On the Relation of the SEIK and ETKF Assimilation Methods

Alfred Wegener Institute for Polar and Marine Research

Lars Nerger, Tijana Janjić, Wolfgang Hiller, and Jens Schröter

Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

Contact: Lars.Nerger@awi.de · http://www.awi.de

XY351

Introduction

During the recent years several ensemble-based Kalman filter algorithms have been developed that have been classified as ensemble square-root Kalman filters. Parallel to these developments, the SEIK (Singular 'Evolutive' Interpolated Kalman) filter [1] has been introduced. Some publications note that the SEIK filter is an ensemble Kalman filter or even an ensemble square-root Kalman filter.

We discuss the relation of the SEIK filter to ensemble square-root Kalman filters in more detail. For this, we compare the SEIK filter with the Ensemble Transform Kalman Filter (ETKF) [2]. The comparison is conducted on the algorithmic formulations as well as in an application to the nonlinear Lorenz96 [6] model.

Comparison of Filters

The equations for the SEIK and ETKF algorithms are displayed on the right hand side. Because the equations are very similar, one has to be careful when comparing the algorithms.

- ETKF uses the ensemble perturbation matrix $\mathbf{Z}$ to represent the estimated error space. In contrast, SEIK uses the basis of the error space in matrix $\mathbf{L}$, which has one column less than $\mathbf{Z}$.
- The ensemble transformation is computed in different spaces. Matrix $\mathbf{A}$ of the SEIK filter is smaller than $\tilde{\mathbf{A}}$ of ETKF by one row and one column. Nonetheless, both contain the same information on the error space.
- The ensemble in the SEIK filter is reduced by one to the basis of the error space. Thus, the last member of the analysis ensemble has to be re-generated from this information. This is performed by the matrix Ω.
- SEIK and ETKF compute the analysis state $\mathbf{x}^a$ using the same error space information. Due to this, the analysis states are identical, if the same forecast ensemble and the same set of observations is used.
- In addition, the analysis ensembles of both filter algorithms will be equal when a particular choice for the matrix Ω is used. It is obtained when the Householder reflection orthogonal to the vector $(1,\ldots,1)^T$ is applied to the identity matrix.
- When Ω is chosen to be a random matrix, it serves for the randomization of the analysis ensemble which is sometimes suggested to avoid a loss of rank in the analysis ensemble.

Conclusion

- The SEIK filter is an ensemble square-root filter like the ETKF. ETKF uses all ensemble perturbations to represent the error space, while SEIK directly uses a basis of it.
- With deterministic transformations, SEIK and ETKF become equivalent. Then, they result in the same analysis ensemble. This is the case if both filters use the symmetric square root of the transformation matrix $(\mathbf{A}, \tilde{\mathbf{A}})$ and SEIK uses a matrix Ω that projects from the error space to the ensemble space.
- An assimilation experiment with the Lorenz96 model showed small differences in the estimated state for both the SEIK and ETKF filters. They are caused by the finite numerical precision of the computations, in particular singular value decompositions.
- These findings unify the separate developments that have been performed for the ensemble square-root Kalman filters and the SEIK filter.

