

Nonlin. Processes Geophys. Discuss., referee comment RC1
<https://doi.org/10.5194/npg-2021-8-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on npg-2021-8

S.G. Penny (Referee)

Referee comment on "Multivariate localization functions for strongly coupled data assimilation in the bivariate Lorenz 96 system" by Zofia Stanley et al., Nonlin. Processes Geophys. Discuss., <https://doi.org/10.5194/npg-2021-8-RC1>, 2021

General Comments:

The development of localization schemes for coupled dynamics is an important activity that needs increased attention as operational forecast centers transition to greater reliance on coupled Earth system forecast models. The authors provide a promising advancement to address the localization of cross-domain error correlations. I believe the work should be published after the authors explore a larger parameter space for their experimental results, as described below. In exploring a larger parameter space, it may be sufficient to focus on one or two leading methods (e.g. GC and BW).

One concern is the choice of model, and how well the results can transfer to more realistic scenarios, given the near linear relationship between the slow and fast components in this system (e.g. see S. Rasp note referenced below). Do the authors have confidence that the results can translate in some way to more sophisticated systems? I would be interested to know how the results change as the coupling strength between the slow and fast components is weakened or strengthened from the baseline state used by the authors.

A second concern is the restriction to observing only the fast dynamics. I would like to see

a complete investigation examining the observation of fast-only, slow-only, and the full slow-fast coupled system. I'd like to see Figure 3 repeated for a few different scenarios, including those just mentioned, but also potentially varying parameters of the EnKF, such as the frequency of observations, the density of observations, the amount of observation noise, the length of the analysis cycle, etc. Not all results need to be reported in figures, but some indication that the authors have explored more variations in the problem specification would help to build confidence in the robustness of the final reported results.

Minor issue: There are a few instances where the present tense is used when it should be past tense.

Specific Comments:

L 8:

"The functions produce non-negative definite localization matrices, which are suitable for use in

variational data assimilation schemes."

I think the term 'positive semidefinite' is more common, and the one originally used by Gaspari and Cohn (1999). I would suggest changing all instances of this throughout the manuscript.

L 14-16:

"The background error covariance statistics stored in B dictate how information from observations propagates through the domain during the assimilation step (Bannister, 2008)"

The term 'propagates' seems appropriate for 4D-Var, but perhaps not for all DA methods. More generally, the background error covariance provides a structure function that determines how observed quantities affect the model state variables, which is of particular importance when the state space is not fully observed.

L 25:

"Localization is typically incorporated into an ensemble estimate of B through a Schur (or element-wise) product."

I would change this to say that localization is typically incorporated into the data assimilation in one of two ways - either through the B matrix using a Schur product, or through the observation error covariance R (e.g. Greybush et al., 2011). You are focusing on the localization applied directly to the B matrix.

Greybush et al., 2011: Balance and Ensemble Kalman Filter Localization Techniques.

<https://journals.ametsoc.org/view/journals/mwre/139/2/2010mwr3328.1.xml>

L 32-33:

“In Earth system modeling in particular, coupled DA shows improvements over single domain analyses (Penny et al., 2017; Zhang et al., 2020)”

Additional sources that determined this point clearly are Sluka et al. (2016) and Penny et al. (2019):

Sluka, T., S.G. Penny, E. Kalnay, and T. Miyoshi, 2016: Using Strongly Coupled Ensemble Data Assimilation to Assimilate Atmospheric Observations into the Ocean. *Geophys. Res. Lett.*, 43, doi:10.1002/2015GL067238.

Penny, S.G., E. Bach, K. Bhargava, C-C. Chang, C. Da, L. Sun, T. Yoshida, 2019: Strongly coupled data assimilation in multiscale media: experiments using a quasi-geostrophic coupled model. *Journal of Advances in Modeling Earth Systems*, 11.
<https://doi.org/10.1029/2019MS001652>
<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019MS001652>

L 35-37:

"Schemes that include cross-domain [error] correlations in the B matrix are broadly classified

as strongly coupled, which is distinguished from weakly coupled schemes where B does not include any nonzero cross-domain [error] correlations. The inclusion of cross-domain [error] correlations in B offers advantages"

To be more precise, the term "cross-domain error correlations" should be used if referred to the error covariance matrix B.

