Dr. M. E. Gharamti
Mohn-Sverdrup Center for Operational Oceanography (NERSC)
C: (47) 947 222 09 E: mohamad.gharamti@nersc.no - Bergen - Norway

September 26th, 2016

Dear Dr. Hendricks Franssen,

Please find enclosed our revised manuscript "On the efficiency of the hybrid and the exact second-order sampling formulations of the EnKF: A reality-inspired 3D test case for estimating biodegradation rates of chlorinated hydrocarbons at the port of Rotterdam." The reviewer's comments were carefully considered in the revised manuscript. When we did not agree with his comments, we gave a thorough explanation. The reviewer's suggestion concerning the comparison between the serial and the original implementation of the adaptive hybrid scheme is very useful and we gratefully acknowledge it. Enclosed please also find a detailed response to the four points raised by the reviewer.

Essentially, we have compared the current serial implementation of the hybrid filter with the previously published one by Gharamti et al. (2014). We now further provide a discussion on the usefulness of the serial algorithm, compared to the original one, in section 4.1.2. The comparison is given in the new Figure 14.

We thank you for giving us a chance to resubmit a revised version of our manuscript and to reply to the reviewer's comments. We are also grateful for all your editorial efforts.

Please do not hesitate to contact me should you need anything else.

Sincerely yours,
M. E. Gharamti

Dr. M. E. Gharamti
Mohn-Sverdrup Center for Operational Oceanography (NERSC)
C: (47) 947 222 09 E: mohamad.gharamti@nersc.no - Bergen - Norway

Reply to Reviewer 3

We would like to thank the reviewer for his comments and suggestions. We followed the reviewer's recommendations and revised the manuscript accordingly. Our detailed replies to the reviewer's comments are given below.

1. I still have some concerns about the novelty of parts of the presented work. More than half of the results section deals with the comparison of EnKF and EnKF-OI and a sensitivity analysis on the weighting factors for EnKF-OI. This has already been presented in Gharamti et al. (2014) and I do not see why this discussion should be repeated in this manuscript (given that the conclusions and the overall model setup are very similar, see below). The authors should instead focus more on aspects that are built on top of the findings from Gharamti et al. (2014). For example, the authors mention that the sequential treatment of observations in the EnKF-OI scheme is a novelty in this work. However, if this is an innovative aspect of the paper why is this new method then not compared to the previous implementation? In the current state of the manuscript, the reader is left alone wondering if this change in the method really provides the supposed improvements.

The overall objective of the manuscript is to evaluate the contribution of the EnKF-OI and the $\mathrm{EnKF_{ESOS}}$ to the analysis step of the standard EnKF scheme. We propose to implement the hybrid scheme serially and allow the adaptive algorithm to select different weighting factors for each individual observation. We run similar assimilation experiments, as in Gharamti et al. (2014), to test the validity of the new serial hybrid scheme. We agree with the reviewer that comparing the current algorithm to the previously published one may be useful. We therefore followed the reviewer's suggestion and compared the performance of the new scheme (serial hybrid) to the previous one (batch hybrid). We included a new figure (Fig. 14) and discussed it in a new paragraph at the end of section **4.1.2**

We found out that the serial algorithm performs slightly better than the original scheme. In terms of relative improvements, the batch and the serial hybrid schemes suggest around 52% and 57% more accurate state and parameter estimates than those of the standard EnKF. We also note that processing the observations serially leads to a smoother selection of the weights between the ensemble-based and the static background covariances. In the batch scheme, the optimization is relatively more erratic and exhibits stronger variations over time.

From an algorithmic point of view, optimizing the parameters α and β in a serial sense is computationally more efficient and does not involve any matrix inversion, in contrast with the batch processing which requires the inversion of the matrix $\left(H_k P_k^f H_k^T + R_k\right)$ for every iteration of the optimization procedure

2. Additionally, a point that is stressed throughout the manuscript is that the authors intend to deal with a more realistic model setup. Real-world problem are generally characterized by uncertainty of subsurface parameters and boundary conditions, spatial heterogeneity, the presence of model structural errors and an increased level of complexity. However, most of the results presented in the paper are based on simplified model assumption, e.g., perfect knowledge on the flow field and hydraulic parameters and a lack of model

dynamics and heterogeneity. Due to these perfect model assumptions, the usage of a 3D model (instead of a 2D case) is also not likely to add much complexity to the overall setup. The new results including uncertainty in hydraulic parameters (presented now in section 4.3) are quite interesting and relevant for the practical application of these kinds of methods. So I would recommend the authors to extend these investigations and focus more on the practical issues of the method that arise under real-world conditions (as an ideal case was already discussed in their previous paper).