Filter Equations for SEIK and ETKF

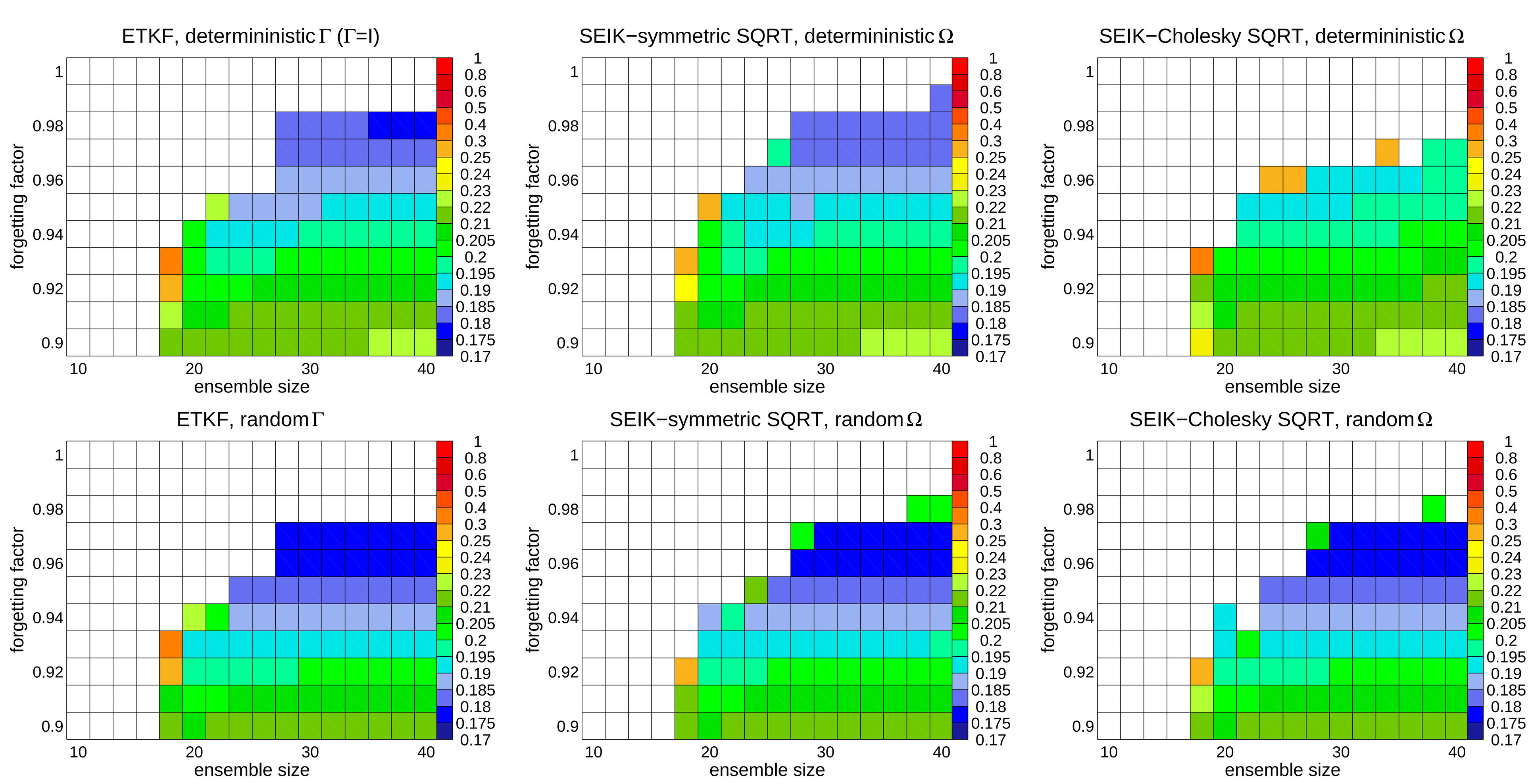
(The equations mostly follow the notations of [4] and [5])

