

Observational error covariance specification in ensemble based Kalman filter algorithms

Tijana Janjić¹, Alberta Albertella², Sergey Skachko³, Jens Schröter¹, Reinhard Rummel²

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 (Brankart et al. 2003; Ott et al. 2004; Nerger et al. 2006; Hunt et al. 2007; Miyoshi and Yamane 2007), 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.

Because the assimilation is performed independently in each local region, the smoothness of the analysis fields are of concern with domain localization methods. In order to resolve the problem with smoothness of analysis fields, the method of observational error covariance localization was proposed by Hunt et al. 2007. This method 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.

The implementation of this localization scheme requires an additional loop that for each observation calculates a weight. The weight depends on the distance of the observation from the analysis location. Given the weight calculated using any of the correlation functions, the observation error covariance matrix is appropriately modified.

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. The assimilation results are compared in spectral space too. The effects on non-observed fields are considered, as well as the impact on consistency when comparing steric height changes induced by assimilation of dynamical ocean topography data.

2. Covariance function of observed dynamic ocean topography

The dynamical ocean topography is obtained by combining altimetric data with a high resolution reference geoid. Recent models of the gravity field are available together with the complete statistical information. If the complete error variance-covariance matrix is given, it is possible to compute not only the accuracy of the derived geoid but also the correlation between the geoid undulations in all points in which the geoid is computed. 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. Using the full variance-covariance matrix of the gravity model $ITG_{Grace03}$ computed at the University of Bonn (Mayer-Guerr 2006) up to the harmonic degree 60, the correlations between the geoid undulations in the South Atlantic area ($\phi = [-10^\circ, -80^\circ]$, $\lambda = [-50^\circ, +30^\circ]$) are computed. In Fig. 1 the correlations of the geoid at the point $\phi = -35^\circ$, $\lambda = -10^\circ$ in latitudinal direction (upper panel) and in longitudinal direction (lower panel) are shown. The correlations in latitude direction are zero at a distance greater than 1200 km from the considered point and are slightly not symmetric. The correlations in longitude direction have more oscillations and are significantly different from zero also at a distance of 3500 km.

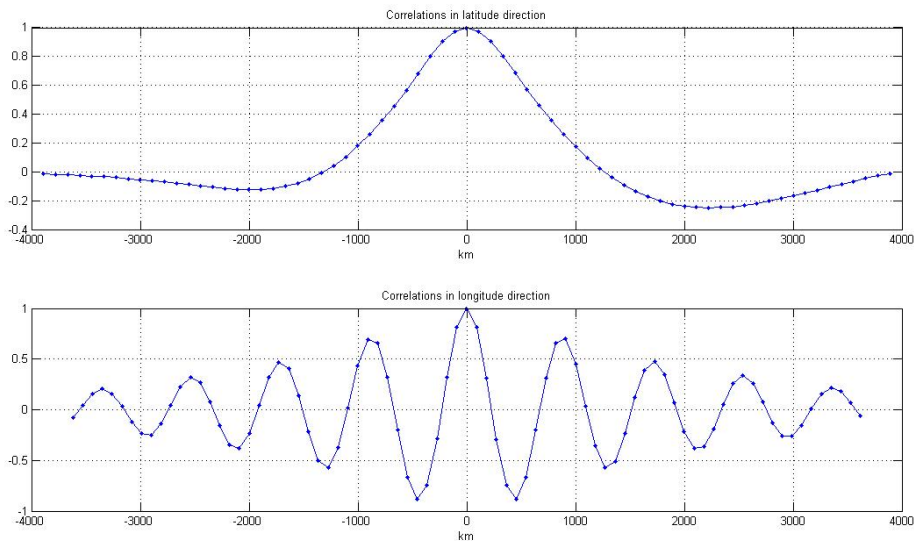


FIG. 1. Correlation between the geoid at the point $\phi = -35^\circ$, $\lambda = -10^\circ$ and points of the South Atlantic area in latitude direction (upper panel) and in longitude direction (lower panel) .

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) (Wang et al. 2008) 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.

