Brown, J. D. and Seo, D.-J.: Evaluation of a nonparametric post-processor for bias correction and uncertainty estimation of hydrologic predictions, Hydrol. Process., 27, 83–105, doi:10.1002/hyp.9263, 2013.

Brown, J. D., Demargne, J., Seo, D.-J., and Liu, Y.: The Ensemble Verification System (EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations, Environ. Modell. Softw., 25, 854–872, doi:10.1016/j.envsoft.2010.01.009, 2010.

Chow, V. T., Maidment, D., and Mays, L.: Applied Hydrology, McGraw-Hill, New York, USA, 1988.

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

Crow, W. T. and Ryu, D.: A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals, Hydrol. Earth Syst. Sci., 13, 1–16, doi:10.5194/hess-13-1-2009, 2009.

Driessen, T. L. A., Hurkmans, R. T. W. L., Terink, W., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model, Hydrol. Earth Syst. Sci., 14, 651–665, doi:10.5194/hess-14-651-2010, 2010.

Dunne, S. and Entekhabi, D.: Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment, Water Resour. Res., 42, 1–15, doi:10.1029/2005WR004334, 2006.

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

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003.

Evensen, G.: Data Assimilation: the Ensemble Kalman Filter, Springer, Berlin, Heidelberg, Germany, 2009.

Evensen, G. and Van Leeuwen, P.: An ensemble Kalman smoother for nonlinear dynamics, Mon. Weather Rev., 128, 1852–1867, 2000.

Hazenberg, P., Leijnse, H., and Uijlenhoet, R.: Radar rainfall estimation of stratiform winter precipitation in the Belgian Ardennes, Water Resour. Res., 47, 1–15, doi:10.1029/2010WR009068, 2011.

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

Karssenberg, D., Schmitz, O., Salamon, P., De Jong, C., and Bierkens, M. F. P.: A software framework for construction of process-based stochastic spatio-temporal models and data assimilation, Environ. Modell. Softw., 25, 1–14, doi:10.1016/j.envsoft.2009.10.004, 2009.

Komma, J., Blöschl, G., and Reszler, C.: Soil moisture updating by ensemble Kalman filtering in real-time flood forecasting, J. Hydrol., 357, 228–242, doi:10.1016/j.jhydrol.2008.05.020, 2008.

Lee, H., Seo, D.-J., and Koren, V.: Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: effects of uncertainties in the data and initial model soil moisture states, Adv. Water Resour., 34, 1597–1615, doi:10.1016/j.advwatres.2011.08.012, 2011.

Li, Y., Ryu, D., Western, A. W., and Wang, Q. J.: Assimilation of stream discharge for flood forecasting: the benefits of accounting for routing time lags, Water Resour. Res., 49, 1887–1900, doi:10.1002/wrcr.20169, 2013.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, 1997.

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: toward an integrated data assimilation framework, Water Resour. Res., 43, 1–19, doi:10.1029/2006WR005756, 2007.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, Hydrol. Earth Syst. Sci., 16, 3863–3887, doi:10.5194/hess-16-3863-2012, 2012.

McMillan, H. K., Hreinsson, E. Ö., Clark, M. P., Singh, S. K., Zammit, C., and Uddstrom, M. J.: Operational hydrological data assimilation with the recursive ensemble Kalman filter, Hydrol. Earth Syst. Sci., 17, 21–38, doi:10.5194/hess-17-21-2013, 2013.

Moore, R. J.: The PDM rainfall-runoff model, Hydrol. Earth Syst. Sci., 11, 483–499, doi:10.5194/hess-11-483-2007, 2007.

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

Moradkhani, H., Sorooshian, S., Gupta, H., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28, 135–147, 2005b.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S.: Applying sequential Monte Carlo methods into a distributed hydrologic model: lagged particle filtering approach with regularization, Hydrol. Earth Syst. Sci., 15, 3237–3251, doi:10.5194/hess-15-3237-2011, 2011.

Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S.: Dual state-parameter updating scheme on a conceptual hydrologic model using sequential Monte Carlo filters, J. Jpn. Soc. Civil Engin. B1, 67, I1–I6, 2011b.

