Observational error covariance specification in ensemble based Kalman filter algorithms

Tijana Janjić[1], Alberta Albertella[2],Sergey Skachko[3], Jens Schröter[1], Reinhard Rummel[2]

[1] Alfred Wegener Institute, Germany tpfander@awi.de, [2] Institute for Astronomical and Physical Geodesy, TU Munich, Germany, [3] Department of Earth Sciences, University of Quebec in Montreal, Canada

1 Introduction

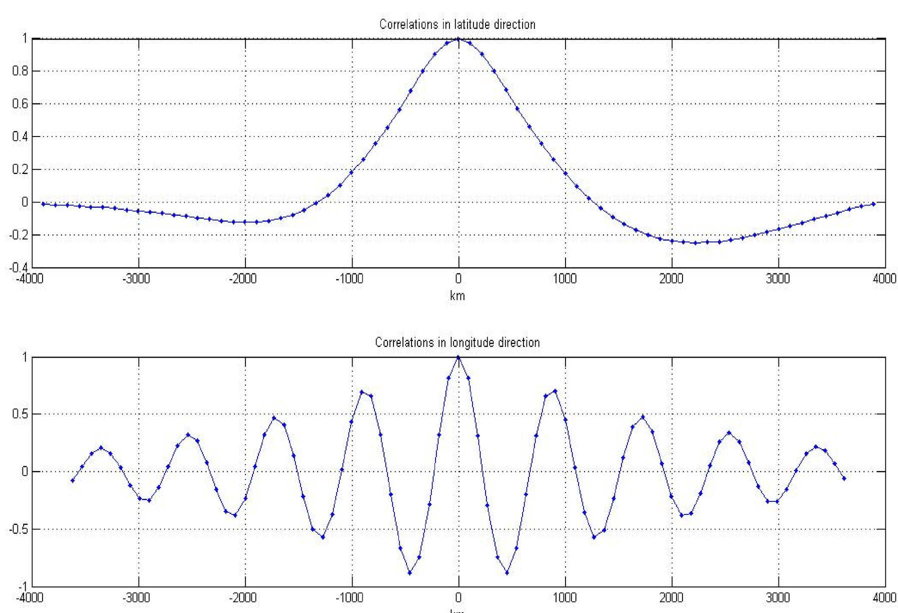
The study focuses on the effects of different observational error covariance structures on the assimilation in ensemble based Kalman filter with domain localization. With the domain localization methods, disjoint domains in the physical space are considered as domains on which the analysis is performed. Therefore, for each subdomain an analysis step is performed independently using observations not necessarily belonging only to that subdomain. Results of the analysis local steps are pasted together and then the global forecast step is performed.

The method of observational error covariance localization (Hunt et al. 2007) modifies the structure of the observational error covariance matrix for the subdomain depending on the distance of observation to the analysis point. We investigate use of different correlation structures together with this method in order to examine the relationship between correlational function used for weighting of observations and true observational error covariance of the data being assimilated.

Comparisons are done for estimation of ocean circulation via assimilation of satellite measurements of dynamical ocean topography (DOT) into the global finite-element ocean model (FEOM). The DOT data are derived from a complex analysis of multi-mission altimetry data combined with a referenced earth geoid. We are using domain localized SEIK algorithm with observational error covariance localization and different correlation models for localization.

2 Covariance function of observed dynamic ocean topography

The dynamical ocean topography is obtained by combining altimetric data with a high resolution reference geoid. In first approximation the altimetric data can be considered uncorrelated and the correlations of the DOT can be identified with the correlations of the geoid plus a constant variance of the altimetry part.



Correlation between the geoid at the point $\phi = -35^o$, $\lambda = -10^o$ and points of the South Atlantic area in latitude direction (upper panel) and in longitude direction (lower panel).

3 Experimental set-up

The data cover the period between January 2004 and January 2005. They are interpolated onto the model grid so that the observations are available at every point of the model grid every ten days. In the polar areas, part of Indonesian region and in Mediterranean sea, the observational data were substituted by the values of the RIO05 mean dynamical topography (MDT). These areas are characterized by low data accuracy due to presence of ice or of complex coastal/bottom topography.

