Geosci. Model Dev. Discuss., 6, C1696–C1703, 2013
www.geosci-model-dev-discuss.net/6/C1696/2013/

Geoscientific
Model Development
Discussions

Interactive comment on "divand-1.0: n-dimensional variational data analysis for ocean observations" *by* A. Barth et al.

A. Barth et al.

a.barth@ulg.ac.be

Received and published: 15 October 2013

Reply to Referee #1

The reviewer provided critical remarks to the manuscript but they were constructive and helped us to improved the manuscript. We believe that the revised manuscript does a much better job in integrating our work with the data assimilation literature (see in particular our reply to questions 1 and 2). The reviewer also questions the robustness of the application in the results section. Based on the specific question below it seems that the reviewer assumes that the conjugate gradient algorithm (without preconditioning) was used to generate the results which is not the case as we explain our answer to question 5.

There seems to be also a confusion about the role of the model climatology (also called reference climatology in the manuscript). We hope that our answer to question 6 can clarify the setup of our tests which is similar to twin experiments in data assimilation. The analysis technique is applied to "pseudo-observations" coming from a numerical model. The quality of the analysis is assessed by comparing the analysis with the model climatology (which is the true state in this setup).

Comments of the referee appear in bold while our reply is in normal typeface. The full citations of all references used in the repy to the reviewers can be found in the revised manuscript.

Reply to the comments:

1. A good part of the paper is dedicated to the method of covariance modelling, with the only reference being a PhD thesis from 1996. There is no indication as to how this relates to other covariance modelling commonly used in data assimilation e.g. see references below.

The references of the reviewer (and related articles, in particular the review article by Weaver and Mirouze, QJRMS, 2013) were very useful to better place our work in the context of other covariance modelling techniques. We thank the reviewer for pointing us in this direction. The manuscript was improved in particular by showing the following points:

- Clarify that our approach is based on norm splines
- Other techniques such as based on the diffusion operator (solved explicitly or implicitly) and methods using the recursive techniques are introduced.

- Clarify that the correlation function belong to the Matérn family of correlation which are also obtained by an implicit diffusion operator.
- We place the use of coordinate stretching to implement anisotropic covariances in the context of previous studies
- At multiple places in the manuscript we emphasis the role of preconditioning which is also related to covariance modelling. See also our answer to question 5.

2. The algorithms presented in section 3 are very common and there is nothing particularly new in their implementation for this system. There is little reference to the large body of work that already uses these methods.

The presented software implements several well known algorithm which are commonly used in meteorology and oceanography. However there are also several novel aspects of the software package. The main technical advance is that the package is able to compute the analysis in an n-dimensional space where n can be arbitrarily high (thus no separate code for 2, 3 or 4 dimensional analyses). It also provide a uniform framework where different solution strategies can be implemented and tested.

From a theoretical point of view, we were also able to compute the analytical kernel of the norm spline approach for n dimensions for the class of scalar products described in the manuscript. The revised manuscript makes a connection to the implicit diffusion approach and covariance based on the second- and third order autoregressive models. We also showed a link between the dimensionality and the highest derivatives that need to be used. For instance we showed that when the highest derivative in the cost-function is a laplacian, the analysis can a have a discontinuous derivative in 3D. To our knowledge these are original result.

The mentioned references are indeed relevant and have been incorporated along with

others in manuscript. Our approach is compared to other methods found in the literature.

The section describing the primal formulation has been expanded by a discussion about the preconditioning with relevant citations to the literature. The paragraph and a new experiment was also added to address the 5 comments of the reviewer.

3. The algorithms as presented assume a linear observation operator and no there is discussion of nonlinearity (maybe not needed if only interpolating?).

The observation operator must be linear for the divand tool. For the generation of climatologies based on in situ measurements, the observation operator is indeed only an interpolation operator as pointed out by the reviewer. The manuscript is updated the make this clear.

4. Eq. (16) - I believe that $\mathbf{y}^o$ should be $\mathbf{y}^o - \mathbf{H}\mathbf{x}^b$ in the minimum of the cost function.

Indeed. In earlier version of the manuscript we can set $\mathbf{x}^b$ to zero (which can be done without loss of generality). However we changed the manuscript before submission to make apear $\mathbf{x}^b$ explicitly. Unfortunately we forgot to change equation 16.

5. The methods presented for solving the problem in section 3 are likely to be highly ill-conditioned. There is no discussion of this or as to whether any preconditioning is necessary.

In section 3 two classes of solver are discussed: the direct solver (section 3.1.1 and 3.1.2) and iterative solver (section 3.1.3, 3.2 and 3.2.1). For the direct solver based on a sparse Cholesky factorization no preconditioning is necessary. This has been clarified in the introduction.

The results (section 6) were obtained using the direct solver based on the CHOLMOD library in the SparseSuite package (Chen et al., 2008; Davis and Hager, 2009). To verify the accuracy of the solution we computed the residual $\mathbf{r}$:

$$\mathbf{r} = \mathbf{P}_{\mathsf{inv}}(\mathbf{x}_a - \mathbf{x}_b) - \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}\mathbf{x}_b)$$

This residual is directly related to the gradient of the cost function and should be zero. The maximum residual for the 2D case is 5.457e-12 (units are degrees C^{-1}) and 5.5593e-10 for the 3D case without advection constrain. Therefore we think that numerical resolution is valid and that our results are robust.