Noh, S. J., Rakovec, O., Weerts, A. H., and Tachikawa, Y.: On noise specification in data assimilation schemes for improved flood forecasting using distributed hydrological models, J. Hydrol., 519, 2707–2721, doi:10.1016/j.jhydrol.2014.07.049, 2014.

OpenDA: The OpenDA data-assimilation toolbox, available at: http://www.openda.org, (last access: 1 November 2014), 2014.

OpenStreams: OpenStreams, available at: www.openstreams.nl, (last access: 1 November 2014), 2014.

Pan, M. and Wood, E. F.: Inverse streamflow routing, Hydrol. Earth Syst. Sci., 17, 4577–4588, doi:10.5194/hess-17-4577-2013, 2013.

Pauwels, V. R. N. and De Lannoy, G. J. M.: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, J. Hydrometeorol., 7, 458–477, doi:10.1175/JHM490.1, 2006.

Pauwels, V. R. N., and De Lannoy, G. J. M.: Ensemble-based assimilation of discharge into rainfall–runoff models: a comparison of approaches to mapping observational information to state space, Water Resour. Res., 45, 1–17, doi:10.1029/2008WR007590, 2009.

Pauwels, V. R. N., De Lannoy, G. J. M., Hendricks Franssen, H.-J., and Vereecken, H.: 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, doi:10.5194/hess-17-3499-2013, 2013.

PCRaster: PCRaster Environmental Modelling language, available at: http://pcraster.geo.uu.nl, (last access: 1 November 2014), 2014.

Rafieeinasab, A., Seo, D.-J., Lee, H., and Kim, S.: Comparative evaluation of Maximum Likelihood Ensemble Filter and Ensemble Kalman Filter for real-time assimilation of streamflow data into operational hydrologic models, J. Hydrol., 519, 2663–2675, doi:10.1016/j.jhydrol.2014.06.052, 2014.

Rakovec, O., Hazenberg, P., Torfs, P. J. J. F., Weerts, A. H. and Uijlenhoet, R.: Generating spatial precipitation ensembles: impact of temporal correlation structure, Hydrol. Earth Syst. Sci., 16, 3419–3434, doi:10.5194/hess-16-3419-2012, 2012a.

Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy, Hydrol. Earth Syst. Sci., 16, 3435–3449, doi:10.5194/hess-16-3435-2012, ~~2012.~~ 2012b.

Reichle, R. H.: Data assimilation methods in the Earth sciences, Water Resour., 31, 1411–1418, doi:10.1016/j.advwatres.2008.01.001, 2008.

Ridler, M. E., van Velzen, N., Hummel, S., Sandholt, I., Falk, A. K., Heemink, A., and Madsen, H.: Data assimilation framework: linking an open data assimilation library (OpenDA) to a widely adopted model interface (OpenMI), Environ. Modell. Softw., 57, 76–89, doi:10.1016/j.envsoft.2014.02.008, 2014.

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

Salamon, P. and Feyen, L.: Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter, J. Hydrol., 376, 428–442, 2009.

Teuling, A., Lehner, I., Kirchner, J., and Seneviratne, S.: Catchments as simple dynamical systems: experience from a Swiss prealpine catchment, Water Resour. Res., 46, 1–15, doi:10.1029/2009WR008777, 2010.

van Deursen, W.: Afregelen HBV model Maasstroomgebied, Tech. rep., Rapportage aan RIZA. Carthago Consultancy, Rotterdam, the Netherlands, 2004.

Verkade, J. S., Brown, J. D., Reggiani, P., and Weerts, A. H.: Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales, J. Hydrol., 501, 73–91, doi:10.1016/j.jhydrol.2013.07.039, 2013.

Vrugt, J. A. and Robinson, B. A.: Treatment of uncertainty using ensemble methods: comparison of sequential data assimilation and Bayesian model averaging, Water Resour. Res., 43, 1–15, doi:10.1029/2005WR004838, 2007.

Wanders, N., Bierkens, M. F. P., de Jong, S. M., de Roo, A., and Karssenberg, D.: The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models, Water Resour. Res., 50, 6874–6891, doi:10.1002/2013WR014639, 2014a.

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrol. Earth Syst. Sci., 18, 2343–2357, doi:10.5194/hess-18-2343-2014, 2014b.