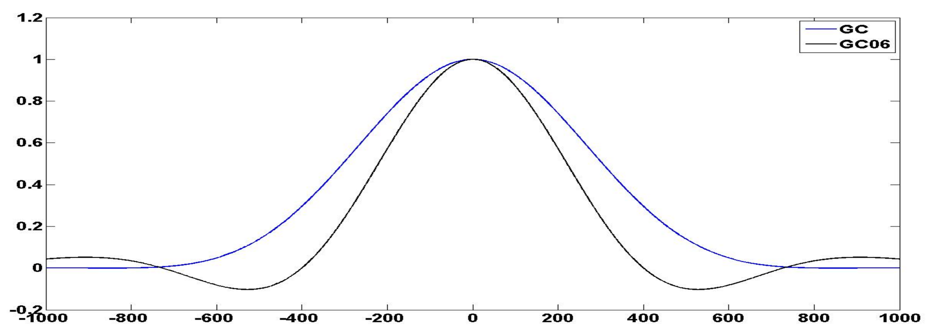
The study was performed using the Finite-Element Ocean circulation Model (FEOM) [3] configured on a global almost regular triangular mesh with the spatial resolution of 1.5°. There are 24 unevenly spaced levels in the vertical direction. FEOM solves the standard set of hydrostatic ocean dynamic primitive equations using continuous linear representations for the horizontal velocity, surface elevation, temperature and salinity.

The local SEIK filter algorithm as implemented within PDAF [4] is used in order to update the full ocean state, consisting of temperature, salinity, SSH and velocity fields at a given time at all grid points.

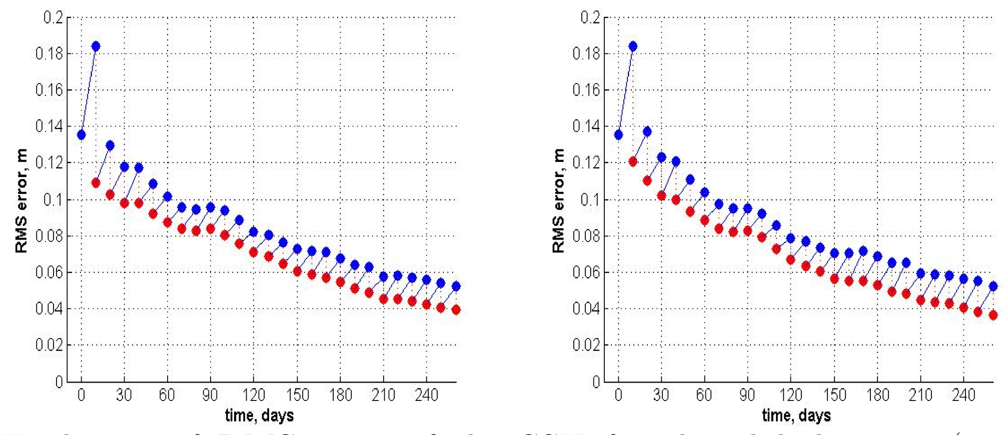
The analysis for each water column of the model depends only on observations within a specified influence region. Thirteen ensemble members are used in the implementation of the local SEIK algorithm. Observational error variance was set to 25 cm^2.

4 Effects on accuracy of weighting of observations

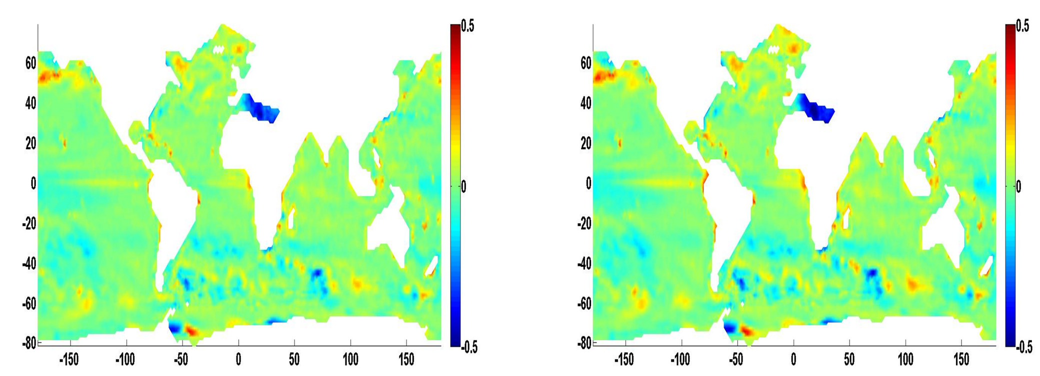
The two correlation models used are plotted in Fig. 4 and compared to uniform weighting. The use of correlation models improves the accuracy of forecast compared to uniform weighting (figure not shown here).



