Authors' reply to second Short Comment by Albrecht Weerts
We thank Dr Weerts for his additional reflections on our discussion paper and our previous comments in reply to his first Short Comment. We provide responses to each individual point below. For clarity, comments are given in italics, our responses in plain text.

I like to thank the authors for the response to my comments/remarks

The example provided about the physically based modeling of the routing process was not to give criticism on the model used in the manuscript. It is clear that the model implements a similar description for routing as Rakovec et al. (2012). It was only an attempt to get clear where the strange/erroneous behavior when using the Kalman Filter in Clark et al (2008), Menodoza et al. (2012) and this manuscript is coming from. In the reply it is states that "The first is through a delay introduced to surface storage, which depends on the distribution of distances to the stream, and the overland flow velocity." How is the updating (when using EnKF) affecting this part of the model states? Could the strange behavior when applying the EnKF be a result of the way this process is being modeled (state t does not only depend on state t-1?)?

We believe the reviewer is incorrect in his assertion that our model does not have the 'Markov Property' that state t only depends on state t-1; our model routing is not equivalent to a unit hydrograph approach. This part of our model (the delay to surface storage to represent the unresolved stream network) is modelled as a frequency distribution (histogram) of water volumes according to their remaining residence times, i.e. time until they will enter the river network. Each bar of the histogram, which is resolved at 1-hour intervals, is stored as a separate model state, and can be updated by the Kalman filter. The size of the updates in the REnKF case is shown in Fig 10 of the paper. At each time step, the histogram is updated with inflows from the catchment (e.g. saturation excess flow), which enter the bins of the histogram according to a pre-specified distribution, and the water in the histogram is shifted by 1 hour to represent time passing. Outflow from this surface store is that portion of water with remaining residence time of 0. Hence the model states at time t do only depend on model states at time t - 1. This description of the surface routing is given fully in Clark et al (2008) which we cited, and hence is not repeated in the paper.
We also note that strong failures of KF-based approaches have been found by other research groups (Charles Perrin, *pers. comm.*).

Ideally, when applying EnKF (or any other algorithm) all model states are updated. I assume that focus of applying the data assimilation is providing accurate forecasts mainly for the short term (48 hours)? Rakovec et al (2012) showed improvement in accuracy over 48 hours of leadtime for a catchment of +/- 1600 km2 (which indicates that other stores than the routing stores are updated as well).

We note the following quote from Rakovec et al (2012): "Like in the synthetic experiment, hardly any change between the forecasted and the updated histograms is observed for soil moisture, upper zone storage and lower zone storage (Figs. 10 and 11), but visible changes can be seen in the routing storages, water level and discharge." Hence although the non-routing stores could be updated, this was not generally the case.

I am not sure what is meant by the remark/comment "We also note the HESSD comment to the Rakovec paper which questions their approach which does not include the physically realistic lagged relationship between hydrological model states and runoff at the catchment outlet."? In the reply and final paper this comment is adequately addressed. Moreover, in the paper a twin experiment was carried out to show the correct working of the DA setup followed by a real world experiment: : :....

Our comment was not meant to suggest that Rakovec et al (2012) did not address reviewers' comments or did not have a correct DA setup, only to highlight the discussion in the community regarding the best model stores to update during data assimilation, and the need to include the lagged relationship between model states and runoff.

I am not sure what is meant by "In all, we believe that our method provides a more physically realistic and sustained correction to model states" compared to which method/application are the author's referring (REnKF vs EnKF in the application described in the manuscript?)?

[Note that the quoted statement is from our previous reply-to-reviewer and is not from the paper]. Here we were comparing our method where soil moisture, groundwater and surface storage states were updated, to methods such as used by Rakovec et al (2012) and other papers we cited in the introduction (Randrianasolo et al., 2010; Berthet et al., 2009) which favour states more immediate to the gauging location such as in-channel water volume. As discussed, the latter methods provide one option to minimise the challenge of catchment lag (as the in-channel water has only a short lag before arriving at the gauging location). Our method takes an alternative approach whereby we update states further removed from the gauging location, but explicitly account for the lag.

"It is not easy to interpret the exact cause of the artefacts under the EnKF, although we believe we already provided physically realistic representations of time delay (kinematic routing) and time/space correlated perturbations, so these are less likely to be the cause. In Figure 13 we showed that the oscillations in the ensemble median flow under the EnKF were due to water being added/removed from the water table in that case (and this was replicated in other cases; not shown). Since the problem is corrected by use of the REnKF, we interpret that the artefacts are removed due to explicit representation of the lag time, and the iterated application of the EnKF. However, we agree that this is a likely explanation rather than a proof of the cause of the artefacts, and we will change the wording to reflect this." I am happy that the authors will rephrase the wording (no guessing!). However, I still feel this should be fully clarified before publishing the final results (and I think this could be clarified by carrying out a twin experiments).

We are glad that Dr Weerts is happy with the proposed rephrasing of this conclusion in our paper. We assume that a 'twin experiment' means to re-run the filter with synthetic data with known error characteristics. In our case, we preferred to evaluate the method using a year sample of actual precipitation and gauged streamflow observations, to test the performance of the filter in the conditions found in our operational system.

The comment with regard to variational algorithms (for instance Maximum Likelihood Ensemble Filter) was an indirect way to ask if you considered those variational algorithms as they might be less computational demanding and maybe some even more suitable for hydrological models than EnKF (as was recently discussed at the HEPEX DA workshop in Korea) .

Thank you for the clarification of this comment. We are quite aware that there are a variety of data assimilation methods available to hydrologists, including variational algorithms, particle filters and Kalman smoothers as suggested by Reviewer 2. One of the authors (C. Zammit) also attended the recent 2012 Beijing HEPEX workshop, where the techniques were also discussed. We will consider such methods as possible future directions, but in the case of this paper report on our existing choice of the EnKF with explicit consideration of catchment lag. We will comment more fully in the introduction regarding the alternative methods.