Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-232-AC1, 2017
© Author(s) 2017. CC-BY 3.0 License.

Geoscientific Model Development Discussions
Open Access
EGU

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.

Yosuke Niwa et al.

yniwa@mri-jma.go.jp

Received and published: 28 March 2017

RC

This paper further develops and tests the inverse system introduced in part I of the paper. It contains interesting information, of which the POpULar system is particularly interesting. Unfortunately, the ability to provide complex correlation structures is not tested in the manuscript. Moreover, the study uses only artificial observations to test the system, which is OK for the current paper.

Printer-friendly version

Discussion paper

CC BY

AC

We really appreciate the reviewer, Prof. Dr. Maarten Krol, for taking his time to review our paper and giving us many valuable comments and suggestions. Described below are our replies to the reviewer's comments with page and line numbers of the attached supplementary manuscript, in which the changes from the original manuscript are colored in blue and red.

RC

I have a few major/minor remarks, which need to be addressed before the paper can move on the GMD. First, I have problems with the conclusion that the "non-exact" adjoint outperforms the linear model that uses the exact adjoint to calculate the derivative of the cost function. I think the result may be driven by the fact that the pseudo-observations have been produced by the model version with the flux-limiter activated. Thus it comes as no surprise that the non-exact flux-limited version of the model finds the solution more easy (the authors use a fixed number of iterations of 60). The problem with a non-exact adjoint appear when you want to estimate error-reductions, for which you to the best of my knowledge need an exact derivative (to approximate the Hessian). Anyhow, it would be interesting to include the views of the authors on the quantification of the error reduction and posterior co-variance calculation. As a note, I am not claiming that the non-exact adjoint produces erroneous results, but I am afraid the authors want to propagate the use of non-exact adjoint, which comes with disadvantages also. It would be instructive to repeat some experiments with pseudo-observations calculated with the flux-limiter turned off.

CC BY

As the reviewer expected, the pseudo-observations are produced by the model with the flux limiter. Nevertheless, the model with the flux limiter may improve the model accuracy and, in fact, we found the better performance in the comparison with real observations as shown in the previous accompanying paper (Niwa et al. 2017). However, such improvement seems small. Furthermore, as the reviewer mentioned, the accuracy reduction induced by turning the flux limiter off could be compensated by smaller model time steps. Therefore, we have shifted the focus from the superiority of NONLINEAR to the comparable optimization performance, i.e., similar convergence speed, of NONLINEAR and LINEAR. Relevant modified/added sentences are as follows:

"It is found that a system of forward and adjoint models that has smaller model errors, but with nonlinearity has comparable optimization performance to that of another system that conserves linearity with an exact adjoint relationship." **[Page 1, Lines 7-10]**

"In fact, the flux limiter could improve the model accuracy due to its non-oscillatory property, but such improvement seems small in a CO_2 transport simulation (Niwa et al., 2017). Also, the error induced by switching off the flux limiter could be compensated by smaller time steps, though it increases the computational cost." **[Page 6, Lines 10-13]**

"The forward simulation to construct the pseudo observations is performed with the online NICAM-TM with the flux limiter of the advection scheme turned on." **[Page 8, Lines 23-24]**

"This is due to the smaller transport error of the continuous adjoint and the flexibility of the POpULar optimization method against the model nonlinearity. Even with LINEAR, POpULar shows high optimization capability, though its error is slightly larger than when used with NONLINEAR." **[Page 17, Lines 11-14]**

Furthermore, as the reviewer suggested, we have repeated the experiment with pseudo-observations calculated with the flux limiter turned off. The inserted discus-

CC BY

sion is as follows.

“This is because the observations are constructed by the model with the flux limiter, which is more compatible with the NONLINEAR model set; that is, NONLINEAR has smaller model errors than LINEAR. In fact, when the observations are constructed without the flux limiter, the GRMSE value from the LINEAR case becomes smaller and compa- rable to that of NONLINEAR (not shown). In this case, the GRMSE value from NONLINEAR does not change so much from the control case, suggesting that NONLINEAR is relatively insensitive to its model error in the optimization.” **[Page 14, Lines 1-5]**

We have also added some prospects about the posterior error estimate in Conclusions as

“For instance, the perfect adjoint relationship is desirable to approximate as accurately as possible the inverse Hessian that is defined as a symmetric matrix. In fact, an accurately approximated inverse Hessian can be considered as the posterior error co- variance and can be applied to estimate error reductions and quantify observational impacts. A study into the application of the inverse Hessian matrix is in progress and will be reported elsewhere.” **[Page 17, Lines 17-20]**.

