EnKF paper, Anonymous Referee #1

In this paper, the authors update the states of a distributed hydrologic model through the assimilation of discharge data. Generally, the paper is well written, and interesting. What I would like to see is a more in-depth discussion on a number of results.

Answer: We thank Anonymous Referee #1 very much for her/his review and comments. Below we provide answers to the points, which were raised.

Major comments:

1) One *HUGE* issue that needs more discussion is the routing. Basically, the authors employ a kinematic wave model. This implies, as the authors correctly note, that one can avoid the need to update the states of the model in the past in order to assimilate the discharge in the present. However, physically (no hydrologist in the world is going to dispute this), runoff is determined by the soil moisture in the past. The consequence of this fact, combined with the routing model that avoids this, is that the modeled discharge is going to be much more sensitive to the state of the routing model than to the state of the hydrologic model. This also becomes evident in the results in the paper. As the modeled runoff is much more sensitive to the state variables of the routing model, the Kalman filter willl mainly update the routing model state variables (thus H and Q). Thus, which is confirmed in the paper, the results of the hydrologic model are hardly going to change at all through the assmilation of the discharge.
In the model description, the authors need to discuss the time-delay issue, the fact that they avoid it (and how), and the consequences this is going to have on the results. In the abstract, the results, and the discussion and conclusion, this really needs to be brought more forward (thus that the states of the routing model are updated, rather than the states of the hydrologic model, and why).

Answer: We thank the reviewer for this comment and we will bring this important point more forward in the revised manuscript as follows.

We will add more discussion (after line 27, p. 3968):
“In previous published papers (Weerts and El Serafy, 2006 and Pauwels and De Lannoy, 2009) a time delay issue was noted due to the use of the unit-hydrograph, where the discharge at time t depends on several previous calculated discharges (for instance at t-1, t-2, t-3, t-4, etc for hourly models to often 24h or more in case of the Ourthe). By using a physically based model the time delay and attenuation is modelled more realistically and the discharge and states x(t) depend only on the states x(t-1) (Markov model). The time delay is thus explicitly taken into account in the model.” This information will be added to the manuscript.

We will add/replace (l4-8,p3979) in discussion the following (same as our answer to Anonymus Referee #2, comment #19):
“In this study, mainly the pdf’s of the two routing model storages were affected by the Kalman filter update, while the other model states (SM, UZ, LZ) were found to be less sensitive to the EnKF scheme. This is because the current formulation of the EnKF (see Eq. 4) does not explicitly consider the strong correlation between soil moisture states in the immediate past and streamflow at the time of forecast. Therefore, it may be difficult to build a covariance matrix among the water balance model states (i.e., SM, UZ, LZ) via assimilating discharge observations.
Based on our results we can state that, given a measured discharge downstream, it is difficult to adjust (and justify) the soil moisture upstream (in a spatially distributed coherent manner) using an EnKF. Other filters like the Ensemble Kalman Smoother (EnKS), which calculate the analysis from several previous time steps (Evensen and Leeuwen, 2000), may result in better adjustment of the spatially distributed soil moisture states, which may improve forecasts for even longer lead times. In this study, however, with a larger number of assimilated discharge gauges, both the forecasted and updated pdf’s of SM, UZ and LZ had more narrow peaks around their actual true values.”

2) Another issue that also needs more discussion is the setup of the synthetic experiment. The authors retain one ensemble member as the synthetic truth. No parameters or initial conditions are changed to generate the ensemble, only the rainfall was changed. Essentially, noise is added to the rainfall. This implies that the authors assume that the only cause of uncertainty is the rainfall, and that for example model parameters and formulation are not a source of error. Most synthetic experiments that I know about, use different model parameters to generate the synthetic truth. This will introduce more differences between the model results and the observations, which is (arguably) closer to reality. I would recommend the authors to justify more the approach that they have used (thus not disturbing parameters to generate the synthetic truth).

Answer: The approach we took for the synthetic experiment is similar to the approach used by Weerts and El Serafy (2006), with the main differences being that we employ a realistic stochastic representation of the spatially distributed precipitation (see Rakovec et al., 2012, same SI). As stated clearly at several locations in the manuscript, we limit the analysis of the synthetic and real world experiment to input uncertainty only. The main reason for this is that we wanted to fully understand and investigate the filter process using a distributed hydrological model and realistic precipitation fields, we believe this is already challenging enough without initial state and model parameter / structural uncertainty. However, we agree with the reviewer this is necessary and our next step is to include/test/develop realistic stochastic representations of the model parameters and model structure as sources of uncertainty and their effects on the forecast performance.