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

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., and Heynert, K.: The Delft-FEWS flow forecasting system, Environ. Modell. Softw., 40, 65–77, doi:10.1016/j.envsoft.2012.07.010, 2013.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Elsevier, San Diego, 2006.

Xie, X. and Zhang, D.: A partitioned update scheme for state-parameter estimation of distributed hydrologic models based on the ensemble Kalman filter, Water Resour. Res., 49, 7350–7365, doi:10.1002/2012WR012853, 2013.

Zhou, Y., McLaughlin, D., and Entekhabi, D.: Assessing the performance of the ensemble Kalman filter for land surface data assimilation, Mon. Weather Rev., 134, 2128–2142, 2006.

Zupanski, M.: Maximum Likelihood Ensemble Filter: theoretical aspects, Mon. Weather Rev., 133, 1710–1726, doi:10.1175/MWR2946.1, 2005.

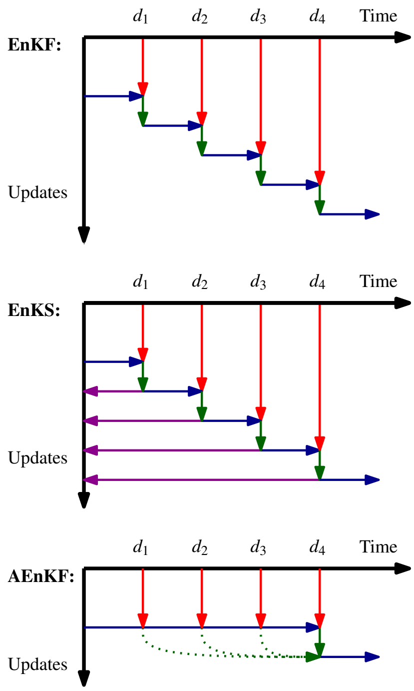


**Figure 1.** Illustration of the model updating procedure for the Ensemble Kalman Filter (EnKF), the Ensemble Kalman Smoother (EnKS), and the Asynchronous Ensemble Kalman Filter (AEnKF). The horizontal axis stands for time, observations ($d_1$, $d_2$, $d_3$, $d_4$) are given at regular intervals. The blue arrows represent forward model integration, the red arrows denote introduction of observations and green arrows indicate model update. The magenta arrows represent the model updates for the EnKS, therefore they go backward in time, as they are computed following the EnKF update every time observations become available. The green dotted arrows denote past observations being assimilated using the AEnKF. The schemes for the EnKF and the EnKS are after Evensen (2009).