	SEIK	ETKF
Some definitions		
State vector	$\mathbf{x}^a \in \mathbb{R}^n$	equal to SEIK
Ensemble of N members	$\mathbf{X}^a = \left[\mathbf{x}^{a(1)},\ldots,\mathbf{x}^{a(N)}\right], \quad \mathbf{X}^a \in \mathbb{R}^{n\times N}$	equal to SEIK
Perturbation matrix	$\mathbf{Z}^a = \mathbf{X}^a - \overline{\mathbf{X}^a}, \quad \overline{\mathbf{X}^a} = [\overline{\mathbf{x}}^a,\ldots,\overline{\mathbf{x}}^a]$	equal to SEIK
Analysis covariance matrix	$\mathbf{P}^a = \frac{1}{N-1}\mathbf{Z}^a(\mathbf{Z}^a)^T$	equal to SEIK
Error subspace basis	$\mathbf{L}^f = \mathbf{X}^f\mathbf{T}, \quad \mathbf{L}^f \in \mathbb{R}^{n\times(N-1)}$	not used in ETKF
T-matrix	$\mathbf{T} = \begin{pmatrix}\mathbf{I}_{(N-1)\times(N-1)} \\ \mathbf{0}_{1\times(N-1)}\end{pmatrix} - \frac{1}{N}\left(\mathbf{1}_{N\times(N-1)}\right)$	not used in ETKF
Analysis covariance matrix	$\mathbf{P}^a = \mathbf{L}^f\mathbf{A}(\mathbf{L}^f)^T$	$\mathbf{P}^a = \mathbf{Z}^f\tilde{\mathbf{A}}(\mathbf{Z}^f)^T$
with transformation matrix	$\mathbf{A} \in \mathbb{R}^{(N-1)\times(N-1)}$	$\tilde{\mathbf{A}} \in \mathbb{R}^{N\times N}$
	$\mathbf{A}^{-1} = (N-1)\mathbf{T}^T\mathbf{T} + (\mathbf{H}\mathbf{L}^f)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{L}^f$	$\tilde{\mathbf{A}}^{-1} = (N-1)\mathbf{I} + (\mathbf{H}\mathbf{Z}^f)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}^f$
State analysis	$\mathbf{x}^a = \overline{\mathbf{x}^f} + \mathbf{L}^f\overline{\mathbf{w}}^{SEIK}$	$\mathbf{x}^a = \overline{\mathbf{x}^f} + \mathbf{Z}^f\overline{\mathbf{w}}^{ETKF}$
with weight vector	$\overline{\mathbf{w}}^{SEIK} = \mathbf{A}(\mathbf{H}\mathbf{L}^f)^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f}\right)$	$\overline{\mathbf{w}}^{ETKF} = \tilde{\mathbf{A}}(\mathbf{H}\mathbf{Z}^f)^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f}\right)$
Square-root of analysis covariance matrix	$\mathbf{Z}^a = \mathbf{L}^f\mathbf{W}^{SEIK}$	$\mathbf{Z}^a = \mathbf{Z}^f\mathbf{W}^{ETKF}$
with weight matrix	$\mathbf{W}^{SEIK} = \sqrt{N-1}\,\mathbf{C}\Omega^T$	$\mathbf{W}^{ETKF} = \sqrt{N-1}\,\tilde{\mathbf{C}}\Gamma$
and square-roots $\mathbf{C}$, $\tilde{\mathbf{C}}$	$\mathbf{C}\mathbf{C}^T = \mathbf{A}$	$\tilde{\mathbf{C}}\tilde{\mathbf{C}}^T = \tilde{\mathbf{A}}$
Matrices Ω and Γ	Ω can be an arbitrary $N\times(N-1)$ matrix with orthogonal columns orthogonal to $(1,\ldots,1)^T$.	Γ is a random rotation matrix or the identity.
Ensemble transformation	$\mathbf{X}^a = \overline{\mathbf{X}^a} + \mathbf{L}^f\mathbf{W}^{SEIK}$	$\tilde{\mathbf{X}}^a = \overline{\mathbf{X}^a} + \mathbf{Z}^f\mathbf{W}^{ETKF}$

$\mathbf{C}$, $\tilde{\mathbf{C}}$ can be the symmetric square root $\mathbf{C} = \mathbf{U}\mathbf{S}^{-1/2}\mathbf{U}^T$ from the singular value decomposition $\mathbf{U}\mathbf{S}\mathbf{V} = \mathbf{A}^{-1}$. Alternative square-roots like a Cholesky factorization are possible.

Assimilation Experiment

Twin experiments were conducted using the nonlinear Lorenz96 model [6]. Synthetic observations of the full state were generated from a model run. The ETKF and the SEIK filter were used to assimilate the observations at each time step over 50000 time steps. For SEIK, the configuration was used that makes it equivalent to the ETKF (see "Comparison of Filters" on the left) as well as a configuration with a square-root based on Cholesky decompostion. The global formulations of SEIK and ETKF were used. These were sufficient for the Lorenz96 model, but require larger ensembles than localized filters for comparable performance.



The mean RMS errors over 10 experiments depending on ensemble size and forgetting factor (covariance inflation). The SEIK filter configured to be equivalent to ETKF provides an almost identical result to the ETKF. The small differences are statistically not significant and caused by singular value decompositions of the matrices $\mathbf{A}$ and $\tilde{\mathbf{A}}$, which have different condition numbers. Errors from the SEIK filter using a Cholesky decomposition of the transformation matrix $\mathbf{A}$ are larger. This is caused by an inferior ensemble quality in which a small number of ensemble members carry most of the variance. With random transformations, the error levels become equivalent.

References

[1] DT Pham et al. (1998). Singular evolutive Kalman filters for data assimilation in oceanography. *C. R. Acad. Sci.* Series II 326: 255–260

[2] CH Bishop et al.. (2001). Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.* 129: 420–436

[4] L Nerger and WW Gregg. (2007). Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *J. Mar. Syst.* 68: 237–254

[5] SC Yang et al. (2009). Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *Q. J. Roy. Met. Soc.* 135: 251–262

[6] EN Lorenz. (1996). Predictability - a problem partly solved. *Proceedings Seminar on Predictability*, ECMWF, Reading, UK, 1–18.