Hybrid ensemble-variational data assimilation in ABC-DA within a tropical framework

Joshua Chun Kwang Lee, Javier Amezcua, Ross Noel Bannister

We are grateful for all the constructive comments provided by the reviewers which has helped improve the readability and clarity of the manuscript. We hope that the reviewers are sufficiently satisfied by our amendments.

Reviewer 1

General Comments

The paper provides the description of the implementation of a specific data assimilation algorithm (Hybrid En-Var) in an existing toy-model (ABC-model covering an x-z slice of the atmosphere) and exemplarily shows the performance of the DA-system in a tropical setup of the model.

The paper focusses on the description of the algorithms which are not new but when applied to this specific model setup may have potential to investigate specific issues of the formulation, setup, tuning and characteristics of data assimilation systems. Studies of simplified models can provide considerable insight into the performance of data assimilation algorithms under specific conditions and thus I think it is worth describing the system in a scientific paper. The description of the implementation is outlined very clearly and detailed and is reproducible. Indeed I think that some passages may be shortened without loss of information. Maybe the authors could try to strengthen the text in this sense.
We thank reviewer 1 for this and all other comments. We have run through the manuscript to write more concisely.

The performance of the system is outlined by example of a simulation encountering tropical convection. This is certainly a very specific application and in order to allow the reader to grasp the situation it would be nice to give more information on the case and the setup: The evaluation of the system by means of the above example points at some deficiencies of the setup (for instance the choice of the climatological B_c). This is not a detriment of this paper as the system described here basically should be a tool to gain insight into the possibilities tuning of the setup. Also here some more information should be provided on the specific setup of the B_c and B_e matrices, maybe some more (cross-)correlations implied by them, also by the revised B_c matrix which was mentioned (derived after the spin up), or by contrasting the (time depending) B_e

correlations with the actual physical fields at that time. Basically the choices made for B_c and B_e variances would deserve more discussion.
We have added discussion on the tuning of B_c. Some of the changes here overlap with other changes in response to specific comments below.

More details will be proposed in the specific comments below. I think the paper would gain from strengthening the description of the algorithm itself but providing more insight into the details of validation test case.

Scientific significance:

Good, the described data assimilation algorithms algorithms are not new but there implementation in this specific toy-model setup has potential to study details of the data assimilation setup and application to specific situations.

Scientific quality:

Fair, The methods are well described. The description experimental setup would improve with some more details given. The setup of the illustrative test case could gain from further informations and discussion of the tuning parameters involved.
For the implied covariances illustration, we have added further explanation. The test case is meant to demonstrate the localisation and effect of the different transforms on the same training data. For the parameters used in the experiments, we have included justification on the horizontal localisation length-scale.

Scientific reproducibility:

Good for the description of the implementation of the algorithms (reproducible). The testcase cannot be assessed that well.

Presentation quality:

Good

Specific Comments

- 3.2.3 Inter-variable and spatial localisation

The authors state that it is possible to achieve inter-variable localisation within this setup. I think it should be mentioned that the strength of ensemble systems is to provide reasonable time-dependent inter-variable correlations and that therefore one should have good reasons to apply inter-variable localisation in praxis.
We agree that ‘turning off’ a feature introduced by ensemble DA might seem strange. We believe there are good reasons for doing this though, e.g.:

1. Testing a system for consistency – setting zero correlation between two different variables is one tool to do this in a testing phase. This can also be used to isolate important multivariate impacts of ensemble DA, and use different localisation length-scales for each variable. We have added an additional sentence to explain how there might be benefits for convective-scale data assimilation.

2. Testing the impact of (specifically) the multivariate impact of ensemble DA (in isolation), related to the previous point. Applying full inter-variable localisation switch would be useful to study this effect.

3. In circumstances where there is a known (and unaccounted for) bias in one variable, resulting (large and polarised) analysis increments in this biased variable would then likely damage other variables that are correlated. This would be a valid reason to 'turn off' the ensemble correlations between the biased variable and others. Although we don't have this problem here, this tool might be useful to readers. There are examples in non-ensemble schemes where stratospheric biases in T were causing problems with stratospheric water vapour.

