# Reply to reviewer 3

We would like to thank the reviewer for the careful reading of the manuscript and appreciation of the work done in this paper. We did our best to take the remarks into account in a revised version of the manuscript (see explanation below).

# Major comments

- *A detailed presentation of the methodology is provided. One would expect that this methodology be compared with others, such as ES including the correction of the state (at least).*

  The following section has been added to the revised manuscript:

  Instead of optimizing the tidal boundary conditions one could directly try to estimate the best model trajectory (van Leeuwen, 2001; Hunt et al., 2004, 2007; Sakov et al., 2010). Whether one approach is preferred over the other depends on the application. In the present case, the amplitude and phase of the M2 tidal signal is a time invariant field. If tidal amplitude and phase at the boundary are corrected, the model can be rerun for any time period. This is not the case if the model trajectory is corrected. However, in the present approach one can easily try both methods without re-computing the ensemble members.

  The vector $\tilde{\mathbf{x}}^{(k)}$ represents the model trajectory (space and time) and the observation operator $\tilde{h}(\cdot)$ extracts the observed surface currents and elevation tidal parameters from the model trajectory. The rows of the matrix $\tilde{\mathbf{S}}$ are defined in a similar way as previously:

$$\left(\tilde{\mathbf{S}}\right)_k = (N-1)^{-\frac{1}{2}} \left(\tilde{\mathbf{x}}^{(k)} - \langle\tilde{\mathbf{x}}\rangle\right) \tag{1}$$

  The matrix containing the scaled ensemble anomalies of the observed part of the model trajectory $\mathbf{E}$ is the same as before. According to the analysis step of the Kalman filter, the optimal trajectory $\tilde{\mathbf{x}}^a$ is given by:

$$\tilde{\mathbf{x}}^a = \tilde{\mathbf{x}}^b + \tilde{\mathbf{S}}\left(\mathbf{E}^T\mathbf{R}^{-1}\mathbf{E} + \mathbf{1}\right)^{-1}\mathbf{E}^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \tilde{h}(\tilde{\mathbf{x}}^b)\right) \tag{2}$$

  The correction to the trajectory $\tilde{\mathbf{x}}^a$ is thus a linear combination of the rows of $\tilde{\mathbf{S}}$. The coefficients of this linear combination are the same as the ones obtained for optimizing the tidal boundary conditions (22) since:

$$\tilde{h}(\tilde{\mathbf{x}}^b) = h(\mathbf{x}^b) \tag{3}$$

  Therefore with almost no additional effort, one can compute the optimal trajectory $\tilde{\mathbf{x}}^a$. However, unlike the optimization of the boundary condition, one cannot compute the model results obtained for a different time period and compare them to the corresponding HF radar observations. We choose thus to compare the model results to the tide gage data to compare both approaches (as in section 5.3). By correcting the model trajectory, one obtains a RMS error 0.46410 m and 0.11976 m for Cuxhaven and Helgoland respectively. These results are very similar to the RMS error obtained by optimizing the tidal boundary condition (0.46431 m for Cuxhaven and 0.11990 m for Helgoland). One notices that the correction to the model state leads to slightly better results than correcting the boundary conditions but essentially both approaches provide virtually the same results. One could expect exactly the same results if the model dynamics would be linear. The fact that the RMS errors differ only on the order of tenths of millimeters shows *a posteriori* that the tidal propagation is mostly linear. The advantage however for correcting directly the tidal signal is that the new tidal parameters can be used for simulating different time periods.

- *The originality lies in that only the forcings are corrected. The authors justifies this approach in the introduction (correcting the state itself generates noise) but do not provide the least evidence of such behavior in their system.*

  The tidal waves travel through the model domain in only 5h. Therefore the entire model solution is immediately affected by the incoming tidal waves and the model solution is thus essentially forced by the lateral boundary conditions. Correcting the model state without correcting the boundary conditions will thus inevitably produce an inconsistent model state. In preliminary tests, an EnKF (correcting the model state) was implemented. However, this experiment turned out to be unsuccessful because of this issue.

- *The corrections are never illustrated. How large are the corrections to the forcings?*

  A figure with the corrections to the tidal boundary conditions is included in the revised manuscript (figure 10 in the revised manuscript). The elevation corrections at the boundary are of the order of 0.2 m.