L 55:

I'll note that Lorenz himself cited this as (Lorenz, 1996). See comment below regarding line 461.

L 57-58:

"We find that, in our set up, artificially decreasing the magnitude of the cross-domain correlation hinders the assimilation of observations."

This is a positive sign for the advancement strongly coupled DA, but I wonder if this could be partly due to the use of the Lorenz system II, which has some highly linear relationships between the small and large scale systems. Some discussion was given, for example, in this blog post by Stephan Rasp:

<https://raspstephan.github.io/blog/lorenz-96-is-too-easy/>

L 61:

"localization function[s] from the literature."

L 77-78:

"A fundamental difficulty in localization for strongly coupled DA is how to propose a cross-localization function L_{XY} to populate both L_{XY} and L_{YX} "

It might be useful to explain at this point which term controls the effect of system X on Y, and Y on X.

L 102:

"we define two processes Z_j , $j = X, Y$ "

I understood this on the third read through. Perhaps the authors could reword this sentence slightly to make it more clear. For example,

"we define two processes Z_j , where j can represent either X or Y"

Or simply,

"we define two processes Z_j , with $j=X,Y$ "

L 105-106:

" Thus L_{XX} , L_{YY} , L_{XY} form a multivariate covariance function, and hence a multivariate, non-negative definite function"

Based on the terminology defined so far, I'm not sure how to interpret the triple (L_{XX}, L_{YY}, L_{XY}) forming a single function. Perhaps a line or two could be added to explain this step.

L 118:

The way I am interpreting the notation is that the term $(1-r/c)_+$ is zero when the term in parentheses is less than or equal to 0, which would occur when $r \geq c$. Can the authors explain the comment about the convolution being zero at distances greater than $2c$ in line 120, it is not immediately obvious.

L 139:

"who perform[ed] the"

L 140:

"in never develop[ed] multivariate"

L 156:

"Porcu et al. (2013) develop[ed] a multivariate version"

"et al." is short for the Latin term "et alia," meaning "and others." It is strange to reference the actions of Porcu "and others" in the year 2013 using present tense.

L 157:

"Roh et al. (2015) [found] that"

L 159:

"Daley et al. (2015) extend[ed] the work"

L 166:

"with B the beta function"

Could you define this here for clarity.

L 168:

"Daley et al. (2015) [gave]"

L 184:

"This approach leads to a "weakly" coupled scheme, which is not the focus of this work."

I understand this may not be the focus, but it seems that it would be appropriate to compare to this approach given that the weakly coupled DA scheme is the standard approach for current operational forecast systems.

L 184-186:

"Additionally, in our setup we observe only one of the two processes and we find that when the assimilation is not allowed to update the unobserved process the result is prone to catastrophic divergence"

It might be appropriate to perform a few experiments where both components are observed, and results are compared using weakly and strongly coupled DA.

L 200/202/205:

"Lorenz (199[6])"

See comment below for line 461.

L 203:

"using an adaptive fourth-order Runge-Kutta method"

Perhaps provide a citation for the method.

L 204-205:

"The solutions are output with a time interval of 0.005 nondimensional units, or 36 minutes"

It seems strange to say there are non-dimensional units and then indicate that it is the same as 36 minutes. Perhaps repeat some of the justification from Lorenz to indicate the relative error growth rates and its relation to more realistic applications that would be approximately equivalent to 36 minutes in operational prediction in the early 1990's.

Figure 2 caption:

"setup" is a noun that means "the way in which something... is organized, planned, or arranged." This should probably be used in most places where the authors current use two words: "set up".

L 209:

"Increasing the coupling strength leads to larger covariances between the forecast errors in processes X and Y , thereby making the effect of cross-localization more pronounced and easier to study."

I believe this is the case. However, I would like to see some sensitivity study of how the benefit of strongly coupled DA paradigm breaks down as the coupling strength between the two components weakens and asymptotes to 0.