The author state that L_horiz has been found to be not positive definite if the length scale is too large (localisation functions exceeds the cycling domain). The fix applied by the authors (setting negative eigenvalues of U^alfa seems problematic to me as the shape of the resulting L at the origin is not smooth). There are better ways to handle this problems:

1) The original article of Gaspari and Cohn shows how positive definite correlation functions can be designed on the sphere. This also works on a cycled domain.
Gaspari and Cohn (1999) demonstrates this for "space-limited covariance functions on S^2", where the cut-off length-scale does not exceed the diameter of S^2. In our case, we are demonstrating an extreme case where the cut-off length-scale (250 km) exceeds half the domain (violating the criteria to be considered "space-limited"). The fix applied in our paper is an engineered approach to allow smooth transition from full localisation (i.e. Kronecker delta at source point), to no localisation. In practice one should not be using such large length-scales; when the cut-off length-scales are small, we essentially recover the "space-limited" covariance function described by Gaspari and Cohn. In the experiments, we only used a length-scale of 20 km (and 100 km for illustration of implied covariances). We have added a short explanation in the manuscript on this point.

2) another option is to specify U^alfa as a Gaspari Cohn function with half the length scale as L. Then U^a*U^aT will again approximate a Gaussian with the required length scale.

In addition to the above response, regardless of whether we use a Gaussian or Gaspari-Cohn function, we found that as the localisation length-scale tends to infinity, the circulant matrix is not guaranteed positive-definite, because the covariance function is not "space-limited".

- 3.3 Generation of ABC analysis ensemble

When first introducing the EBV method the authors just mention that "the method .. is uninformed about the observational method". I think this issue should be discussed a little bit further, maybe when the ensemble spread is compared to the rmse error. Only if observation density and observation error are properly accounted for in the ensemble generation process (as done in some other ensemble generation processes mentioned in this section) it can be expected that ensemble spread and rmse matches.
We have included additional discussion on this. Such an approach is perhaps most theoretically justified in systems whose forecast errors saturate over the DA cycle period and so would be insensitive to initial condition errors (and hence details of the observation network). The EBV method is a pragmatic "control method" so that we can study the impact of accounting for the observations by comparing it to the EnKF approach (if required, in the future). In one sense it is a half-way-house between having completely static covariances and fully flow-dependent covariances that do explicitly account for the observations.

The same discussion applies to the estimation of the climatological matrix B_c. Also here as far as I see only possible balances are taken into account, but not the actual variance which actually depends on the data assimilation setup.
We have included additional explanation in Section 4.4.

- 3.3.1 Ensemble bread vectors

I thing the normalisation factor E_tot deserves more discussion. This is basically a tuning factor which fixes the ensemble spread. Deriving it from the mean energy norm of the ensemble of differences of independent realisations of the state vector (eq. 7b) would represent the climatological variance, not the uncertainty of the analysis.
The choice of \epsilon_0 based on E_tot is pragmatic since we do not have a long time-series of analyses. The current approach will get the most 'active' directions, but not necessarily the right amplitudes (energies). There may be alternatives, e.g. running a series of experiments and finding the average analysis error, but overall, designing a suitable tuning factor is not straightforward. We have added more comments in Section 3.3.1.

- 4.1 implied background error covariances.

I think the covariances shown in Figure 3 would deserve some more discussion. It is mentioned that balances and multi-variate relationships will be explored in a separate study but some more information would be helpful here.
We have added discussion on the topic.

The (time dependent) B_e covariances could be contrasted with physical fields at the respective time. Can the vertically alternating patterns in the B_c correlations be explained ?
We have added additional explanation. As the covariances do not reflect identifiable features in the physical fields, and to save space, we have not included plots of the physical fields. We hypothesize gravity wave processes are associated with the vertical oscillations.

- 4.2 Details of observing system simulation experiments

Figure 4 shows that the contribution of J_e to the total cost function is very small, even only 20% in case of only B_e used (and 80% contribution of J_o). Doesn't this indicate some insufficient tuning of the variances? It means that the contribution of the background in this experiment is quite limited.
In general, there is greater misfit of the analysis to observations and background (in a least-squared sense), so the ensemble penalty is very small. This is partly because the variances of the implied ensemble-derived background error covariance matrix are small compared to the variances of the implied static background error covariance matrix and observation error variances. We have included some discussion on the tuning of the variances, suggesting that having much smaller variances in the static background error covariances may help to improve the RMSE of the experiments. The small variances are in line with the expectation that the true background errors of the system are generally small, given that the free background run does not deviate substantially from the 'truth' run as the cycling progresses.