Correlation functions used for weighting of observations: 5th order polynomial weighting (blue line), and weighting using correlation model from [1] as shown with black line.



Evolution of RMS error of the SSH for the global ocean (except zones corresponding to RIO05 MDT location in the data). Experiment with observational radius of 900 km and 5th order polynomial weighting (left), and weighting as shown with black line in Fig. 4 (right).



Left: The mean difference between the dynamical topography obtained from the observations and from analysis for experiment with 5th order polynomial weighting and 900 km observational radius.

Right: The mean difference between the dynamical topography obtained from the observations and from forecasted fields. Results for experiment with weighting as shown with black line in Fig. 4 are similar (not shown).

5 Spectral results

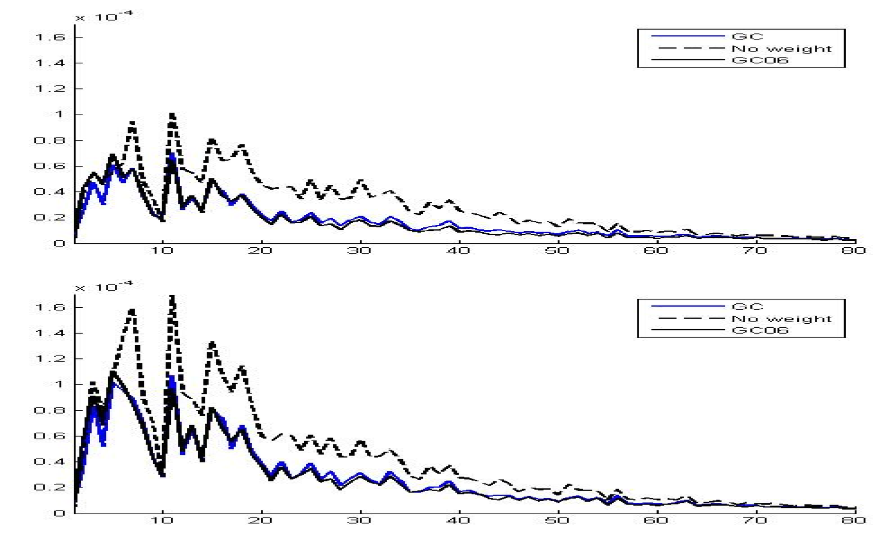
Following the same iterative procedure used in the computation of the geodetic DOT, it is possible to extend also the forecast field and the analysis field over the entire Earth's surface. Then spherical harmonic analysis can be applied to obtain the harmonic spectrum of each field or of their differences. The same procedure is applied to the forecast and to the analysis field.

In this first study only the error in the mean DOT (stationary part) obtained by averaging over 10 day outputs is considered. We consider the differences as function of the harmonic degree and order, computing the following index

$$\epsilon_\ell^{oi} = \sum_m \left(T^o_{\ell m} - T^i_{\ell m} \right)^2$$

where $T^i_{\ell m}$ are spectral coefficients and (i) can be $i = a$ for the analysis result or $i = f$ for the forecast result.
Spectral properties in forecast field show similar structure as for analysis, only the amplitudes have increased. This is true for all three methods.

The weighting by 5th order polynomial shows almost evenly distributed error structures for all scales. Note that since data itself have spectral coefficients only up to order 35, everything above 35 is the error that was introduced by analysis scheme and further amplified by the forecast. For the two weighting results shown this error is quite small.

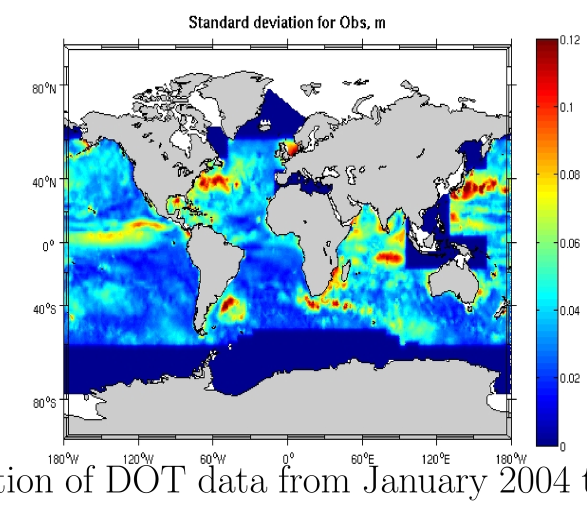


Upper: Log plot of spectral difference between analysis and the data depending on length scale.