RC

The language needs refinement, preferably by a native speaker. Now, at many places, the optimal wording is not used to correctly phrase what the authors (in my opinion) want to say. I made many suggestions in the attached pdf. Especially sections 2.2.1 and 2.2.3 need to be clarified, since the strategy with respect to the biomass burning emissions remains unclear to me. Why is GFED left out of the true fluxes?

CC BY

AC

According to the reviewer's suggestions, we have modified the manuscript; especially, Sections 2.2.1 and 2.2.2 (previously 2.2.3) are thoroughly rewritten. Furthermore, we had a native speaker check again.

In fact, both the prior and true fluxes have biomass burning emissions but from different information sources: the prior biomass burning emission is from VISIT (as the part of NBP) and the true one is from GFED. These two biomass burning data are distinctly different because GFED is based on satellite fire spot data but VISIT uses only prognostic model-derived variables such as fuel load and soil moisture. **[Page 10, Lines 6-11]**

RC

What also should be clarified is the fact to the error setting of B is such that you assign larger errors at places where the flux sets differ. This leads to (artificial) good performance, because the system will adjust fluxes where differences are present, and will not adjust sources where no differences between the flux sets are present. I guess your statement: "Nevertheless, not including the biomass burning(s) in xtrue would help us to elucidate the ability of the inversion system to find large CO2 flux anomalies" gives a clue, but the description is simply messy, and needs improvement. To be more convincing, it would be good to repeat the inversion with a "proportional to emission" error, as commonly done in inversions, or using the "French" approach to set the error based on the emissions in neighbouring grid cells.

CC BY

Why we used the prior-true flux differences for the prior error is that we wanted to simply validate the optimization scheme as a first step and also to evaluate the adjoint effects. Because the adjoint effects might depend on the setting of the prior error, the "true" prior error could provide a sort of benchmark result. As the reviewer pointed out, this leads an overestimation of the inversion performance. We touched on that in the first manuscript **[Page 11, Lines 6-10]**, but we have added more in the revised manuscript to draw readers' attention **[Page 1, lines 6-7; Page 4, lines 13-15; Page 10, Lines 23-27; Page 11, lines 23-25; Page 18, lines 2-3]**.

Nevertheless, we also wanted to see the case when the prior flux errors are wrongly assigned. To this end, we did not use GFED to assign the flux errors. This results in that the general pattern of the prior flux error is "true", but partially not "true", where GFED's emission is dominant. Relevant descriptions we modified/added are as follows:

"Partially, we add errors into the prior flux errors by not including the biomass burning emission of GFED in $\Delta\mathbf{x}$ so that we can also see how the system works with wrongly defined prior errors. Therefore, $\Delta\mathbf{x}$ represents the flux differences between VISIT NBP and the modified CASA NEP for land, and between Iida et al. (2015) and Takahashi et al. (2009) for ocean." **[Page 10, lines 27-30]**

"Moreover, a remarkable performance of the new system is demonstrated by the result that the inversion is able to detect regionally limited flux anomalies caused by biomass burnings that are not represented in the prescribed prior errors." **[Page 18, lines 4-5]**

As the reviewer suggested, the "proportional to emission" error could be another way, but this strongly depends on which flux dataset we choose. In fact, we have already tried a prior error whose distribution is proportional to respiration or GPP component of the prior flux. Then, we found that the inversion performance significantly depends on the flux dataset, which makes it difficult to evaluate the adjoint effects. Therefore, we

CC BY

leave how to design the prior error covariance (not only variances but also covariances) for a future study.

RC

technical, use "souce inversion" of "estimating sourcecs and sinks".

AC

We have modified the text to “source/sink inversions of atmospheric constituents” **[Page 1, line 1]**

RC

this depends on the measurement set-up, which you have not explained in the abstract.

AC

According to the reviewer’s comment, we have added “when 65 weekly flask sampling observations at ground-based stations are used” at the end of the sentence. **[Page 1, lines 11-12]**

CC BY

RC

first explain a bit better....how synthesis inversion works...what this matrix is, etc...