In order to generate the initial error covariance matrix a model run was performed to produce a set of 10-day forecasts from a series of initial conditions distributed at 10-day intervals over a year. The initial error covariance matrix is then approximated with a lower rank matrix using the first thirteen empirical orthogonal functions of the ensemble. The initial field was the same as in Skachko et al. (2008).

The local SEIK filter algorithm as implemented within PDAF (Nerger et al. 2006, 2005) 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. 2. In Fig. 3 evolution of the root mean square error (RMS) is shown depending on the weighting structure of the observational error covariance. In this experiment, the influence region is a circle with a radius of 900 km. Results are shown for weighting using 5th order polynomial correlation function Gaspari and Cohn (1999) in the left panel of Fig. 3, uniform weighting of the observations is shown in the middle and correlation model from Gaspari et al. (2006) on the right. The use of correlation models improves the accuracy of forecast compared to uniform weighting.

Spatial structure of the mean error field can be seen in Fig. 4. In the left column differences between mean analysis over nine months of assimilation and mean of observations are plotted for weighting using 5th order polynomial correlation function (top), uniform weighting of the observations (middle) and weighting as shown in Fig. 2 (bottom). In the right column of Fig. 4 the same is shown for the forecast field. Errors in the forecast field are larger in

amplitude then for the mean analyzed field, however structures of the errors seen on the pictures are very similar for two correlation functions.

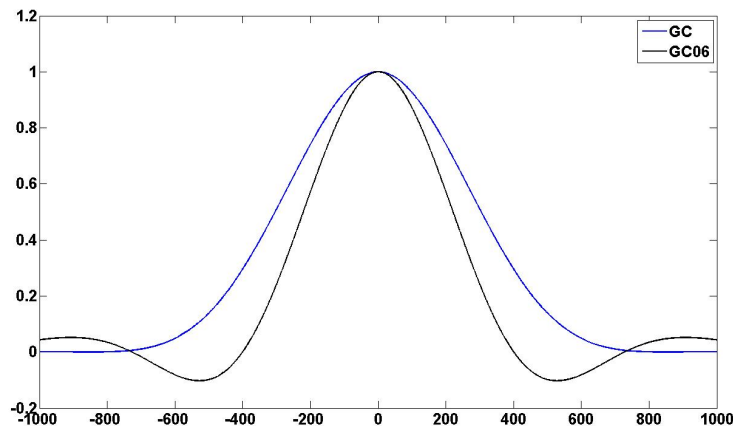


FIG. 2. Correlation functions used for weighting of observations.

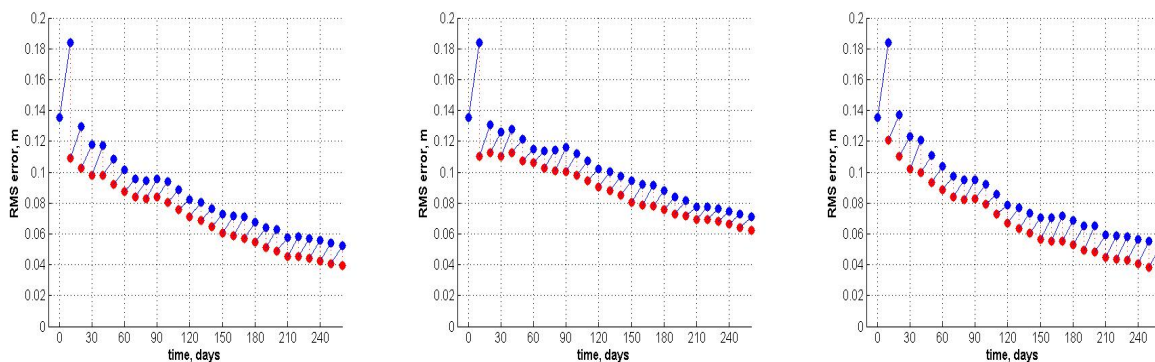


FIG. 3. Evolution of RMS error of the SSH for the global ocean (except zones corresponding to RIO05 MDT location in the data). Experiment with observational radius of 900 km and 5th order polynomial weighting, uniform weighting of observation and weighting as shown in Fig. 2.

