Geosci. Model Dev. Discuss., 5, C941–C950, 2012
www.geosci-model-dev-discuss.net/5/C941/2012/

Geoscientific Model Development Discussions

GMDD

5, C941–C950, 2012

Interactive Comment

Interactive comment on "Quantifying the model structural error in carbon cycle data assimilation systems" *by* S. Kuppel et al.

S. Kuppel et al.

sylvain.kuppel@lsce.ipsl.fr

Received and published: 21 November 2012

The authors would like to thank the Referee for the very helpful comments and suggestions. The comments have been taken into consideration in the revised manuscript. We answer all of them individually in the following.

1 Summary of error characteristics

A general summary (e.g. as a table) of the various error characteristics could help to get a better overview of the results derived in this work (e.g from P2267 L25 – 27; P2268 L14; P2269 L19 – 20). Additionally the authors also could extend

their descriptions of the prior-residuals, prior-parameter errors and observation errors beyond the reporting of median values and they could also describe the spread of these errors derived from the different Fluxnet sites.

In the revised manuscript, was have added Table 1 (Table 2 in the revised manuscript), and modified the end of the first paragraph in Sect. 4:

'The statistics of the prior residuals (i.e., measurements-minus-simulations) and the prior-parameter error allowed us to estimate the structure of the observation error (i.e., model error + measurement error) whose inferred characteristics in different observation are summarized in Table 2, and to subsequently derive the model error based on earlier works regarding the flux measurement error.'

Table 1. Summary of the characteristics of the median observation error (measurement error + model error) in the ORCHIDEE model, projected in several observation spaces.

Observation type	**Structure of the observation error**		
	Standard deviation	Time correlation	Space correlation
Surface carbon flux (NEE)	1.7 $gC.m^{-2}.d^{-1}$	Rapid decrease, below 0.4 after the first day	Exponential decrease, e-folding length of 500km
Atmospheric CO_2 concentration (surface sample)	1.3 ppm	Rapid decrease, below 0.4 after the second day	No specific structure
Atmospheric CO_2 concentration (total column)	0.5 ppm	Rapid decrease, below 0.4 after the second day	Exponential decrease, e-folding length of 1200 km

Note that in Table 1, we have chosen to remain focused on the observation error because it is the main focus of this study. Also, we consider that describing the spread between sites is interesting but would substantially lengthen the text and overload the figures. Meanwhile, to our sense it would not serve significantly the analysis in that our main focus is to calculate an 'averaged' estimate of $\mathbf{R}$ to represent the mean characteristics of the observation error. Therefore, we find relevant to use only the median as a metric for this first study.

2 Limitation to DBF

The authors limit their study to deciduous broadleaf forests and also state this as a limitation. I suggest to further discuss this limitation and where possible the authors could make an attempt to give some quantitative arguments. Specifically this could include:

- **P2263 L6-8 and P2271 L7: How much does DBF really dominate Northern Hemisphere (in Orchidee) and what could this mean for the results presented here?**

  The sentence P2263 L6-8 is somewhat misleading as we did not imply that the Northern Hemisphere is dominated by DBF ecosystems, but rather that it is the dominant vegetation type at the measurements sites used in this study. In the revised manuscript, we have rephrased the sentence:

  'We selected 12 flux tower stations of the Northern Hemisphere located in temperate deciduous broadleaf forests, which correspond to one of the plant functional types (PFT) used in the ORCHIDEE model (Table A1).'

  In ORCHIDEE, the DBF ecosystem is not dominant in the Northern Hemisphere (let alone on continental Earth), so that the extrapolation is a rather strong hypothesis. However, we have adopted it because Krinner et al. (2005) and more

GMDD
5, C941–C950, 2012

recently Wang et al. (2012) have shown that the model-data misfit is not significantly peculiar for the DBF ecosystem as compared to the other plant functional types of ORCHIDEE. We have modified the corresponding sentence in the revised manuscript, at the beginning of section 3.5 :

'Assuming similar model-data mismatches across biomes with the ORCHIDEE model (see for example Wang et al., 2012), the characteristics of the observation error $\mathbf{R}$ *diagnosed in temperate deciduous broadleaf forests (e-folding lengths of 500 km and 1 day for space and time correlation, respectively) is prescribed at the global scale, and we project the inferred* $\hat{\mathbf{R}}^{prior}$ *in the space of atmospheric concentrations using the LMDZ transport model (see Sect. 2.5).'*