The objective of our manuscript differs from that of Gharamti et al. (2014). Apart from the hybrid's serial implementation, we investigated the impact/usefulness of accounting for sampling errors in both the forecast ensemble and the observations. The reviewer seems to miss this essential point and focuses his argument on the similarity to an earlier study. In the previous revision, we have followed the reviewer's recommendation and included a section where uncertainties are included in the groundwater flow component of the coupled system. Investigating other sources of uncertainty, such as those related to the boundary conditions, heterogeneity, structural errors, etc is interesting but does not fall within the scope of the current study. Building on the current findings about the data assimilation method, we will consider quantifying different sources of model uncertainties in a future study. This is now mentioned in the conclusion section. Thank you.

3. Currently, section 4.3 also lacks some quantitative information that underpin the comparison between EnKF and EnKF-OI. It is mentioned, that EnKF-OI outperforms EnKF for a moderate uncertainty of hydraulic parameters but no numbers are given to judge the different performance. Additionally, results for EnKF are missing in Figure 18.

We have now included the results from the EnKF using the perturbed flow scenarios. We report that under moderate flow-uncertainty, the EnKF-OI (α= 0.7) and the EnKF-ESOS respectively yield 22% and 20% more accurate concentration estimates. This gain is obviously less important than the one obtained under perfect flow conditions. Figure 18 (now Fig. 19) has been updated to include the EnKF results, as suggested. Thank you.

4. Additionally, there is also a problem with the initialization of the assimilation experiments. In a real-world case, one would typically derive the initial concentration fields by simulating a warm up period with historical data until the assimilation of measurements starts at a certain time t0 (e.g., after a monitoring network was installed). However, what is done in this paper is that the initial concentration fields are derived by taking snapshots from a forward simulation for exactly the same time period that is used later in the assimilation experiments. The problem here is that the initial ensemble at time t0 already 'predicts' a considerable contamination along the whole (future) trajectory of the plume. On the contrary, the observations at time t0 state that the aquifer is completely clean initially. For a real-world situation this would be quite unrealistic because this huge discrepancy between observed and simulated concentration (clean versus totally contaminated) would heavily questions the credibility of the forward model, i.e. such a forward model would eventually be discarded due to the unrealistic bias. I would recommend

the authors to account for this problem and to adapt their model initialization accordingly.

We disagree with the reviewer's argument. In real-world situations, we often have little knowledge about the initial configuration of the model state variables and the distribution of the parameters. A biased initial ensemble is another complexity that, if present, the EnKF may struggle with. Furthermore, the free-run simulation (from which we collect snapshots to construct the initial ensemble) is subject to various sources of model uncertainties, as compared to the reference run. As mentioned in the manuscript, we perturb the initial conditions and the degradation parameters. Thus, the initial ensemble does not reflect directly the truth or the sampled observations.
Selecting snapshots from a perturbed run is not new. In fact, many operational and near-operational studies follow a similar strategy:

- Sakov, P., Counillon, F., Bertino, L., Lisæter, K.A., Oke, P.R. and Korablev, A., 2012. TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. Ocean Science, 8(4), p. 633.
- Hoteit, I., Hoar, T., Gopalakrishnan, G., Collins, N., Anderson, J., Cornuelle, B., Köhl, A. and Heimbach, P., 2013. A MITgcm/DART ensemble analysis and prediction system with application to the Gulf of Mexico. Dynamics of Atmospheres and Oceans, 63, pp.1-23.
- Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L. and Zheng, F., 2014. Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian Earth System Model: a twin experiment. Tellus A, 66.

The reviewer's suggestion for initializing is now acknowledged and mentioned in the revised manuscript (Section **3.4**). Thank you