Some minor comments:

3) P 3962 L 20: most hydrologic forecasting systems employ lumped models ? I would rephrase this as "most hydrologic discharge assimilation systems ...". There are actually quite a number of papers where distributed models are used for data assimilation (soil moisture).

Answer: Most operational hydrologic forecasting systems (used by agencies like Environment Agency in England and Wales and the National Weather Service in the USA, the Water Management Centre Netherlands in The Netherlands, etc) employ lumped models for hydrologic forecasting. We do agree that there are quite a number of scientific papers using distributed models to assimilate soil moisture, but we think that those are often not (yet) used in real time operational hydrologic forecasting systems. This Special Issue is focused on hydrologic forecasting, hence we added this statement to show the relevance of this work in relation to what is used in operational practice.
Therefore, we propose following phrasing: “most operational hydrological forecasting systems”.

4) P 3963 L 8-10: Sequential assimilation methods ... sequentially. Rephrase (sequential twice in the sentence).

Answer: We propose following rephrasing: “Sequential methods are mostly employed for state updating in hydrological models by assimilating observations when they become available.”

5) Same page L 17: Better to refer to Evensen (1994) when discussing the EnKF.

Answer: Yes, we will refer to the original paper Evensen (1994).

6) Same page line 25: I would argue that discharge measurements are the most widely used OPERATIONALLY FOR FLOOD FORECASTING. Again, there are quite a number of papers that use other data (for example soil moisture or temperature data for weather prediction).

Answer: We agree that there are some papers which use other data than discharge to assimilate. In the revised manuscript we will add a point that temperature observations play an important role in real-time operational forecasting systems, especially in regions with significant snow melt. However, we think that the in-situ soil moisture observations are used very rarely in **real-time operational** forecasting systems.

7) P 3967: Please provide a more detailed explanation on the Broersen and Weerts error correction method.

Answer: Since we did not compare the auto regressive moving average (ARMA) error correction method with the results of the presented EnKF study, we did not provide detailed description of the ARMA method and only refer to the paper of Broersen and Weerts (2005) and references cited therein. We think that a detailed description of a method which is actually not employed in our research would only distract the reader.
However, the methods used by Broersen and Weerts are available in a Matlab package called ARMASA (http://www.mathworks.com/matlabcentral/fileexchange/1330) developed by Broersen and coworkers from Delft University. The methods is also implemented in Delft-FEWS.

8) P 3968: Please provide some more explanation on how google maps is used for channel width estimation. I also do not understand why for non-channel cells, the channel width is equal to the cell width. If a cell is not in the channel, then perhaps the channel width should be zero?

Answer: A) We used Google orthophoto maps (satellite images) for rough estimates by eye to obtain information on the channel widths of corresponding Strahler stream order numbers.
B) For non-channel cells, the channel width is equal to the cell width, because the water is routed from these cells by a kind of sheet flow (see P 3979 L12) on the top of the whole grid. And by making the width of the non-channel cells very large, we are able to decrease the hydraulic gradient of the water in the "channel" of these non-channel cells and therefore increase the response time of these rather slowly responding cells. In case of zero channel width the hydraulic gradient could become infinite. Finally, we admit (P 3980 L 22) that this model structure without flux between neighbouring cells except for the two routing model states is a limitation and we want to investigate this effect in further research.

9) Page 3975, L 2: DA "machinery"... Please use another term.

Answer: We will change DA "machinery" to DA procedure (see comment #6 raised by Anonymous Referee #2)

Overall, if the authors provide a satisfactory response to these questions and remarks, the paper should be publishable.

References:

Broersen, P. M. T. and Weerts, A. H.: Automatic error correction of rainfall-runoff models in flood forecasting systems, in: Instrumentation and Measurement Technology Conference Proceedings, Ottawa, Canada, 17–19 May 2005, 2005

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99(C5), 10,143–10,162, doi:10.1029/94JC00572, 1994

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

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

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

Additional corrections:
A) We would like to thank to Seong Jin Noh from Kyoto University for his personal comment about the expression in the denominator of Eq. (12), where should indeed be N instead of (N-1). This will be corrected in the revised manuscript.

B) Additionally, for the real world experiment in Figs.10 and 11, there are erroneously shown asterisks for “true” model states in SM, UZ, LZ and H model states, for which the “true” model states are of course UNKNOWN. These asterisks for SM, UZ, LZ and H will be removed from Figs. 10 and 11 in the revised manuscript.