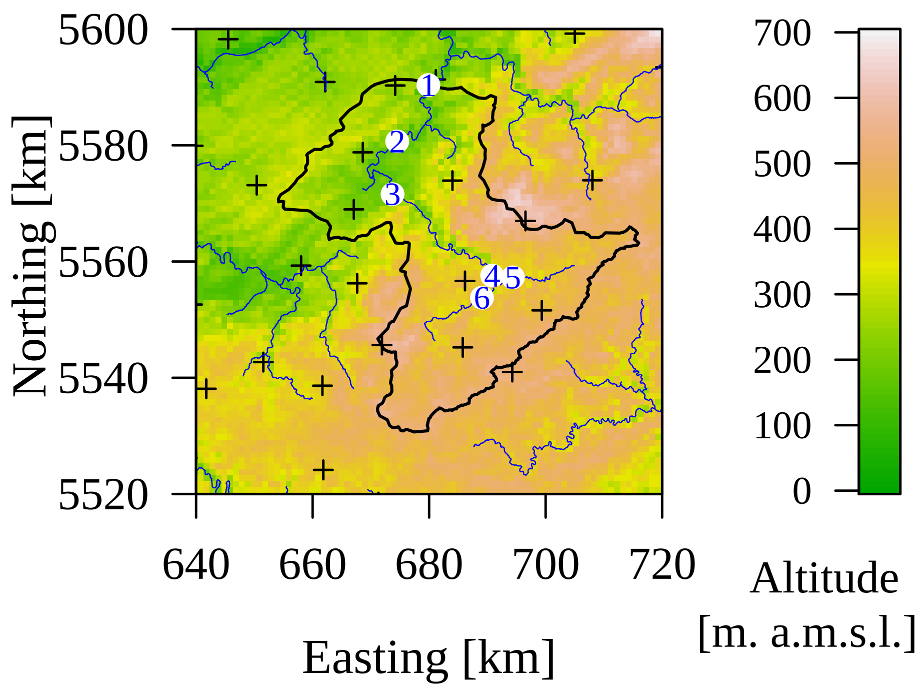


**Figure 2.** Topographic map of the Upper Ourthe (black line) including the river network (blue lines), rain gauges (plusses), six river gauges (white circles labeled with numbers: 1 = Tabreux, 2 = Durbuy, 3 = Hotton, 4 = Nisramont, 5 = Mabompre, 6 = Ortho). Projection is in the Universal Transverse Mercator (UTM) 31N coordinate system. After Rakovec et al. (2012b).

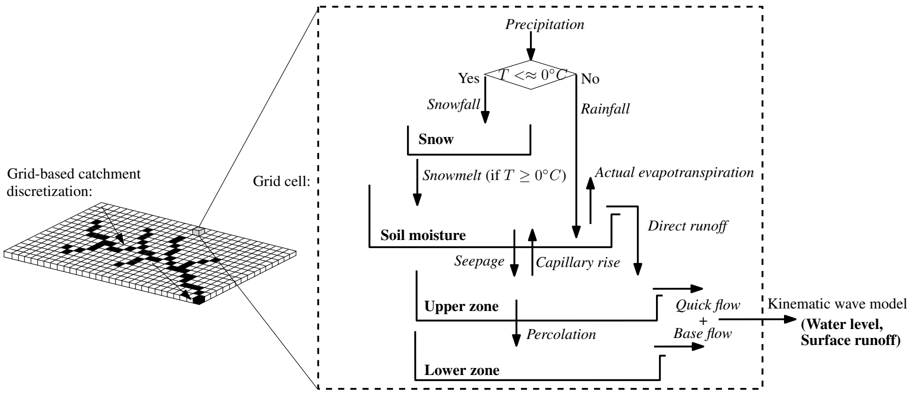


**Figure 3.** Left: catchment discretization using a grid-based approach including the channel delineation. Arrows indicate flow direction. Right: schematic structure of the HBV-96 model for each grid cell. Model states are in bold and model fluxes in italics (after Rakovec et al., 2012b).

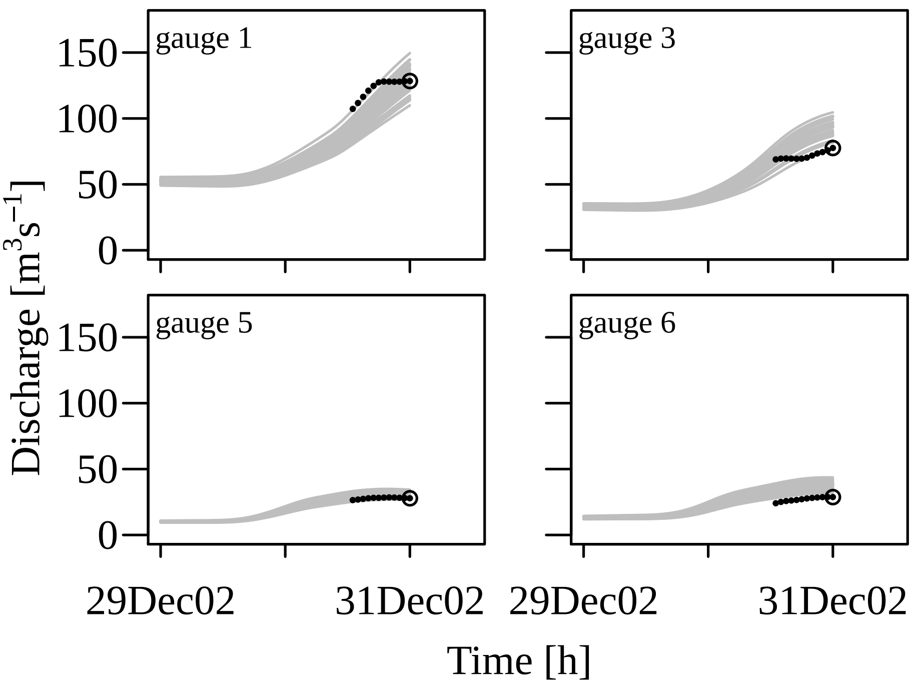


**Figure 4.** Discharge ensemble forecasts (grey lines) and observations (points) at four locations (gauges 1, 3, 5, 6; see Fig. 2). Observations being assimilated using the AEnKF are schematized according to the state augmentation size for two scenarios: assimilation of data from the current time step $W = 0$ (open circle, traditional EnKF approach) and assimilation of data including the previous 11 time steps, $W = 11$ (black dots). The observations are assimilated into the model states on 31 December 2002, 00:00 UTC.

