Nonlin. Processes Geophys. Discuss., 1, C150–C153, 2014
www.nonlin-processes-geophys-discuss.net/1/C150/2014/

Nonlinear Processes in Geophysics
Discussions
Open Access

NPGD
1, C150–C153, 2014

Interactive Comment

Interactive comment on "Data assimilation of two-dimensional geophysical flows with a Variational Ensemble Kalman Filter" *by* Z. Mussa et al.

Z. Mussa et al.

Zubeda.Mussa@lut.fi

Received and published: 21 May 2014

i) " It is unclear from reading the article what more using the QG model with the scheme reveals about the performance of the scheme"

RESPONSES - The two-layer Quasi-Geostrophic model is a chaotic model that in addition can be made large-scale by adjusting the density of its spatial discretization grid. This is the principal difference from the model studied in the work by Solonen et al. (2012), where the first model was chaotic, but low-dimensional and the second was linear, which means that it only measures the filter's performance (by increasing the scale), but not the ability of the filter to remain stable in case of large-scale dynam-

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

ics. Since the main advantage of the VEnKF over EKF is its ability to circumvent the memory issues caused by covariance storage, such approximation as the VEnKF is must be tested using a large-scale chaotic model. For VEnKF such benchmarks have never been performed before. The QG-model was selected as the one that fills the gaps left by the paper by Solonen et al. (2012). In addition, we should emphasize that the very same model is used as a bench-marking case for data assimilation methods in ECMWF (see ECMWF 2010-2011 report)

ii) " From the results shown it seems to me that the point here is rather more to do with the ability of data assimilation to also enable the capture of cross-flows rather than the particular ability or advantage of the VEnKF."

VEnKF doesn't offer any advantage over the EKF except its low-memory storage design, and it is fully the property of data assimilation to account for phenomena that were not properly captured by the prediction model. For this reason the choice of VEnKF is mainly based on its ability to handle large scale problems, as the shallow water model has about 16000 degrees of freedom, which would be run very inefficiently with the EKF.

iii) "Thoughout the paper it is not also not clear from the description given how the observstions are incorporated into the data assimilation scheme."

RESPONSES - a)For data assimilation performed on top of the Quasi-Geostrophic model we begin by selecting a random set of 100 nodes from the spatial grid of the model. From these nodes we collect explicit values of the stream function perturbed by artificial observation noise. Selection of the observation sources is done one time and kept fixed during the whole data assimilation period. We should mention, that decreasing number of observations deteriorate convergence properties of the filter. However, in VEnKF (as well as in the other ensemble-based approximate Kalman filters) the problem can be alleviated up to some extent by increasing the number of ensemble members. However, for this study no special research for the case of the qg-model has been done regarding the number of observations and ensemble cardinality combinations that allow the filter to converge.

b) No, when observation interval was increased the model time step remain the same, defined throughout the paper as delta t = 0.103. What has been changing is the observation frequency (assimilation window). The solution between the observations was used as the initial condition for the next time point forecast, which we must not save as at the end we make comparison between the fit and the measurements. This emphasizes the fact that there are no multiple time steps between measurements.

c) With the VEnKF algorithm it is necessary to interpolate highly sparse observations lest the filter diverge. The reason why the covariance matrix Cp alone is not able to handle this comes from the small size of the ensemble relative to the size of the system, and sparsity of raw, un-interpolated observations, so that there is substantial sampling noise in Cp, unlike with EKF. The interpolation kernel over space and time we use is similar to the various localization techniques employed in almost all Ensemble Kalman filters that are usually justified as a way to avoid spurious long-distance correlations. It bears close resemblance to the interpolation kernels used in classical Optimum Interpolation data assimilation.

iv) "About the final section of the paper"

RESPONSES - Here we admit an error in our calculation for the measurements' average time interval, as the frequency was 7.07X10-1, it follows that the period = 1.4s. This time period is much longer than model time step of 0.1s, hence justifying time interpolation of the measurements. The reviewer is absolutely right in pointing out to the strong constraint that dense observations bring to the assimilation process. When observations are interpolated, they also bring smoothness prior to the Bayesian estimation process in tow, over both space and time. In these tests of filter divergence and density of observations, we try to see if there is some general law that would govern the ratio between the temporal density of observations and the spread of the ensem-

NPGD

1, C150–C153, 2014

ble, since both impose constraints on the local smoothness of the analysis. The result obtained seems to indicate that such a law exists, testifying to the complementary role of ensemble spread and observation interpolation in data assimilation with Ensemble methods.

Interactive comment on Nonlin. Processes Geophys. Discuss., 1, 403, 2014.

NPGD

1, C150–C153, 2014

Interactive Comment