- **P2268 l16: The uncertainties from Hollinger and Richardson (2005) are derived for one site which is dominated by an evergreen needleleaf forest. The errors are described to depend on the magnitude of the fluxes themselves. This should be more prominently stated and the consequences of potentially higher measurement uncertainties on Rmod should be discussed.**

Richardson et al. (2008) give a more systematic description of the random measurement error, describing a mean behavior of its standard deviation using ten site-years, four of which in temperate deciduous broadleaf forests. We acknowledge that this error varies with the flux amplitude, ranging at the half-hourly timescale from 1 to 6 μmol.m^{-2}.s^{-1}. In the submitted manuscript, the reported value of 0.4 gC.m^{-2}.d^{-1} has been calculated by assuming a roughly independent error from one hour to another (supported by Lasslop et al., 2008), which from a half-hourly error in the order of 2 μmol.m^{-2}.s^{-1} gives 2/$\sqrt{24}\approx$0.4gC.m^{-2}.d^{-1} for daily averages (with here 1 μmol.m^{-2}.s$^{-1}\approx$ 1 gC.m^{-2}.d^{-1}).

In the revised manuscript we have taken into account the variation of the measurement error mentioned above, although restricting the range from 1 to 4

μmol.m^{-2}.s^{-1} because 4 μmol.m^{-2}.s^{-1} corresponds to a negative flux of around -15 μmol.m^{-2}.s^{-1}, never exceeded on daily average at the sites studied here.

Finally, it leads to an standard deviation ranging from 0.2 to 0.8 gC.m^{-2}.d^{-1} at the daily timescale. As recommended, the consequences on the model error have been better highlighted in the second paragraph of Sect. 3.2 in the revised manuscript :

'The median standard deviation of the observation error is estimated to be 1.7 gC.m^{-2}.d^{-1}. This number combines measurement and model contributions (Sect. 2.3). Using measurements across different types of forest ecosystems, Richardson et al. (2008) found that the random measurement errors range approximately from 0.2 to 0.8 gC.m^{-2}.d^{-1} depending on the flux magnitude, which means that the variance due to the measurement errors accounts for 1 to 25% of the total observation variance. Additionally, Lasslop et al. (2008) showed that no significant measurement error correlation remains at the daily time scale. From these elements, we conclude that the seasonal structure of the model error in ORCHIDEE is very similar to that of the observation error described above (the orange curve in Fig. 2), along with a standard deviation ranging from 1.3 to 1.6 gC.m^{-2}.d^{-1}.'

3 Further considerations

p2261 L1 : Scholze et al (2007) have done this as well.

We have added to reference to this study in the revised manuscript.

P2261 L10: What about uncertainties in the surface characteristics (eg.: pft assignment to grid-cells)

GMDD
5, C941–C950, 2012

This error is in our opinion part of the forcing error, which should indeed be extended beyond the mere meteorological forcing (the latter being here considered part of the model error). We have modified the text accordingly:

'Bayes' theorem provides a rigorous paradigm to build such CCDASs. Its application implies characterizing the uncertainties of each CCDAS component:

- *The measurement error,*
- *The model error, which stems from inappropriate equation forms or from missing processes in the carbon-cycle model structure,*
- *The error brought by the meteorological and vegetation forcing data, here considered as a part of the model error,*
- *The parameter error, arising from inadequate knowledge about a series of parameters.'*

P2265 L3-4: I assume that $\hat{\mathbf{R}}^{prior}$ is the estimation derived from eq. 1 and $\hat{\mathbf{R}}^{eval}$ from eq. 3. The authors should make clearer statements, what they exactly mean with those terms.

To make it clearer, we first of all added this sentence at the end of the first paragraph of section 2.3 containing Eq. (1) :

'Note that the estimation of $\mathbf{R}$ with this diagnosis will hereafter be noted $\hat{\mathbf{R}}^{prior}$'.

Also, the mentioned paragraph has been modified to link more explicitly each term to the corresponding equation:

'The prior and posterior diagnoses proposed in Eqns. (1) and (3) are respectively the starting points toward a prior estimation of the covariance matrix of observation errors (i.e., model errors + measurement errors) $\hat{\mathbf{R}}^{prior}$ and a posterior evaluation $\hat{\mathbf{R}}^{post}$.'