  > Figure 1 shows the elevation correction from the assimilation scheme and the elevation M2 tidal amplitude and phase diagnosed from the assimilative run. Note that the scheme provides the correction over the whole domain but only the corrections at the boundary are actually used by the GETM model. The assimilation increases the amplitude of the incoming tidal wave (in the south of the western boundary) and decreases its phase.

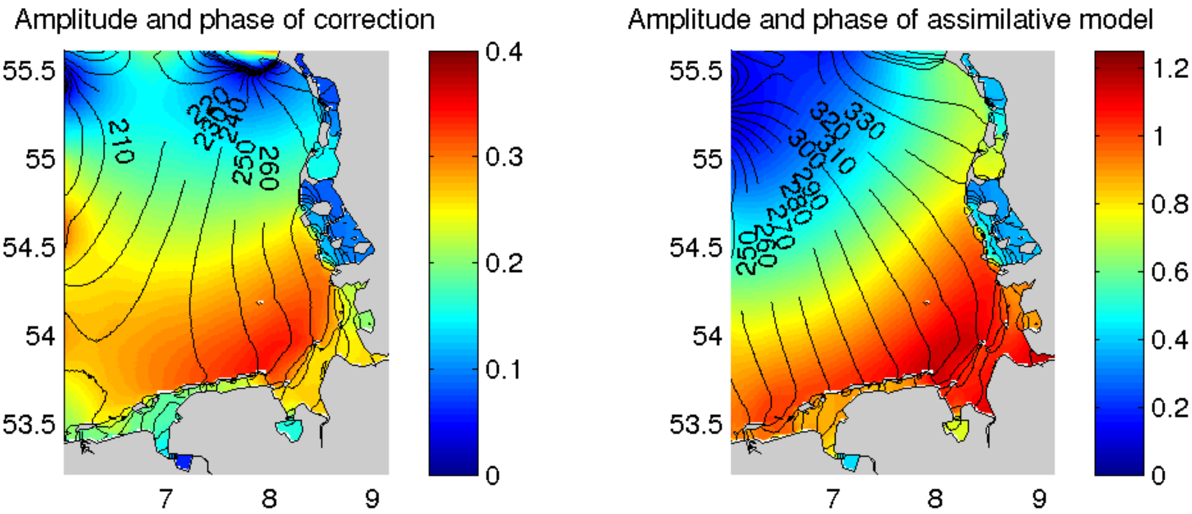


Figure 1: Amplitude and phase of the correction of the M2 tidal elevation (in m, left panel) and of the assimilative model run (right panel). The increment of the phase contour lines is 10°.

- *The authors propose an original application of the ES. They test it but do not find very convincing improvements over the free model when they compare the simulated fields with the observations (e.g. Figures 11 and 13). What must be concluded about the method?*

  The improvements in Figures 11 and 13 are small because they show the total RMS error between the model and HF radar surface current observations. This RMS error includes contributions from other error sources such as wind, bathymetry and other tidal constituents (as mentioned in the original manuscript, section 5.2). Here only the M2 tidal constituents are optimized. Therefore, the RMS error due only to the M2 tidal signal has been computed (figure 12). On this figure, the RMS error is roughly divided by 2. Also the comparison with independent tide gage data confirm an error reduction of the tidal surface elevation (by a factor of 2 of Helgoland and a factor of 1.4 of Cuxhaven).

- *The conclusion does not really conclude.*

The conclusions have been extended to include the comparison with the assimilation scheme optimizing directly the model trajectory (asked by this reviewer). It also includes several future improvements (suggested by reviewer 3): such as the optimization of other tidal constituents, bathymetry and bottom friction. Beyond the improvement of tides, it would be interesting to study if the scheme is also applicable for time dependent fields such as wind forcings and heat flux (by assimilating for example satellite SST).

## Minor comments

- *section 2.2 (EOT): Is it a climatology or a time dependent dataset?*

  The amplitude and phase of the M2 tides is a time independent data set. The manuscript is updated to clarify this.