L 213:

"We choose to place the variable X_k in the middle"

Does the placement of the X variable have any influence on the results of localization? Is there any sensitivity here, or are the results generally the same regardless of how the placement of the X and Y variables are interpreted?

L 218:

"We develop localization functions for EnVar schemes where non-negative definiteness of the localization matrix is essential to ensure convergence of the numerical optimization. Since the minimizer of the 3D-EnVar objective function is the same as the EnKF analysis mean in the case of linear observation (Lorenc, 1986), in this experiment we make use of the EnKF rather than implement an ensemble of 3D-EnVar assimilation scheme (Evensen, 1994; Houtekamer and Mitchell, 1998; Burgers et al., 1998)"

This discussion is a bit confusing. I think it could be cleaned up with a little reorganization, e.g.

We develop localization functions for data assimilation schemes that rely on Schur product modification of the background error covariance matrix B . In our experiments we use the stochastic EnKF (Evensen, 1994; Houtekamer and Mitchell, 1998; Burgers et al., 1998). However, because the minimizer of the 3D-EnVar objective function is the same as the EnKF analysis mean in the case of linear observations (Lorenc, 1986), our results translate to EnVar schemes as well. The positive semi-definiteness of the localization matrix is essential to ensure convergence of the numerical optimization methods used to implement EnVar <cite>.

L 230-231:

"In this experiment we use the adaptive inflation scheme of El Gharamti (2018) and apply the inflation to the prior estimate."

Can this be added to the EnKF equations above for clarity?

L 233-234:

"We run each DA scheme for 3,000 time steps, discarding the first 1,000 time steps and reporting statistics from the remaining 2,000 time steps."

Is this referring to model time steps, or the number of analysis cycles?

L 235-237:

"The observation operator H is such that all of the Y variables are observed, and none of the X variables are observed. In this way we can isolate the effect of the localization on the performance of the filter for the X variable."

This means you are observing the fast dynamics and using this to update the slow dynamics through the error covariance statistics. This has been shown effective in a number of studies exploring strongly coupled DA. Penny et al. (2019) showed that the reverse was also possible, particularly if the size of the analysis window is decreased (or the frequency of observation updates is increased).

L 256:

“, so that we hypothesize that GC allows”

Change to:

“, so we hypothesize that GC allows”

L 266-267:

“ By contrast, the BW and Askey functions show virtually no difference between the multivariate and univariate versions”

The BW method looks slightly improved, and the Askey method slightly degraded.

L 280:

It is up to the authors, but generally the conclusions reads more clearly if this is now written in past tense. E.g.

"In this work, we develop[ed]..."

"We compare[ed] multivariate GC to three..."

"We [found] that, in a toy model..."

"In this work we investigate[d]..."

"We [found] that this..."

L 284:

"the localized estimate of the background [error] covariance matrix"

L 294:

"A natural application of this work is localization in a coupled atmosphere-ocean model. Multivariate GC allows for within component covariances to be localized with GC exactly as they would be in an uncoupled setting, using the optimal localization length scale for each component Ying et al. (2018). In this work we discuss the importance of the cross-

localization radius in determining performance. However, this work does not address the question of optimal cross-localization radius selection, which is an important area for future research”

This is certainly of interest - are there any conjectures that can be made about the applicability of the results here extending to an application like a coupled atmosphere-ocean model?

While the interpretation of localization is clear in the within-component covariances, how would you interpret localization on the cross-component covariances? Could a situation in which it might be desirable for an atmospheric observation to have an influence on the ocean state but not vice versa create difficulties with the symmetry relied on above in forming LXX, LYY, and LXY as a triple, and the need for maintaining positive semi-definiteness?

L 461:

The full citation for Lorenz-96 is not given. It should be Lorenz (1996):

Lorenz, E.N., 1996: Predictability—A problem partly solved. Proc. Seminar on Predictability, Vol. 1, Reading, Berkshire, United Kingdom, ECMWF, 1–18.

Note that Lorenz cited it himself this way in:

Lorenz, E.N., 2005. Designing chaotic models. *J. of the Atmos. Sci.* 62, 1574–1587.
DOI:10.1175/JAS3430.1.