GMDD

5, C941–C950, 2012

P2271 L14-16 Any reason why the surface stations (flasks and continuous) do not show a correlation structure but the total column measurements do.

In the revised manuscript, we have added at the end of section 3.5 a suggestion about what might cause this difference of structure:

'We suggest that the total column smoothens out the surface-originated signals, which results in much smaller variances and, thus, much larger correlations.'

P2271 L21-21: To my understanding, the model structural error is not equivalent to the aggregation error, even though its misrepresentation might have similar consequences for the results of the data assimilation system.

The corresponding sentence in the manuscript was meant to point out the analogous nature of these two errors in terms of truncation, although indeed the truncated spaces are different. In the revised manuscript, we have modified the sentence in order to be clearer:

'This term is analogous to the aggregation error that has been rigorously described in atmospheric inversions (Bocquet et al., 2011; Kaminski et al., 2001; Thompson et al., 2011), in that it arises from truncating a given space of variables.'

P2272 L14-19: This statement is not fully clear to me. Maybe the authors could describe their intention with more details.

This sentence refers to the results presented P2270 L9-20 where we show that:

- using in the optimization a diagonal $\mathbf{R}$ with the standard deviation diagnosed in the second paragraph of section 3.2 ($\hat{\mathbf{R}}^{prior}$) provides good results based

GMDD

5, C941–C950, 2012

on the optimality criterion described in Eq. (6) (Table 1, second row),

- the same optimization yields a posterior estimation $\hat{\mathbf{R}}^{post}$ (Eq. 3) which, when used for a new optimization, improves the results according to the same optimality criterion (Table 1, third row). It suggest that repeating iteratively the posterior diagnosis after each optimization could converges to a more accurate estimation of $\mathbf{R}$.

In the conclusions, we have modified the sentence, which now reads:

'The same inversions also show that the diagnosed standard deviation of the observation error complies fairly well with a common optimality criterion used in data assimilation. We and additionally suggest that an iterative use, in successive inversions, of the standard deviation brought by the posterior diagnosis mentioned above could further improve the estimation of the observation uncertainty.'

4 References

- Bocquet, M., Wu, L., and Chevallier, F.: Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism, Quarterly Journal of the Royal Meteorological Society, 137, 1340-1356, Doi 10.1002/Qj.837, 2011.
- Hollinger, D. Y., and Richardson, A. D.: Uncertainty in eddy covariance measurements and its application to physiological models, Tree Physiol., 25, 873–885, 2005.
- Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, Journal of Geophysical Research-Atmospheres, 106, 4703-4715, 2001.

- Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, Artn Gb1015, Doi 10.1029/2003gb002199, 2005.
- Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences, 5, 1311-1324, 2008.
- Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., Reichstein, M., Stauch, V. J., Braswell, B. H., Churkina, G., Kruijt, B., and Hollinger, D. Y.: Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals, Agricultural and Forest Meteorology, 148, 38-50, DOI 10.1016/j.agrformet.2007.09.001, 2008.
- Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic carbon cycle data assimilation system simulations, J. Geophys. Res.- Atmos., 112, D17305, doi:10.1029/2007jd008642, 2007.
- Thompson, R. L., Gerbig, C., and Rodenbeck, C.: A Bayesian inversion estimate of N(2)O emissions for western and central Europe and the assessment of aggregation errors, Atmospheric Chemistry and Physics, 11, 3443-3458, DOI 10.5194/acp-11-3443-2011, 2011.
- Wang, T., Brender, P., Ciais, P., Piao, S., Mahecha, M. D., Chevallier, F., Reichstein, M., Ottlé, C., Maignan, F., Arain, A., Bohrer, G., Cescatti, A., Kiely, G., Law, B. E., Lutz, M., Montagnani, L., Moors, E., Osborne, B., Panferov, O., Papale, D., Vaccari, F. P.: State-dependent errors in a land surface model across biomes inferred from eddy covariance observations on multiple timescales, Ecological Modelling, 246, 11-25, doi: 10.1016/j.ecolmodel.2012.07.017, 2012.

GMDD
5, C941–C950, 2012

Interactive Comment

Interactive comment on Geosci. Model Dev. Discuss., 5, 2259, 2012.

GMDD
5, C941–C950, 2012

Interactive
Comment