The reviewer is right that the condition number of the matrix $\mathbf{P}_{\mathsf{inv}}$ is very high (larger than 10^5). For such cases, the iterative method presented in section 3 would be indeed impractical without preconditioning (but this was not used to generate the results). Equation 33 of the original manuscript (in section 3.2) shows which preconditioner is implemented in the dual formulation which is based on a square root factorization of the covariance matrix of the innovation vector ($\mathbf{H}^T\mathbf{P}\mathbf{H} + \mathbf{R}$) taking only the diagonal elements of $\mathbf{R}$ into account.

Section 3.1.3 of the original manuscript did not mentioned preconditioning. This was changed in the revised manuscript. We implemented the commonly used preconditioning based on the square root factors of the background covariance matrix (e.g. Lorenc, 1997; Haben et al., 2011) and a modified incomplete Cholesky factorization of the analysed error covariance matrix (Jones and Plassmann, 1995). These preconditioners are tested in an idealised configuration (but with a data density and observation error covariance similar to the ones found in realistic applications). Figure 3 of the revised manuscripts shows the number of iteration needed to reach a target accuracy (here 1e-6). In these experiment, the preconditioner based on the modified incomplete Cholesky factorization needed the fewest iterations for all tested domain sizes.

In the revised tool, the software has also be extended so that the user can provide a

custom preconditioner based on the background error covariance $\mathbf{B}$, observation error covariance $\mathbf{R}$ and observation operator $\mathbf{H}$.

6. The method as applied in the results section requires a climatological field in order to estimate the parameters in the analysis. This means that the procedure is linked to a particular model climatology. In practical data assimilation such parameters are normally based on a physical understanding of the system and it is not clear why a simulataneous optimization of the parameters as described here should lead to anything that is physically realistic. (Note this also explains my score on scientific reproducibility, though I think that is less important than the main point here).

The setup of the result section is actually equivalent to a twin experiment in data assimilation where we know the “true solution” and we want to recover it based on (often incomplete) pseudo-observations extracted from the true solution. In our case, the monthly model climatology is the “true solution” and pseudo-observations are extracted from the daily model output. In data assimilation it is common to determine optimal parameters of the analysis by using such twin experiments (e.g. Nerger et al, 2012) because the experiment is setup in a way, where one knows the “true“ solution. This procedure has also been used before in the context of data analysis by e.g. Guinehut et al. (2002) and Buongiorno Nardelli (2012).

For an analysis procedure with real observations, analysis parameters are often determined with cross-validation (e.g. Wahba and Wendelberger, 1980, Brankart and Brasseur, 1998, Troupin et al, 2010). The general principle is that some observations are left out from the analysis (the cross-validation data set) and the parameters are determined by minimizing the difference from the cross-validation data set. The physical understanding of the system can often give us an order of magnitude of the parameters involved. However, choosing the precise (or optimal) value is very difficult based on the physical understanding alone.

If we would have chosen fixed analysis parameters and observed a difference between the different analysis techniques, we could not be sure if the difference is inherent to the technique or is due to a suboptimal choice of the parameters for a given approach. In the present manuscript we compare the best possible parameters for every approach to each other.

It is true that in some cases (for example very poor data coverage or correlated observation errors which are not taken into account), the minimization of some error estimation can result in unphysical results, but this was not the case for the experiment where the optimal correlation length is about 1200 km (in space) and 5 months (in time). Such orders of magnitude could be expected from the large-scale features present in the ocean model.

7. Eq. (27) - The second term is already included in $\mathbf{P}_{\text{inv}}$ and so appears twice.

The reviewer is right. $\mathbf{P}_{\text{inv}}$ should be $\mathbf{B}_{\text{inv}}$ (which does not already include the second term).

8. In section 6.2 it is mentioned that the parameter a_s is calibrated according to the significance of the advection, but there is no explanation as to how this is done in practice.

In the present case, the parameter a_s is determined in the same way as the other parameters, i.e. it is optimized by minimizing the RMS difference between the analysis and the model climatology (i.e. the “true” climatology). It was mentioned in section 6.1 (last paragraph) and it is now repeated in section 6.2 of the revised manuscript.

9. In Fig 5 (top, left) it is not clear why the cross and circle do not overlap. It is discussed in the text, but if the panel is showing correlation with respect to the point of the cross, then by definition the correlation is equal to one there.

Fig 5 shows the covariance and not the correlation (see also the caption and corresponding description of the original manuscript). The point and the cross do not overlap because the background error variance is not spatially uniform. In the interior of the domain, every grid point is constrained by the four direct neighbors (or more if higher derivatives than the laplacian are used). This is not the case at the boundary of the domain and boundary points are thus less constrained. Therefore the background error variance is higher near the coast. Maybe the original version of the manuscript did not explain this clearly. The explanation in the manuscript has been expanded.

Interactive comment on Geosci. Model Dev. Discuss., 6, 4009, 2013.