5. Spectral results

It has been shown in the previous section that the modification of the observational error covariance by different correlation weighting has an effect on the accuracy of the results obtained. Question arises whether this improvement in the accuracy is the result of additional filtering introduced by modification of the observational error covariance. In this section we

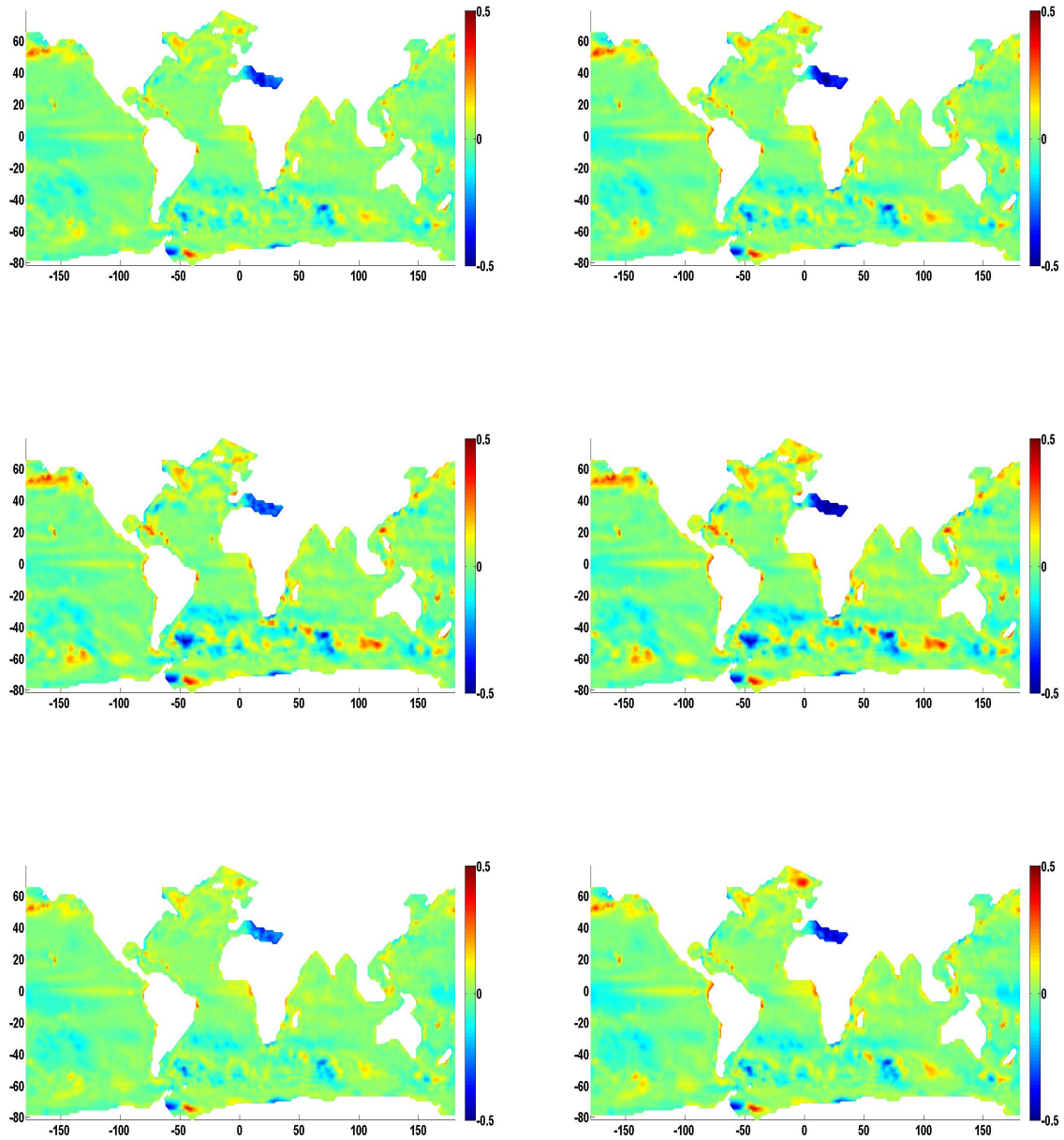


FIG. 4. Difference between the mean field of SSH from observations and mean SSH from analysis (left). Difference between the mean field of SSH from observations and mean forecasted SSH (right). Experiment with observational radius of 900 km and 5th order polynomial weighting (upper), uniform weighting of observation (middle) and weighting as shown in Fig. 2.

discuss spectral properties of errors depending on which observational covariance is used for weighting.

