Hybrid SEIK algorithm for oceanographic data assimilation

Tijana Janjić, Sergey Skachko, Sergey Danilov and J. Schröter
Alfred Wegener Institute, Bussestrasse 24, D-27570 Bremerhaven
Germany.

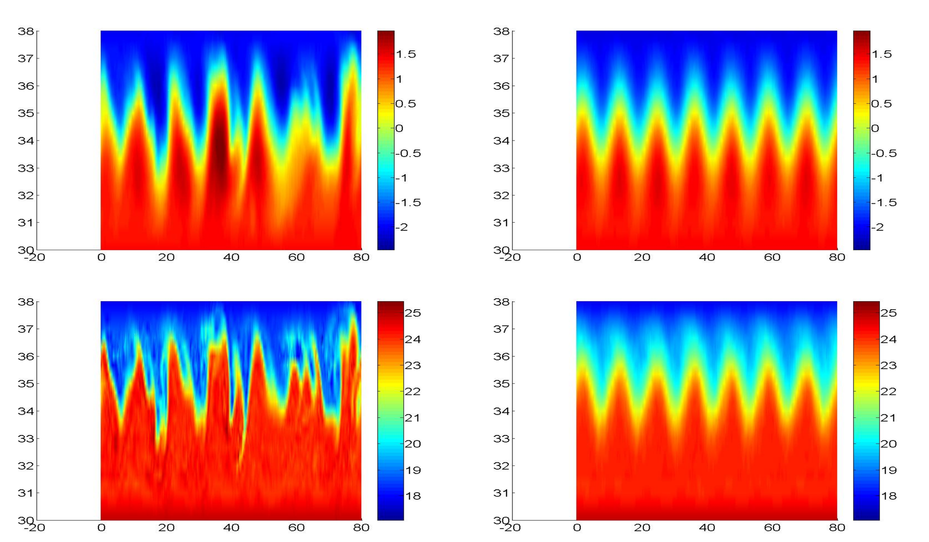
1 Introduction

The singular evolutive interpolated Kalman filter (SEIK) as a variant of the ensemble Kalman filter has been implemented and tested for application in oceanography assimilating altimetric data within twin experiments framework. Previous studies suggest that this filter is reasonably well-behaved in the presence of instability. In the SEIK assimilation algorithm, the analysis error covariance matrix is approximated by a covariance matrix whose rank corresponds to the number of ensemble members used for representing the forecast error covariance. In order to achieve a computationally efficient algorithm, the rank of this covariance matrix is often chosen to be small, leading to problems with the convergence of the filter.

We modified the SEIK algorithm in order to incorporate stationary covariance. Our goal was the improvement of the performance of the SEIK algorithm without increasing the number of ensemble members, since the evolution of the ensemble members is the most computationally demanding part of the algorithm. The performance of our hybrid algorithm has been tested and compared to the SEIK algorithm as well as to the local SEIK algorithm using twin experiment set up and a simplified channel configuration of the Finite Element Ocean Model (FEOM) developed at the Alfred Wegener Institute (AWI). In the experiments, we deal with flows generated by baroclinic instability. In this setting, we compare results of the assimilation when only observations of the sea surface height (SSH) data are present. In particular, we consider the distribution of the errors with height in all model fields and this way investigate the filter's ability to propagate information from the measurements of SSH to larger depths.