Having small variances does not necessarily represent a deficiency in the minimization and we are not aware of any reason why a small relative contribution from the background (or ensemble) term necessarily indicates a problem with the statistics. For instance, an example operational Met Office run that we looked at had J_b / (J_o + J_b) at the minimum = 0.14. Theoretically, a perfectly tuned DA system with one observation (easy to analyse) has expected value of J_b / (J_o + J_b) at the minimum between 0 and 1, depending on B and R, where a small value isn't necessarily a problem. We *do* know that the expected value of (J_o + J_b) at the minimum should be ~No. of obs / 2, which is what we emphasized in the paper.

- 4.3 Sensitivity to weighting of B_c and B_e

Figure 5 shows a variation of the RMSE on a time scale of 8h, figure 6 of the assimilated values itself, can that be explained ?

In the RMS of the zonal wind (u) and scaled density perturbation (rp) analysis fields (Fig. R1, below), one can see that there are oscillations in both the 'truth', the free background, and the experiment (EBVd) runs. EBVd tends to deviate from the 'truth' the most when the RMS peaks in the u and rp analysis fields, leading to the variation of RMSE on a time-scale of 8 hours. The 'peaks and troughs' in all u and rp RMS of analysis fields all have a period of 8 hours, suggesting that this is due to a dominant dynamical mode of the system (rather than associated with a mode that is not well observed, if it were visible only in the error plots). One can also notice the out-of-phase nature between the RMS of u and rp analysis fields. This indicates that the system is exchanging kinetic and elastic energy with a time-scale of 8 hours, and is consistent with wave dynamics. Visual inspection of the u analysis field in the free background run indicates local maxima (or minima) occurring every 16 hours, which gives rise to the time-scale of 8 hours (peak to trough, trough to peak) in the RMS fields. According to normal mode analysis, the period of 16 hours is within the range of low-zonal-wavenumber gravity waves, suggesting that this signal is due to dominant gravity waves in this system. We have added some comments on the periodicity in the manuscript, following the above discussion.

If the RMSE values of analysis errors are compared to the nominal observational errors the former appear to be very large. It would be illustrative to show the distribution of observations to better understand the performance of the data assimilation procedure. What are the forecast (background) errors.

We have added a comment on the number of observations as a percentage of the degrees of freedom of the state. The distribution is described at the start of the paragraph. The analysis errors are reducing over the 50 cycles, which indicate that the trajectories are still in the process of gradually converging around the "truth", even though the RMSE values start larger than nominal observation errors.

It is stated that B_c was re-calibrated using other training data after the spin-up process. Wouldn't it be appropriate to show the covariances for this matrix in figures 3, as they are actually used in the assimilation experiment ?

We have added clarification that calibrating B_c using other training data was to test if the issue with the v analysis errors could be resolved. We still used the original B_c for the experiments (shown in Figure 3, together with raw localized covariances).

