Reply to Reviewer #2

We thank the reviewer for commenting on the manuscript. Below please find our detailed response to the reviewer's concerns.

The paper presents an assimilation approach to subsurface contaminant transport problem inspired by the port of Rotterdam in the Netherlands. A multi-dimensional and multi-species reactive transport model is coupled to simulate the migration of contaminants within subsurface flow model. The biodegradation chain of chemicals is modeled for five decades. An artificial measurement data for the concentration is build using a synthetic setup and then used for updating the concentration and degradation rates in presence of model and observational errors. An adaptive hybrid ensemble Kalman filter is evaluated along side the exact second-order sampling formulation introduced by one of the authors in an earlier publication. The paper is well written and the presented numerical results are interesting. However, the test setup assumed perfect knowledge of the distributed subsurface parameters (permeability and porosity), which is generally unknown except at few locations.

We thank the reviewer for his positive feedback. Concerning the spatial distribution of the permeability (hydraulic conductivity) and porosity, we agree that quantifying their uncertainties is essential; however, it is beyond the scope of our work. In the revised manuscript, we now provide details on the offline estimation procedure that lead to a 3D parameterization of the flow parameters. Please refer to Section **3.2.2**.

I find the results convincing but would recommend that the authors add the following to the numerical study:

1- The utilized models and state parameter estimation techniques are limited to online updating systems which in many cases are known to under-perform iterative schemes (ensemble smoother where all the data is assimilated at once) specially within an annealing framework in what is known as ensemble smoother with multiple data assimilation. Can the author include that in their numerical study.

The objective of the current work is to test the usefulness of accounting for [1] EnKF forecast under-sampling issues (forecast step) and [2] EnKF observation sampling errors (analysis step). We then draw conclusions on how each of these two issues affects the accuracy and reliability of the resulting state and parameters' estimates. We realize that iterative ensemble schemes are convenient to apply in subsurface applications, but this is not the focus of our study. An ensemble smoother (ES) can still suffer from under-sampling issues during the forecast step because of the limited ensemble size. Adding the MDA scheme to the analysis step may help to improve the fit to the data when all are assimilated at once. It is further computationally more demanding and may suffer from convergence issues. In general, it can be subject to the same problems related to under-sampling of the background and observational error covariances as the EnKF. This study only considers the filtering problem. The smoothing framework could be considered in future studies. Thank you.

2- Could the authors re-run the model with the estimated parameters from the initial time step without data assimilation to assess the quality of the estimated parameters.

The reviewer is raising a good point here. Rerunning the simulation from the initial time using the estimated parameters is useful in real experiments in which the true parameters are not known. In our twin-experimental setup, the true parameters are known (given by Suarez and Rifai, 1999) and thus the quality of the estimated parameters is directly assessed by how far are those from the true ones. This has been analyzed in figures 8, 12, 15 and 16. Nevertheless, we followed the reviewer's suggestion and we did a forward model run using the estimated parameter. We compared, please see Fig. 1 below, the resulting MSE for concentration to that obtained using the initial parameters (initial ensemble mean from the EnKF runs). As can be seen, the estimates of the concentration improve when using the estimated biodegradation parameters in the FTR-Model. Overall, the gain in concentration accuracy is about 24%. We will be happy to add this figure in the revised manuscript if the reviewer still thinks that it can be useful. Thank you.

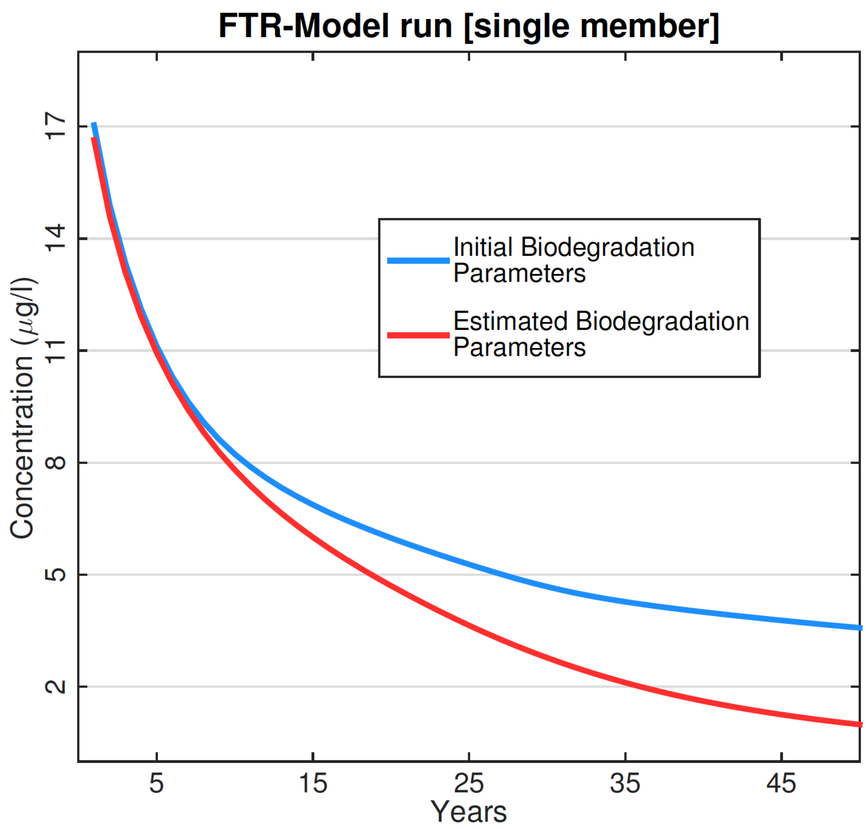

Figure 1: Free model run using the initial parameters and those estimated by the tuned hybrid EnKF-OI.

3- Uncertainties in spatial parameters (permeability and porosity) is a very interesting topic. Can the authors include some elements of that in their study even in a simplified way?

Following the reviewer's suggestion, we added a Section (**4.3**) in which we analyze a new set of results based on perturbed flow hydraulic parameters (conductivity and porosity). An ensemble of these hydraulic parameters is created and used to run the coupled FTR-model. We found that imposing large uncertainties on the hydraulic parameters strongly degrades the performance of all filtering schemes. Given that the performance of the hybrid EnKF-OI depends on the quality of the background statistics, satisfactory results were obtained only when the uncertainty imposed on the background information is relatively moderate. Further details can be found in the revised manuscript. Thank you.