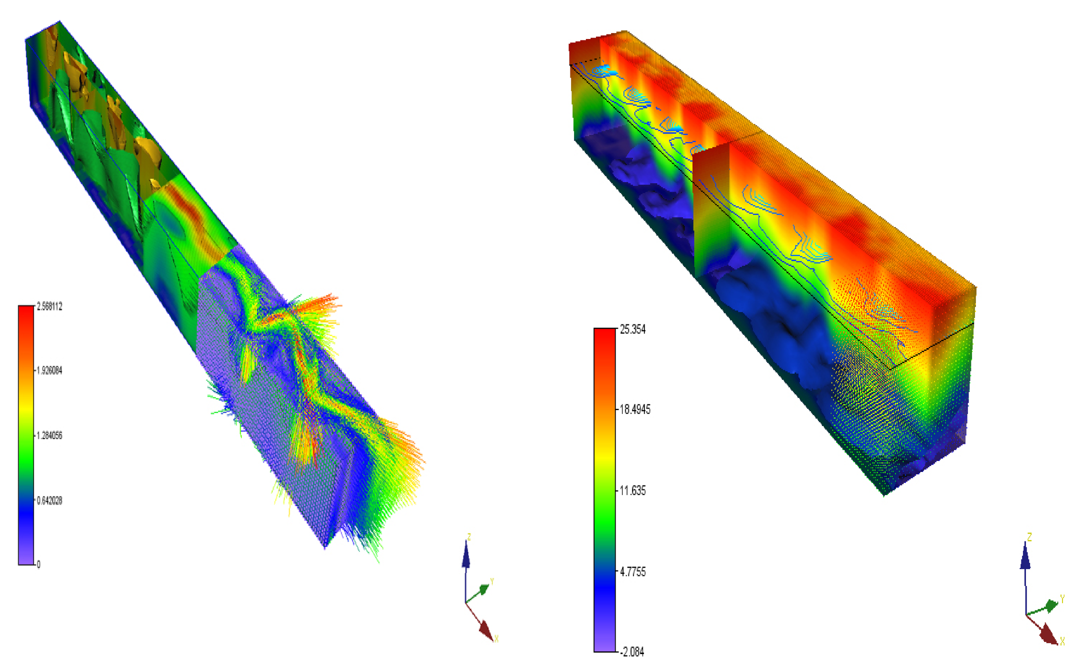
2 Twin experiment setup

- A reference model run is used as "true" field for verification and for the construction of observations.
- Observations at time t_k are obtained by adding random Gaussian measurement error with mean zero and standard deviation of 5 cm to the "true" SSH field at time t_k.
- Observations of the SSH are assimilated every 10 days.
- In order to initialize the filter we make a long model run, and using different model states, we construct the initial covariance and the initial guess. The initial analysis is calculated as the average over model realizations, and the initial covariance matrix is then estimated via EOF analysis on that sample as $\mathbf{P}_0^a = \mathbf{V}_0\mathbf{U}_0\mathbf{V}_0^T$. Here $\mathbf{V}_0 = [\mathbf{v}_1 \dots \mathbf{v}_r]$ is $n \times r$ matrix whose columns are the r eigenvectors $\mathbf{v}_i, i = 1, \dots, r$ corresponding to the r largest eigenvalues of the computed covariance matrix on the sample. The matrix $\mathbf{U}_0$ is $r \times r$ diagonal matrix with these r eigenvalues on the diagonal.
- all experiments use $r = 15$ EOFs.
- In the experiments, we deal with flows generated by baroclinic instability. We use a simplified channel configuration of the Finite Element Ocean Model (FEOM) developed at the Alfred Wegener Institute (AWI).



Left: True SSH (upper panel) and SST fields (lower panel) at $t = 0$.

Right: Estimated SSH (upper panel) and SST fields (lower panel) at $t = 0$.



Left: True velocity field at $t = 200$ days.

Right: True temperature field at $t = 200$ days.

3 Hybrid SEIK Filter Algorithm

Hybrid ensemble variational schemes were developed to incorporate error covariances estimated from an ensemble into the variational framework [2]. In theses schemes, the background error covariance is replaced by a weighted sum of the 3DVAR background error covariance and the sample ensemble covariance.
We developed a hybrid scheme starting from the SEIK algorithm. In the SEIK algorithm (see [4]), $\mathbf{x}_k^f$ is calculated as the average over the ensemble members $\mathbf{x}^{f,i}(t_k), i = 1, \dots, r+1$ and $\mathbf{P}_k^{f,SEIK}$ as corresponding covariance matrix. From such calculated $\mathbf{x}_k^f$ and $\mathbf{P}_k^{f,SEIK}$ the analysis $\mathbf{x}_k^a$ is obtained using formulas:

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{x}_k^o - \mathbf{H}_k\mathbf{x}_k^f),$$

$$\mathbf{K}_k = \mathbf{P}_k^a\mathbf{H}_k^T\mathbf{R}_k^{-1}. \tag{3.1}$$

The analysis covariance matrix $\mathbf{P}_k^a$ is easily calculated in its reduced form from the reduced form of the forecast error covariance matrix. Such calculated $\mathbf{P}_k^a$ is used in (3.1) to calculate the $\mathbf{x}_k^a$ as well as for generating new ensemble members. The new ensemble members are redrawn in such a way as to have their mean equal to $\mathbf{x}_k^a$ and the covariance equal to $\mathbf{P}_k^a$. Each of the ensemble members is evolved to the new time step to obtain new $\mathbf{x}^{f,i}(t_k), i = 1, \dots, r+1$.
In the hybrid algorithm $\mathbf{P}_k^f$ is modeled as

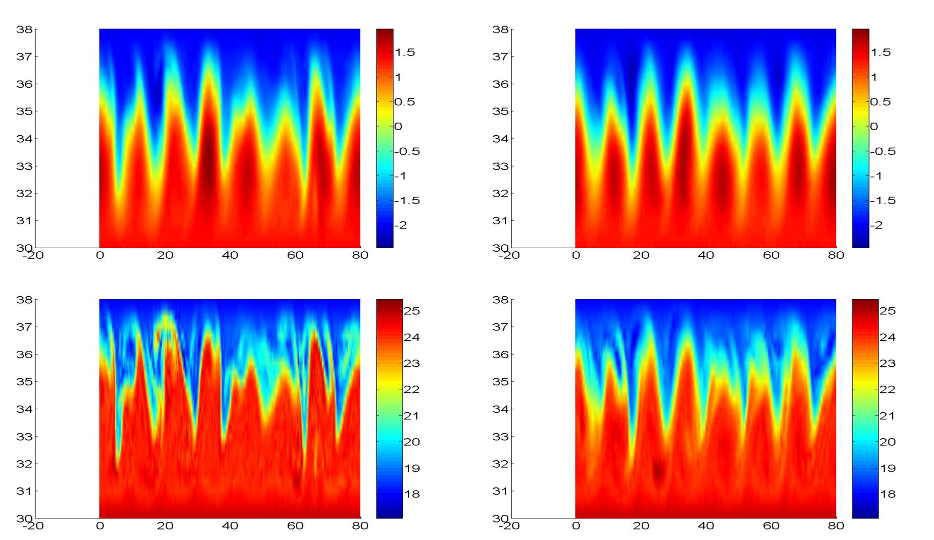
$$\mathbf{P}_k^f = \alpha\mathbf{P}_k^{f,SEIK} + (1-\alpha)\mathbf{B}, \tag{3.2}$$

where $\mathbf{B} = \mathbf{P}_0$ is the covariance matrix that is constant in time and α is a coefficient between zero and one.
The analysis error covariance matrix $\mathbf{P}_k^{a,SEIK}$ is calculated again in reduced form from (3.2) and resampling is done from the covariance given as linear combination