AC

We have added the sentence as follows.

"The synthesis inversion method takes an approach based on the Green's function matrix, in which sub-continental scale regions are prescribed and the total flux for each region is set as a parameter." **[Page 2, lines 17-19]**

RC

I think the argument is that you have to limit the degrees of freedom, something you accomplish by introducing error correlations among the grid-cells

AC

According to the comment, we have added the sentence as follows.

"Furthermore, introducing the flux error correlations reduces the degree of freedom and this may provide positive influences, e.g. reduction of noises, in a flux estimation, especially when observational networks are sparse." **[Page 3, lines 20-22]**

CC BY

RC

? why would that be? Difficult Eigenvalue problem?

AC

We wanted to say that a prior error covariance matrix with a higher dimension would require a much larger computational cost in an eigenvalue decomposition. To be more specific, we modified as “However, eigenvalue decomposition that was used in Chevallier et al. (2007) or Meirink et al. (2008) would become difficult when the specified prior error covariance is complicated or time-consuming when the spatiotemporal resolution of fluxes is increased.” **[Page 3, lines 22-24]**

RC

this statement is not true: this has not been proven if you refer to only one study.

AC

According to the comment, we modified the text to be more general as, “Therefore, which adjoint is better may depend on the assimilation settings.” **[Page 4, lines 9-10]**

CC BY

RC

unclear if the flux-limiter is used here (see main comment)

AC

We have added a sentence saying that the pseudo observations are constructed by the online NICAM-TM with the flux limiter. "The forward simulation to construct the pseudo observations is performed with the online NICAM-TM with the flux limiter of the advection scheme turned on. **[Page 8, lines 23-24]** Furthermore, to introduce the flux limiter beforehand, we have moved the sentences for the flux limiter (and also for the continuous and discrete adjoint modes) from Section 2.3 **[Page 12, lines 14-18]** to Section 2.1.2 **[Page 6 lines 6-lin 10]**.

RC

not entirely true....I mean you also use a background term in the cont function.

AC

Accordingly, we modified as ". . .the prior fluxes, which are expected to get closer to the true fluxes." and deleted the following sentence "The validity of the. . ..". **[Page 8, lines 21-22]**

CC BY

RC

at this stage it is good to mention that you also perturb the observations...!

AC

According to the reviewer's suggestion, we have moved the whole paragraph in the previous Section 2.2.2 (Pseudo observations) to the second paragraph of Section 2.2. **[Page 8, lines 24-29; Page 10, lines 14-21]**

RC

This is tricky: in fact, these numerical artifacts could be avoided by smaller time steps etc.

AC

According to the reviewer's comment, we have added the following sentences into the third paragraph of Section 2.1.2, which was previously located in 2.3. "In fact, the flux limiter could improve the model accuracy due to its non-oscillatory property, but such improvement seems small in a CO_2 transport simulation (Niwa et al., 2017). Also, the error induced by switching off the flux limiter could be compensated by smaller time steps, though it increases the computational cost." **[Page 6, lines 10-13]**

CC BY

RC

chi2 ??? overfitting?

AC

We have added discussion about the chi-square test from **Page 14, line 34** to **Page 15, line 10**.

RC

well, this sounds not really convincing. See my main comment (which I wrote before reading and understanding this part.).

AC

We have modified the text as “These results suggest that, under the assumption that the prior fluxes are well known, our new inversion system is capable of reproducing continental flux patterns by using only the surface observations.” **[Page 16, lines 3-5]**

RC

non-optimized case: what about more iterations?

CC BY

AC

These values do not significantly change when more iterations are performed. We can see that this experiment reaches almost close to the convergence after 60 iterations in Figure 3, which is newly added to show the cost function.

RC

what about regions with zero prior? error scales with emissions??? e.g. AUstralia... OR IF THE ERROR IS "MISSING BB ...LOGICAL RESULT

AC

We did not scale the errors with the prior emissions, but used the prior-true flux differences as described in the text. In a future inversion study with real observations, we will use some prior flux information to construct prior errors (e.g., proportional to a prior flux), including biomass burning.

Other trivial changes

Because the words “off-line”, “on-line”, “non-linear” are changed to “offline”, “online”, “nonlinear”, respectively, in the previous accompanying paper (Niwa et al., 2017) through the publication process, we have changed the text of this manuscript as well.

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

CC BY