**Table 1.** Overview of the periods used in this study.

| Period | Number of events | Maximum observed discharge [$m^3\ s^{-1}$] |
|---|---|---|
| 23 Oct 1998–15 Nov 1998 | 1 | 210 |
| 15 Feb 1999–5 Mar 1999 | 2 | 195 |
| 15 Jan 2002–6 Mar 2002 | 4 | 340 |
| 21 Dec 2002–7 Jan 2003 | 1 | 380 |

**Table 2.** Four partitioned state updating schemes (indicated in the first column) for 5 model states (indicated in the first row) being updated and thus included in the model analysis. Model states are described in Sect. 2.1 and Fig. 3 and have following acronyms: discharge ($Q$), water level ($H$), soil moisture storage (SM), snow storage (SN), upper zone storage (UZ), and lower zone storage (LZ).

| Name | $Q$ | $H$ | SM | SN | UZ | LZ |
|---|---|---|---|---|---|---|
| No update | | | | | | |
| all | √ | √ | √ | √ | √ | √ |
| noSM | √ | √ | | √ | √ | √ |
| HQ | √ | √ | | | | |

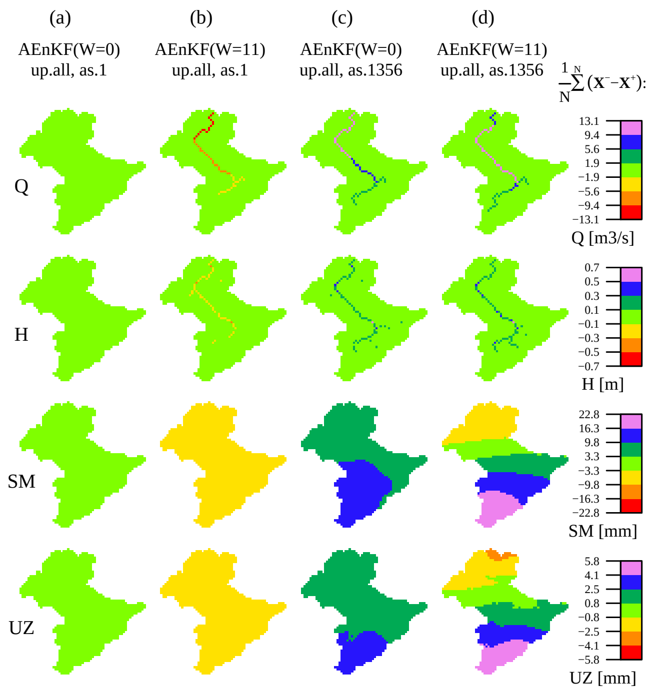


**Figure 5.** Mean difference between the forecasted ($\mathbf{X}^-$) and updated ($\mathbf{X}^+$) model states on 31 December 2002 at 00:00 UTC for different scenarios (shown in vertical panels). We show only 4 sensitive model states: discharge (Q), water level (H), soil moisture (SM) and upper zone (UZ). We dropped out the insensitive lower zone (LZ). Notation $W = 0$ and $W = 11$ indicates the size of the state augmentation. Notation up.all indicates that all of the model states are updated. Notation as."xx" indicates the gauges which are assimilated, see Fig. 2 for their locations. The corresponding ensemble of model forecasts and observations being assimilated are shown in Fig. 4.