$$\alpha\mathbf{P}_k^{a,SEIK} + (1-\alpha)\mathbf{B}. \tag{3.3}$$

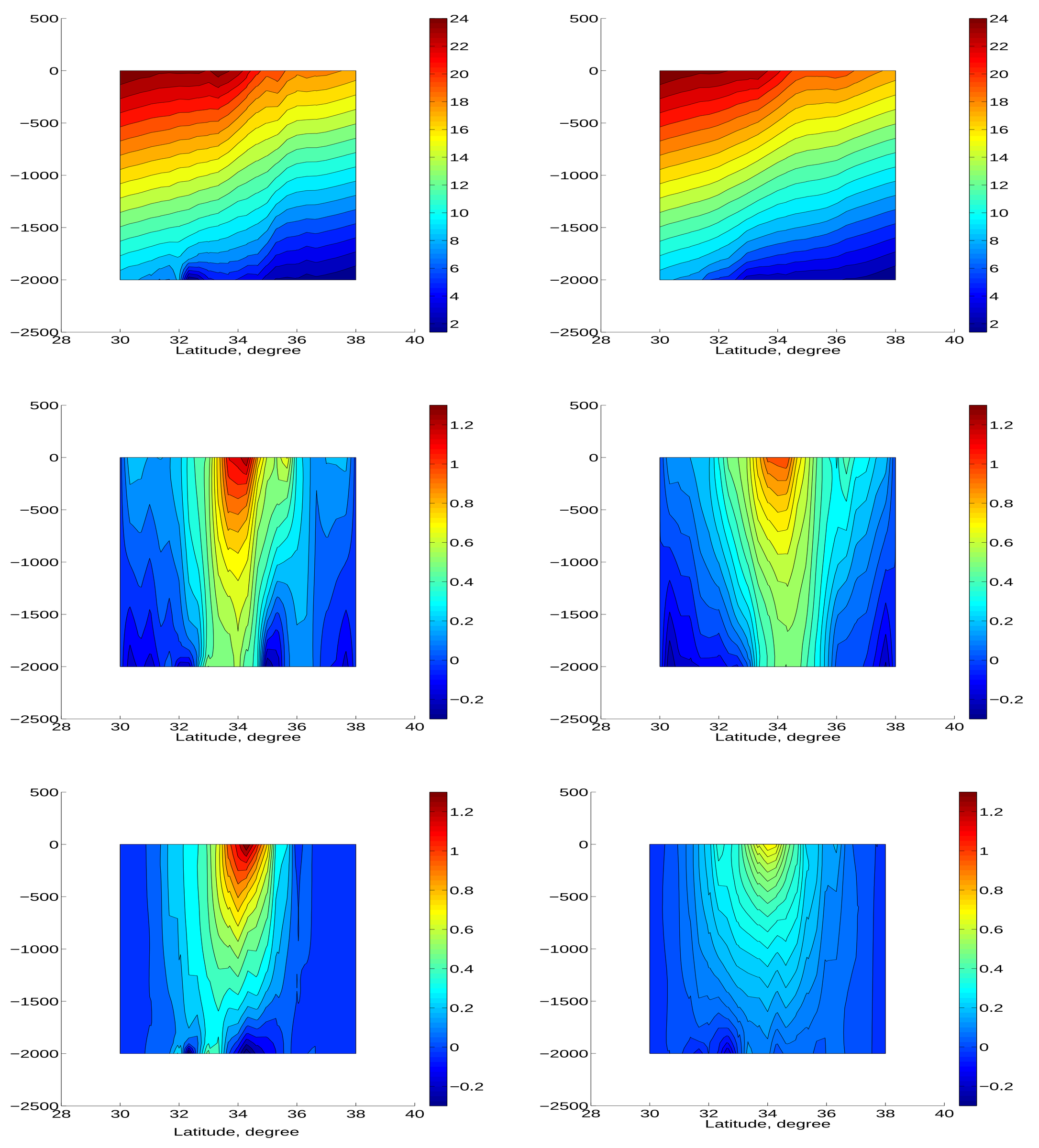
This procedure provides ensemble analysis states for the next step of the SEIK filter algorithm.
The performance of this hybrid algorithm has been tested and compared to the SEIK algorithm, use of constant covariance matrix as well as local SEIK algorithm (see [3]) using twin experiment set up from above.

4 Twin experiment results



Left: True SSH (upper panel) and SST fields (lower panel) at $t = 200$ days.

Right: Estimated SSH (upper panel) and SST fields (lower panel) at $t = 200$ days (after 20 analysis steps) using hybrid algorithm with $\alpha = 0.8$.