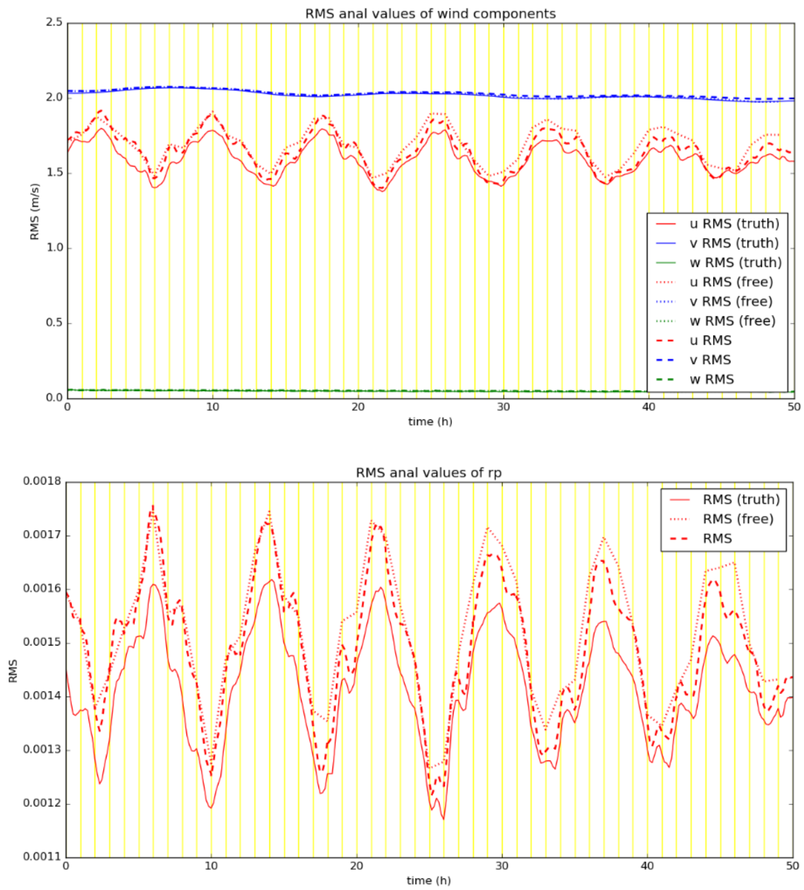

Figure R1: Root-mean-square (RMS) of analysis fields of zonal wind (u), meridional wind (v), vertical wind (w; top panel), and scaled density perturbation (rp; bottom panel) for 'truth' (solid), free background (dotted), and EBVd (dashed) runs.

Technical Corrections

- 3.2.3 Inter-variable and spatial localisation

It should be stated how exactly the length scale h for the localisation function is defined, there a several options: There the Gaspari?Cohn function goes to zero, based on the second derivation at the origin (as defined by Daley, ...

The exact equation used is included in the manuscript (Eq. 4.10). We have added a line to mention that h is equivalent to c in Eq. 4.10 (using a=1/2).

- 3.3 Generation of ABC analysis ensemble

The EBV-method is first mentioned in Section 3.3 but the synonym is defined not before section 3.3.1

We have made the corresponding changes.

- Figure 6:

It is hard to see the (gray) ensemble trajectories. The figures could be stretched in the vertical to better resolve this.
We have amended this to better show the grey trajectories. However, the spread of the ensemble is generally very small compared to the local variation (in RMS) of the fields, so the grey lines may not be immediately clear for all variables.

Reviewer 2

General comments:

A hybrid ensemble-variational DA algorithm is described here for the ABC model. The manuscript is mostly written clearly, and I enjoyed reading it.

The main aspect that is currently lacking, in my opinion, is a description of this study's importance, and the motivation. I think this is a critical part of any manuscript, and currently it is too much left for the reader to guess.
We thank reviewer 2 for highlighting this and for all other comments. We have added additional justification on the importance and motivation behind the study.

Specific comments:

1. As stated in the general comments above, I think the main change that is needed here is a clear description of this study's importance, and the motivation. It is needed in both the Abstract and Introduction.

What is the problem that you are trying to address? Why is it important to publish this manuscript?

A reader might guess, from what is currently written, that hybrid ensemble-variational DA algorithms are promising but relatively new, and that more studies are needed to understand implementation choices, properties, and performance. Any additional documentation of studies is then a valuable contribution to the literature.

Is that correct? If so, could you please add some description like this to the Abstract and Introduction? If not, could you please describe what, in your view, is the importance of this study, and the motivation for it?

The Summary section seems to include some statements along these lines, at Lines 626-630: "Given the rapid adoption and broad shift towards hybrid ensemble-variational methods in convective-scale numerical weather prediction, we hope that the ABC-DA system can prove useful in providing further insights and highlight other potential issues that may arise in such methods. Particularly for the tropics, further work is required to better understand the characteristics of the ensemble-derived background errors, such as disentangling its flow-dependency or designing the localisation to isolate or identify important multi-variate relationships."

That type of information from the Summary section should be made very clear to the reader in the Abstract and Introduction. Do not make your reader guess. Tell the reader exactly what you have in mind for the importance and motivation of your study.

Following the general comments, we have amended the abstract and introduction to emphasize that the main motivation of this study is to explore hybrid methods applied in the tropical context, which thus far has barely been explored.

2. I would not refer to your ABC model as a "toy" model. When I hear the phrase "toy" model, especially in the context of forecasing, I think of very low-degree-of-freedom models, often just ordinary differential equations, such as the Lorenz 63 system.

Your model is a model of fluid dynamics. I would not refer to it as a toy.

Because of the word "toy" in the Abstract and Introduction, I did not even realize that you were using the equations of two-dimensional fluid dynamics, until I saw the equations themselves in Section 2.

I would recommend removing the word "toy" from the manuscript, and describing the ABC model in a way that makes it clear that it involves two-dimensional fluid dynamics.
We have made the amendments accordingly, omitting “toy”.