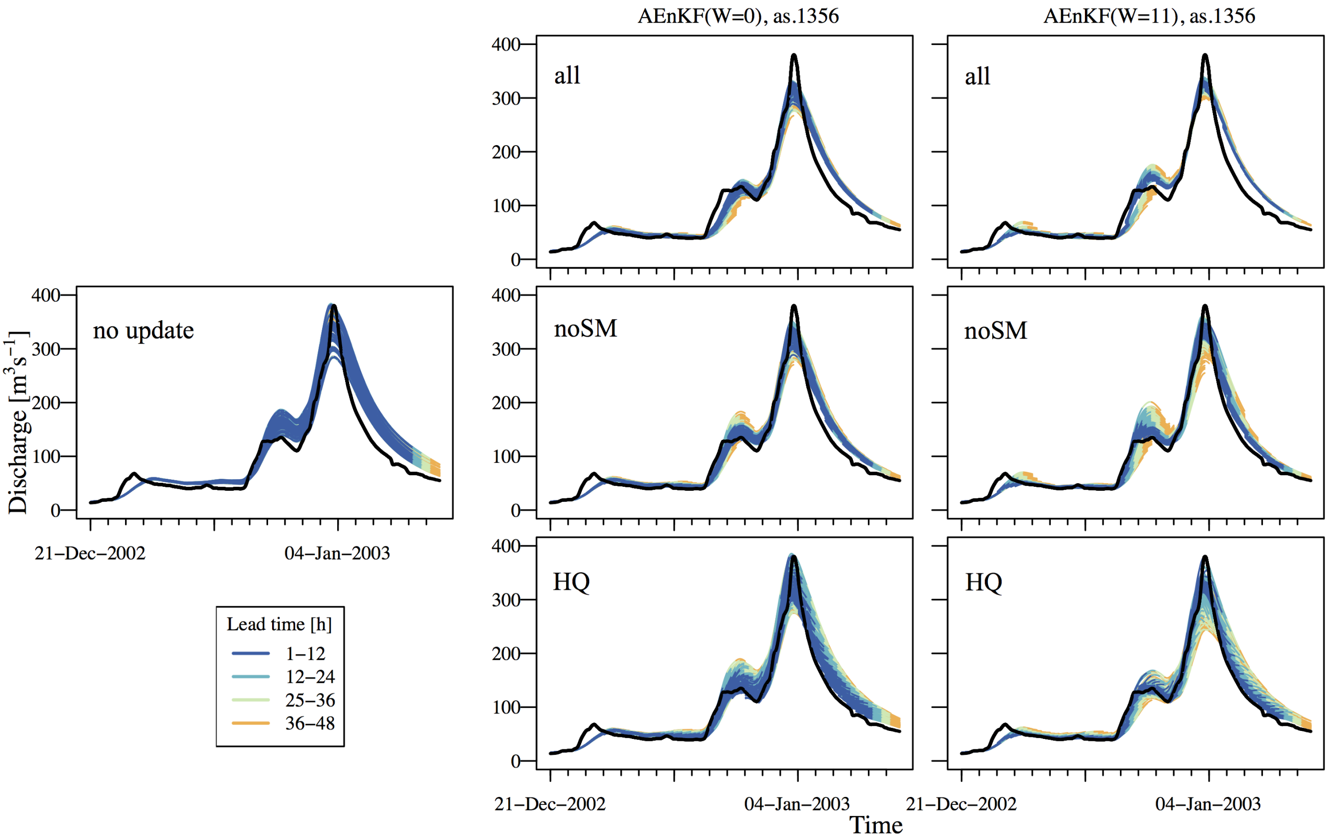


**Figure 6.** Ensemble of discharge forecasts for a typical event at the catchment outlet (Tabreux, gauge 1) for ~~different~~ three updating scenarios~~.~~ ~~The control run~~ : all, noSM, HQ (~~with no update~~see Table 2 for definition)~~is shown in the top panel~~. The combined effect of the model states being updated (3 scenarios shown in rows) and the length of the state augmentation vector ($W$) of past observations being assimilated (2 scenarios in columns) is presented. ~~The observations are shown in black.~~ Gauges 1, 3, 5, and 6 are assimilated. The control run (with no update) is shown in the left panel. The observations are shown in black.