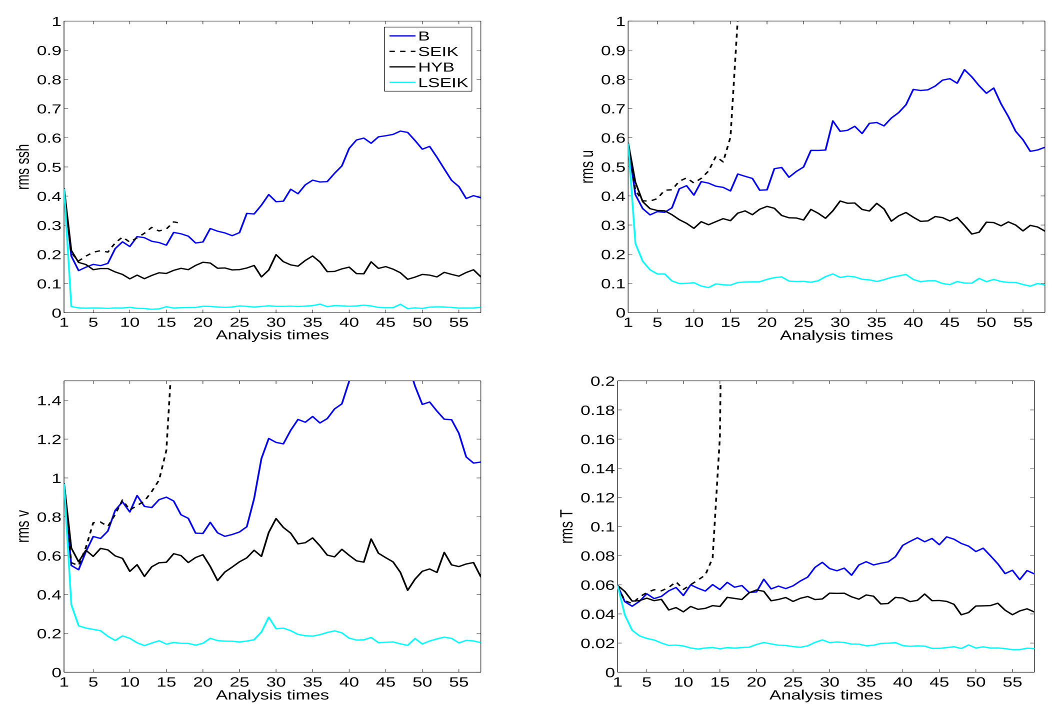
Left: Cross section of true state at $t = 200$ days and approximately 40^o lon. From top to the bottom: Temperature, zonal velocity, meridional velocity.

Right: Cross section of estimated state at $t = 200$ days and approximately 40^o lon using our hybrid algorithm with $\alpha = 0.8$.

Finally, we compare four algorithms by calculating the relative root mean square error at analysis times:

$$\frac{||\mathbf{x}_k^t - \mathbf{x}_k^a||_2}{||\mathbf{x}_k^t||_2}.$$

Here, $\mathbf{x}_k^t$ and $\mathbf{x}_k^a$ are true and analysis fields respectively at the analysis time t_k.



Upper Left : Relative RMS error of SSH field for constant covariance matrix, SEIK, local SEIK and our hybrid SEIK algorithm with $\alpha = 0.8$.

Upper Right: Relative RMS error of zonal velocity field for constant covariance matrix, SEIK, local SEIK and our hybrid SEIK algorithm with $\alpha = 0.8$.

Lower Left : Relative RMS error of meridional velocity field for constant covariance matrix, SEIK, local SEIK and our hybrid algorithm with $\alpha = 0.8$.

Lower Right: Relative RMS error of temperature field for constant covariance matrix, SEIK, local SEIK and our hybrid SEIK algorithm with $\alpha = 0.8$.

5 Conclusion

The covariances derived from the ensemble Kalman filter methods are nonstationary and anisotropic. However, in order to have good estimation results a large number of ensemble members need to be evolved in general. We have modified the SEIK algorithm in order to incorporate stationary covariance. This hybrid SEIK filter algorithm performs better than either SEIK or use of constant error covariance during assimilation. However, with the same number of ensemble members the accuracy of the local SEIK is superior to the accuracy of the hybrid algorithm developed here.

References

[1] I. Hoteit and D. T. Pham and J. Blum (2002): A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific, *J. Mar. Syst.*, **36**, 101-127.

[2] T. M. Hamill and C. Snyder (2000): A hybrid ensemble Kalman filter-3D variational analysis scheme, *Mon. Wea. Rev*, **128**, 2905–2919.

[3] L. Nerger and S. Danilov and W. Hiller and J. Schroter (2006): Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter, *Ocean Dynamics*, **56**, 634-649.

[4] D. T. Pham (2001): Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Mon. Wea. Rev.*, **129**, 1194-1207.

[5] X. Wang and T. M. Hamill and J. S. Whitaker and C. H. Bishop (2007): A Comparison of Hybrid Ensemble Transform Kalman FilterOptimum Interpolation and Ensemble Square Root Filter Analysis Schemes, *Mon. Wea. Rev.*, **135**, 1055-1076.

[6] X. Wang and C. Snyder and T. M. Hamill (2007): On the Theoretical Equivalence of Differently Proposed Ensemble3DVAR Hybrid Analysis Schemes, *Mon. Wea. Rev.*, **135**, 222-227.