The iterative method of Albertella et al. (2008) makes it possible to calculate the spectral content of the ocean surface in spherical harmonic expansion. 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. In Fig.5 spectral error distributions are shown in terms of length scale. That is we consider the differences as function of the harmonic degree and order, computing the following index

$$\epsilon_{\ell}^{oi} = \sum_m (T_{\ell m}^o - T_{\ell m}^i)^2$$

where $T_{\ell m}^i$ are spectral coefficients and (i) can be $i = a$ for the analysis result or $i = f$ for the forecast result.

In Fig. 5 the index ϵ_{ℓ}^{oi} for the three weight functions is shown. 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.

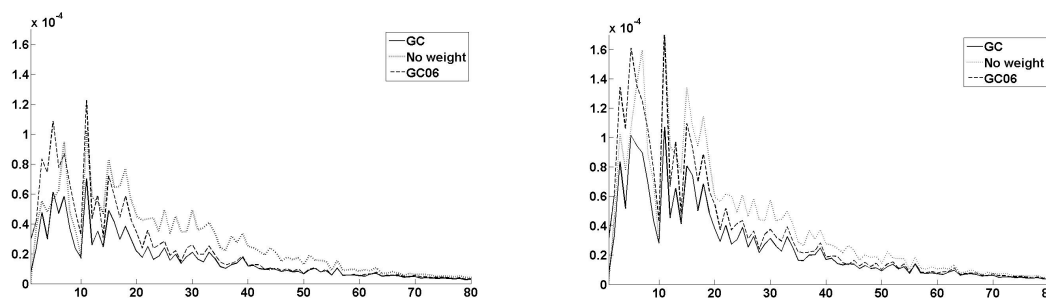


FIG. 5. Log plot of spectral difference between analysis and the data (left) and forecast and the data (right) depending on length scale.

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 (figure not shown). 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.

7. Conclusion

The method of weighting of observations that is often used together with domain localization was examined in the case of assimilation of dynamic ocean topography data into the global finite element ocean model. In this example, 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 but show differences in spectral results. Depending on which correlation function is used for weighting, the spectral properties of the solution can be improved. 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.

References

- Albertella, A., R. Savcenko, W. Bosch, and R. Rummel, 2008: Dynamic ocean topography - the geodetic approach. *IAPG/FESG-Schriftenreihe*, (27).
- Brankart, J.-M., C.-E. Testut, P. Brasseur, and J. Verron, 2003: Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean models: Application to a 1993-1996 hindcast of the North Atlantic ocean circulation. *JGR*, **108**, 19–1–19–20.
- Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Gaspari, G., S. E. Cohn, J. Guo, and S. Pawson, 2006: Construction and application of covariance functions with variable length fields. *Quart. J. Roy. Meteor. Soc.*
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.
- Mayer-Guerr, T., 2006: Gravitationsfeldbestimmung aus der analyse kurzer bahnboegen am beispiel der satellitenmissionen CHAMP und GRACE, Dissertation University of Bonn.

- Miyoshi, T. and S. Yamane, 2007: Local ensemble transform Kalman filter with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.
- Nerger, L., S. Danilov, W. Hiller, and J. Schröter, 2006: Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. *Ocean Dyn.*, **56**, 634–649.
- Nerger, L., W. Hiller, and J. Schroeter, 2005: Pdaf the parallel data assimilation framework: Experiences with kalman filtering. *Use of High Performance Computing in Meteorology, Proceedings of the 11. ECMWF Workshop*, 63–83.
- Ott, E., et al., 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.
- Skachko, S., S. Danilov, T. Janjić, J. Schroeter, D. Sidorenko, R. Savchenko, and W. Bosch, 2008: Sequential assimilation of multi-mission dynamical topography into a global finite-element ocean model. *Ocean Sci.*, **4**, 307–318.
- Wang, Q., S. Danilov, and J. Schroeter, 2008: Finite element ocean circulation model based on triangular prismatic elements with application in studying the effect of topography representation. *J. Geophys. Res.*, **113**, 5015.