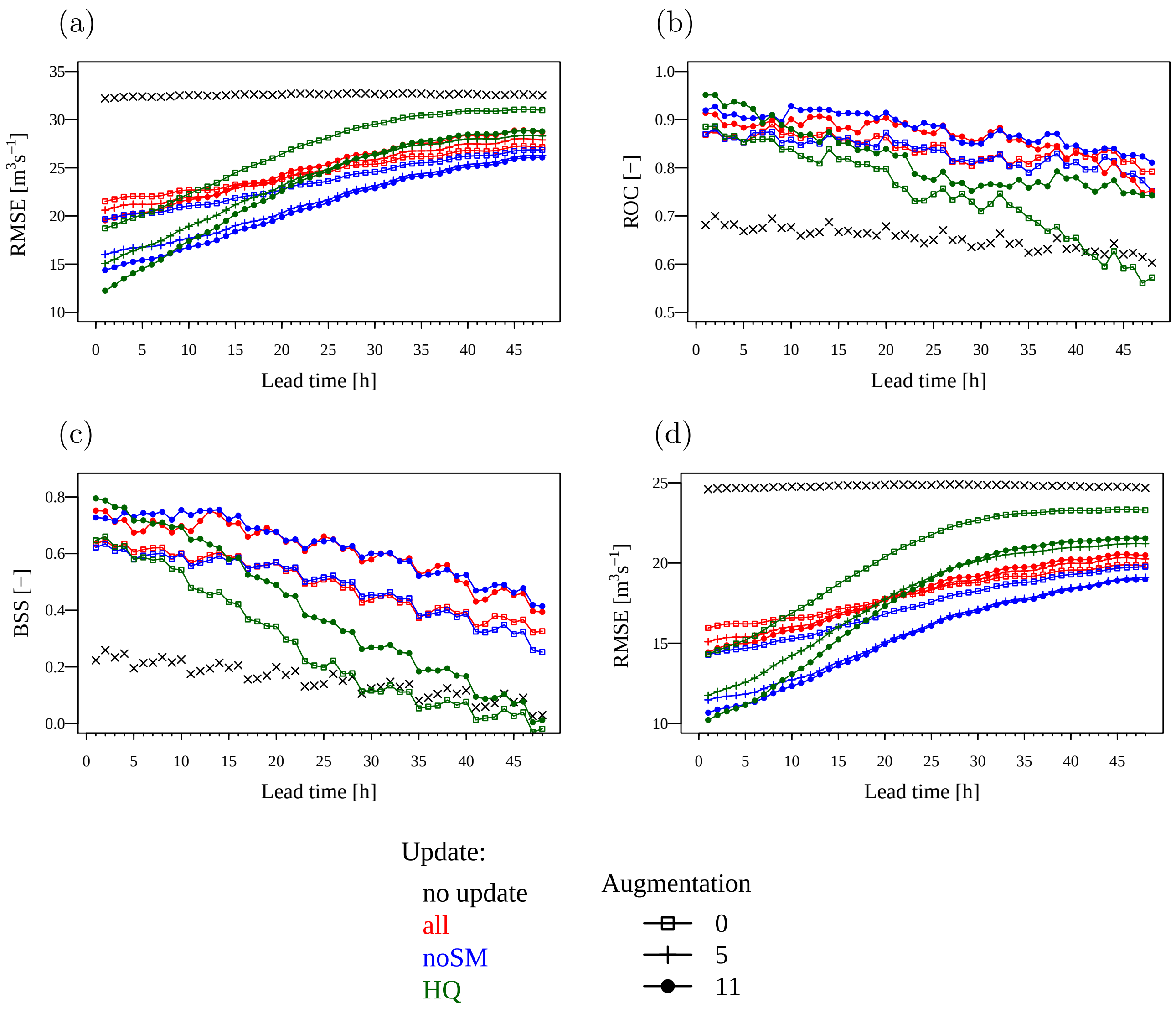


**Figure 7.** **(a)** Root-mean-square-error (RSME), **(b)** Relative operating characteristic (ROC), and **(c)** Brier skill score (BSS) at Tabreux (gauge 1) for different discharge observation vectors for which different model states are updated and with different lengths of the state augmentation vector ($W$) of past observations being assimilated. The results incorporate a set of 8 flood events shown in Table 1. Gauges 1, 3, 5, and 6 are assimilated. For BSS, the reference forecast is the sample climatology and only values larger than the 25th percentile of the whole sample are considered. **(d)** Same as **(a)** but the results are presented for Durbuy (gauge 2), a validation location which is not assimilated.