- *section 4.1: The constraints are exposed clearly but not the sampling method. Is there some vector randomly sampled here? Which one, and how?*

  The 2nd order exact re-sampling strategy of the SEIK Filter is used (Pham, 2001). For the sake of briefness, this approach is only outlined in the present manuscript. For more details, the reader is referred to the article of Pham (2001).

- *What is actually perturbed initially to form the ensemble? The initial state and the forcings (lateral and atmospheric) all over the time window? What is corrected precisely by the assimilation? Is the initial state corrected here? This must clarified.*

  The purpose of this study is to correct only the M2 tidal signal. Therefore, initial conditions, other lateral boundary conditions and atmospheric forcings are not perturbed. Only the boundary condition of the M2 tides are optimized and not the initial state. This is clarified in the revised manuscript.

- *section 4.2: the data assimilation expert would feel more comfortable if the method is summarized as "the ES of Van Leeuwen and Evensen (1996) in which only the forcings are corrected" (or another sentence equivalent or more appropriate). This is somewhat lost in the text, in lines 16-22.*

  The proposed method certainly borrows concepts from the Ensemble Smoother of Van Leeuwen and Evensen (1996), but there are some differences as well. Beside the fact that we optimize the forcings instead of the model trajectory (as pointed out by the reviewer), the ES scheme used perturbed observations. In the present scheme a deterministic analysis step is used. Therefore, we think that calling the method "the ES of Van Leeuwen and Evensen (1996) in which only the forcings are corrected" is not entirely appropriate. See also the reply to reviewer 2 (1st item) for a comparison with 4D-EnKF and AEnKF.

- *I do not really see the need of the parallel with 4Dvar throughout the paper.*

  4D-Var is mentioned in the manuscript since the scheme has some aspects in common with it: 1. 4D-Var is most often implemented to estimate unknown forcing fields (or initial conditions). The proposed approach is also implemented to derive forcing fields (here tidal boundary conditions); 2. past and future data are used in the analysis; 3. The final solution does satisfy the model equations exactly (per construction). 4. The perturbations are derived using a cost function. This cost function corresponds to the background constrain in a 4D-Var scheme. The data constrain is now implemented using an ensemble scheme to avoid the formulation of an adjoint of the GETM model.

  Formally, the scheme can be seen as an optimal interpolation scheme with the time dimension embedded in the estimation vector and observation vector. One could re-derive the 4D-Var scheme from 3D-Var by using the same re-definition of the estimation vector and observation vector.

- *section 5: I expected to find some considerations about the length of the time window. 60 days seem a long period and I doubt the observations at the end of the interval have any impact on the corrections of the forcings (and initial state?) at the beginning of the interval. If the impact is significant, is it reliable? Are 51 members able to correctly represent correct error cross-correlations between two vectors separated by 60 days? Probably not.*

  The duration of the assimilation run is 40 days (not 60 days). The assimilation schemes aims to optimize the complex M2 tidal parameters (directly related to amplitude and phase of the M2 tides). Since the elevation at the boundary is a periodic function, the covariance between the surface currents inside the model domain and the complex tidal parameters is also periodic and never decays to zero with time.

  The amplitude and phase of the M2 tidal signal is actually a time invariant field (the underlying tidal signal varies of course in time, but not its amplitude and phase). A velocity measurement at any time is thus related to the tidal boundary conditions.

- *section 5.1 and figure 8: the conclusion is optimistic. It seems the assimilation of EOT data affects very poorly the current velocities but there is no interpretation of this in the text.*

  Actually, the assimilation of EOT data did improve the analysis. This was the purpose of figure 8. In this experiment, only EOT data is assimilated and HF radar surface currents are used for an independent validation. For all values (from 0.001 m to 1 m) of the error standard deviation for the EOT data, the RMS error compared the HF radar velocity (blue line) is smaller than the error of the free run (red line).

# References


Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. A., and Zimin, A. V.: Four-dimensional ensemble Kalman filtering, Tellus, 56A, 273–277, 2004.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D, 230, 112–126, 2007.

Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, Monthly Weather Review, 129, 1194–1207, 2001.

Sakov, P., Evensen, G., and Bertino, L.: Asynchronous data assimilation with the EnKF, Tellus, 62A, 24–29, 2010.

van Leeuwen, P. J.: An Ensemble Smoother with Error Estimates, Monthly Weather Review, 129, 709–728, 2001.