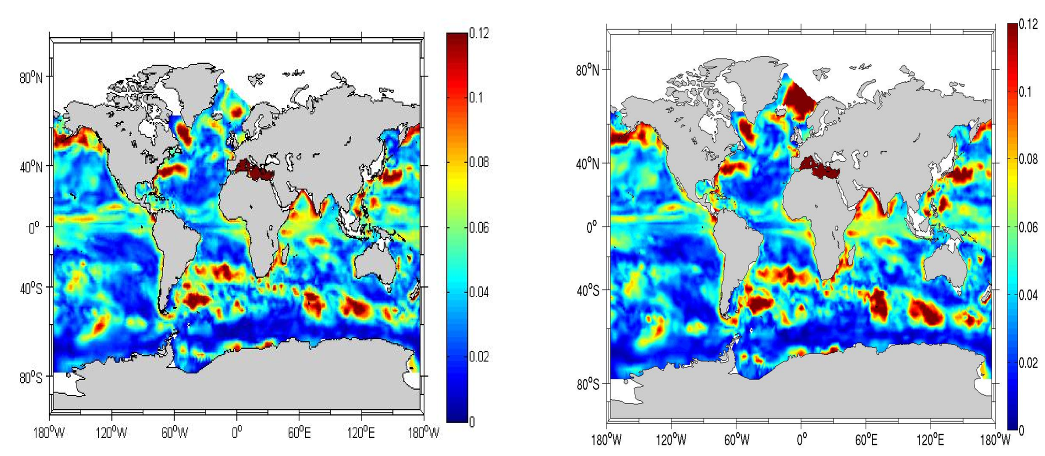
Lower: Log plot of spectral difference between analysis forecast and the data depending on length scale.

6 Effects on non-observed model variables

Finally we consider the effects on temperature field and steric height. We compared the standard deviation of the observations and standard deviation of the steric height calculated from the analysis. In many regions of the world ocean there is a good correspondence between these two fields. However also structures that are not present in the observations appear in the steric height standard deviations indicating scales introduced by assimilation.



Standard deviation of DOT data from January 2004 till January 2005. The deep-blue rectangular areas correspond to the locations where the RIO05 MDT was substituted in the data (no variability).



Standard deviation of steric height of analysis field from

Left: Experiment with 5th order polynomial weighting.

Right: Experiment with weighting as shown with black line in Fig. 4 .

7 Conclusion

- The use of observational weighting improves the accuracy of the analysis. In the case considered the true covariance matrix of the observations has a long length scales and the best possible approximation for use with domain decomposition method is searched after.
- The results of two correlation models used are similar, and more accurate then use of uniform weighting.
- Weighting of the observations by the 5th order polynomial produced spectral results that are closest to the data. The weighting by 5th order polynomial shows almost evenly distributed error structures for all scales.
- Structures that are not present in the observations appear in the steric height standard deviations indicating scales introduced by assimilation.

References

[1] G. Gaspari and S. E. Cohn and J. Guo and S. Pawson (2006): Construction and application of covarinace functions with variable length fields, *Quart. J. Roy. Meteor. Soc.*, **132**, 1815-1838.

[2] B. R. Hunt and E. J. Kostelich and I. Szunyogh (2007): Efficient Data Assimilation for Spatiotemporal Chaos: a local Ensmble Transform Kalman filter, *Physica D*, **230**, 112-126.

[3] Q. Wang and S. Danilov and J. Schröter (2008): Ocean circulation Model based on triangular prismatic elements with application in studying the effect of topography representation, *JGR*, **113**, C05015.

[4] L. Nerger and L., W. Hiller and J. Schröter (2005): PDAF - the Parallel Data Assimila- tion Framework: Experiences with Kalman filtering. In W. Zwieflhofer and G. Mozdzynski, editors, Use of High Performance Computing in Meteorology - Proceedings of the 11. ECMWF Workshop, pages 63–83. World Scientific.