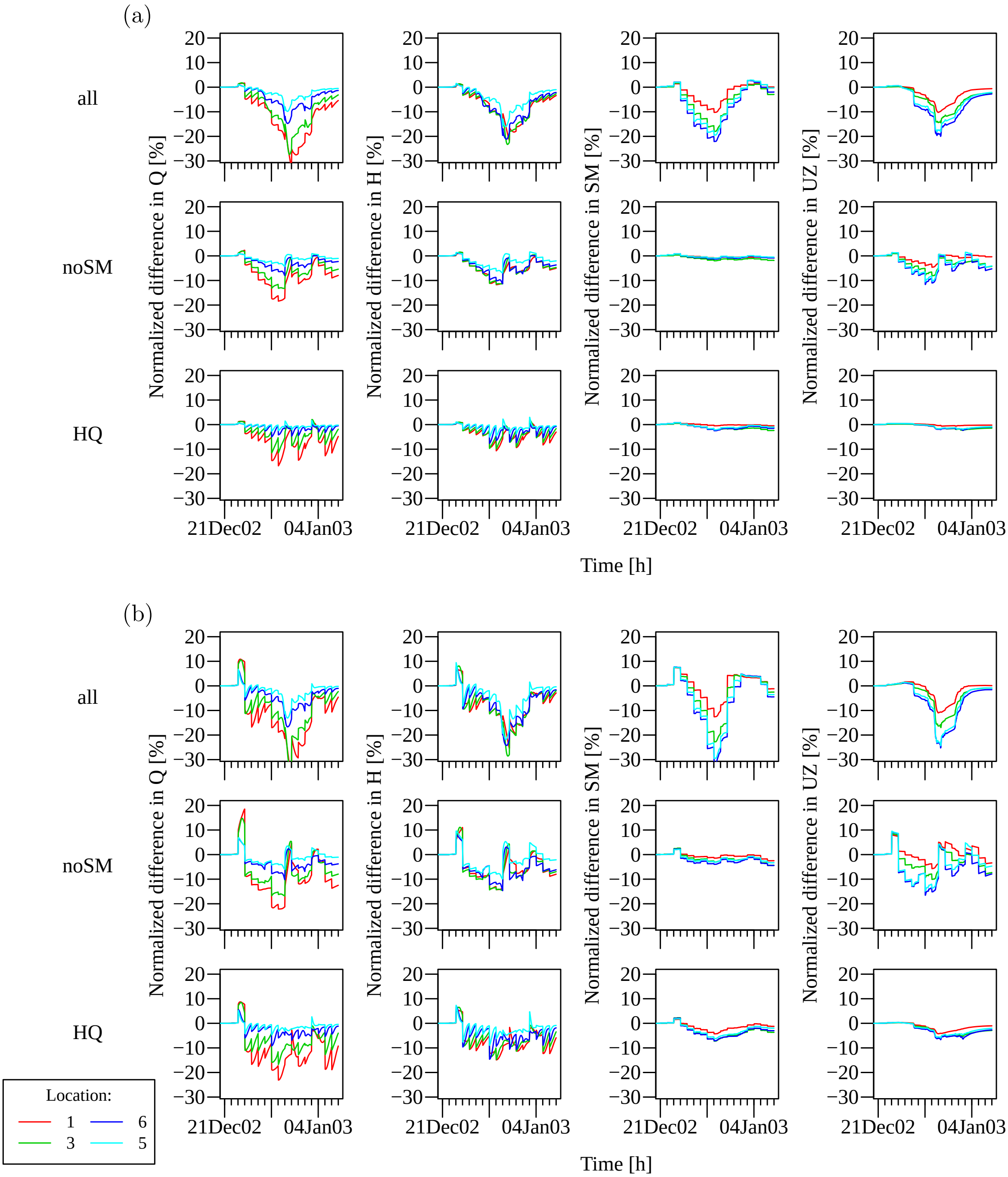


**Figure 8.** Scaled difference between the ensemble mean for the 3 partitioned update schemes and the control run without data assimilation at 4 gauging locations (shown by different colors) within the Upper Ourthe catchment using the AEnKF with **(a)** $W = 0$ and **(b)** $W = 11$. We dropped out the insensitive lower zone (LZ). Gauges 1, 3, 5, and 6 are assimilated. The results correspond to the same period as presented in Fig. 6.