Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-232-RC2, 2017

Geoscientific Model Development Discussions

Interactive comment on "A 4D-Var inversion system based on the icosahedral grid model (NICAM-TM 4D-Var v1.0): 2. Optimization scheme and identical twin experiment of atmospheric CO_2 inversion" *by* Yosuke Niwa et al.

Anonymous Referee #2

Received and published: 14 February 2017

This paper describes the application of the POpULar minimization scheme within the 4dVAR analysis of CO2 fluxes. The topic is worthy of publication as a number of modelling issues are raised, however the use of an identical twin experiment with synthetic observations has its limitations. While such an approach is appropriate for testing a system, care needs to be taken regarding the strength of the conclusions. There are a number of corrections and clarifications that should be dealt with prior to final acceptance. Some revision is warranted prior to final acceptance.

Scientific Significance: Good – applies techniques from one area (NWP, oceanography,

etc.) to another (CO2 inversion). New area of application provides sufficient differences to make the problem of wide interest. Also explores two different ways of designing adjoints.

Scientific Quality: Fair to Good – the overall approach and methods are valid. However there is a need for stronger caveats around some of the conclusions.

Scientific Reproducibility: Good - Full description of algorithm and data sets is provided. Detail s of models are in other papers.

Presentation Quality: Fair – There are a number of typos etc. and providing displatys of differences is preferable to eye-ball comparisons between full fields.

Specific issues:

1. A number of choices were made in the construction of the identical twin experiment that could unfairly favour one of the configurations over the other systems tested. These choices include:

- Using the same initial concentrations in the true and experimental run

- Having the true and prior land fluxes represent emissions with the same sources The effect of these choices on the conclusions should be assessed.

2. Initial concentrations are held fixed, and only fluxes are adjusted. This choice needs to be justified, especially if this system is to be cycled in an assimilation-forecast mode.

3. The background error variances are derived using the same truth as the experiment, Eq. 24, but in general the variances would also contain error, and probably significant errors since the method as described has no method of updating these.

4. To better understand Fig 3 (Global RMSEplots vs. iteration) there should also be lines showing, cost function and its gradient. This is important for a number of reasons:

- All iterative schemes are run for 60 iterations – the importance of this is not estab-

lished. Is the iterative process for each configuration close to convergence.

- Each configuration will have a different cost function value, so it is important to see how the different configurations compare in terms of cost functions.

- It would also be useful to check the standard Jmin diagnostic, given that the covariances are reasonably well known (by construction in an identical twin experiment)

Figures: When comparing fields, it is much more instructive to compare show differences rather than full fields. Increments/differences are shown in Figure 4, but in 5, 6 & 7 only the full fields are shown, so it is quite possible to miss some important differences.

Figures 8 & 9: Not sure why the lines need to be so thick. I would find it clearer if thinner lines were used – and easier to actually read values.

Minor corrections:

Page 1, line 11: "...difficult decomposition of a matrix...". The standard decompositions used in variational assimilation are generally well-established not especially "difficult" in comparison with the development of other models described in this paper. The constraints on matrix structure imposed by the decompositions can be a problem – and so the key point is in the following sentence regarding the freedom to use a more general error covariance structure.

Unfortunately the paper uses a simple Gaussian correlation function, with one length scale for land and another for ocean, and no cross-correlation so the advantage is not investigated

- My underlying concern is that having the freedom to more generally specify error covariances is still some way from constructing a more general matrix and showing that it adds value – both in terms of accuracy and computational cost. There are a number of further potential problems that are not explored, and these may interact with the comparisons of the two adjoints explored in this paper

Page 2, line 17: While 4D-Var was implemented in operational NWP earlier than most other geophysical models, the theoretical basis goes back further, and so to say the 4D-Var was originally developed for NWP is incorrect. There were also 4D-Var ocean systems developed at a similar time.

Page2, line 22: "..and need to be resolved" is redundant

Page 3, line 2 & 3: It should be pointed out that nearly all global NWP systems also employ some level of ensemble covariances. These are considered important to:

- Capture some variation in time and space of background error variances and correlation

- Provide analysis error estimates.

Also while the error covariances are diagonal, these are in transformed variable space, which allows for some non-linear balances. The error covariance matrix of untransformed (physical) variables has off-diagonal elements. These latter variables are more akin to the fluxes being examined here.

Page 3, line 5: There are problems with using a spectral decomposition for surface fluxes, but not allowing variables to be mutually correlated when they are close by is not one of them. Indeed the error correlation used here can be handled spectrally. The points that follow are valid, however the construction of useful, general correlations can add computational cost that is not explored here – and the effect of non-linear vs linear calculations could potentially be different.

Page 3, line 27: Guo and Sandu found that the results varied between tests using idealized data and real data. This should be a flag that care must be taken with interpreting results from the identical twin experiments conducted here.

Page 3, line 31. It is not obvious to me where there was significant investigation into how the observation network could be better exploited – apart from the suggestion that the current network can provide sub-continental scale information, which seems a

rather obvious comment to me.

Page 4, line 11: "...x represents the increment..."

Page 11, line 19: typo: quadratic

Page 13, line 2: "Upon closer inspection..." would be better

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-232, 2016.