Nonlin. Processes Geophys. Discuss., 2, C373–C376, 2015
www.nonlin-processes-geophys-discuss.net/2/C373/2015/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation" *by* M. Bocquet et al.

Anonymous Referee #2

Received and published: 3 September 2015

Comments on "Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation" by Bocquet, Raanes and Hannart.

This paper amends and extends the EnKF-N algorithm previously described in several other papers by the first author and collaborators. The context of the algorithm is very nice in that it explicitly recognizes and samples the errors in the EnKF due to the use of finite size ensembles and nonlinear dynamics. By using a hyperprior assumption to account for these errors, the authors are able to bring analytic insight to a number of more ad hoc techniques (inflation and localization) that are commonly used in EnKF applications.

Parts of the mathematics in the presentation are beyond my capability to check so I cannot attest to the correctness of the derivations. However, the demonstrated low-

order model performance is compelling and the improvements to the earlier erroneous derivations are obvious.

In my opinion, the authors would do a service if they could supply a little bit more information about the expected computational cost in both time and storage to apply the EnKF-N with the various hyperprior assumptions as a function of the number of degrees of freedom in the dynamical model. A comparison to the cost of the EnKF algorithm would be enlightening.

Minor comments and questions include:

1. P. 1092, Line 26: Localization can be required even if not rank deficient for EnKFs or in presence of non-gaussian/non-linear effects. 2. P. 1093, Line 18: Traditional methods only make assumptions about first two moments, not all moments as implied here. 3. P. 1094, line 10: I don't know what the authors mean by doing the analysis in "ensemble" space. This should be clarified here before the subsequent use. I am aware of terminologies like "model" and "observation" space, but don't know how this relates. 4. P. 1094, line 16: Not clear why the posterior should relax to the prior for "quasi-linear". 5. P. 1095, line 6: Don't see why non-linear dynamics has to be partially responsible. Could this not all be due (at least for second moments) to a degenerate ensemble? 6. P. 1099, line 20: This assumption is only justified in the case that the ensemble is approximately non-degenerate. Is that ever the case in any real applications you have in mind? 7. P. 1106, line2: "Not as performant…" You should probably say a little bit more about what you mean by this statement and what experiments you performed to explore it. In particular, it seems that the EnKF-N corresponds to a single inflation value for the whole ensemble, while some of the other methods allow different values for different variables. In a nonlinear model system, the latter could clearly have advantages. 8. P. 1107, paragraph starting at line 14. This paragraph was unclear. Obviously violation of the EnKF assumptions can come either from nonlinearity or degeneracy, and which dominates (or even exists) depends on the application. Here, you seem to be saying that nonlinearity is somehow the

generic cause. As an aside, what happens if you apply the EnKF-N in a case where the EnKF is sufficient (linear, Gaussian, ensemble big enough to span growing and neutral directions)? 9. P. 1109, bottom. You need to give a little bit more details about these comparison experiments, in particular stating that the ensemble sizes were the same, and possibly commenting on the relative computational cost. 10. P. 1110, line 14: Nonlinearity is the "profound cause" only because the ensemble size of 20 is larger than the number of positive local Lyapunov exponents? 11. P. 1110, line 24: You need to give a bit more information about the 2D barotropic model. 12. Section 6, line 1: This barely nonlinear regime is truly problematic for deterministic ensemble filters, but is known to be a problem for EnKFs in low-order models (see for instance Anderson 2010 on nongaussian filter updates). However, the problem there generally goes away with larger models. Is that anticipated for the issue here? 13. Section 6.2, start: This notion of "relaxing to prior" seemed confusing to me. The observations have no information, so the posterior is the prior. I'm not sure what you are "relaxing" from in this discussion. 14. P. 1115, line 20: I think that this use of "climatological" is misleading. The required statistics do not come from the climate of the model as this terminology would normally imply. Instead, they are statistics from the "climatology" of the prediction system including the assimilation. To sample them, one would need to run a high-quality assimilation system (large ensemble, well-tuned) and sample the statistics from that. 15. P. 1124, first paragraph. The practical use of the methodology is to avoid the need to tune multiplicative inflation for perfect model experiments. This avoids the cost of doing multiple runs to tune the inflation. However, given that this is a major result of the paper, there needs to be a little bit more discussion of the computational cost (time and storage) to implement the EnKF-N compared to a basic EnKF. Some of the adaptive inflation methods already in the literature that are referenced in the paper are generally able to produce smaller RMSE that the best tuned single inflation value for the dynamical systems examined here. These methods have very small incremental cost compared to the base EnKF. It would also be important to indicate the expected scaling of the computational cost for the EnKF-N versus the EnKF as the model size

grows.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 1091, 2015.