14.2. Multivariate Data Assimilation of Carbon Cycle Using Local Ensemble Transform Kalman Filter

[1]Ji-Sun Kang, [1]Eugenia Kalnay, [2]Junjie Liu, [2]Inez Fung, and [1]Ning Zeng
[1]University of Maryland, College Park, Maryland
[2]University of California, Berkeley, California

1. Introduction

In order to project a future climate, it is necessary to take into account of the impact of greenhouse gases which are being emitted more and more due to the human activities. This study has focused on CO_2, one of major greenhouse gases. In addition to the anthropogenic emission, it is absorbed or released from the surface (land and ocean), depending on the biogeochemical conditions as well as its interaction and feedback with the climate. Estimating surface CO_2 fluxes is essential to project the global budget of CO_2 because natural variability of surface CO_2 fluxes can answer how much of CO_2 will remain in the atmosphere.

Thus, our purpose in this study is to estimate surface CO_2 fluxes as well as atmospheric CO_2 concentration with an advanced data assimilation technique.

2. SPEEDY-C and VEGAS with SLand

First, we modified SPEEDY model (Molteni, 2003) to simulate atmospheric CO_2 concentration. We added one prognostic variable of atmospheric CO_2 which has only advection and diffusion, in addition to the original prognostic variables of wind (U, V), temperature (T), humidity (q), and surface pressure (Ps). Then, the model reads the forcing of surface CO_2 fluxes and transports it by wind in the atmosphere. The model does not change surface CO_2 fluxes. This model will be referred to as "SPEEDY-C".

For the time-varying fluxes of surface CO_2, we coupled a terrestrial carbon model VEGAS (Zeng, 2005), which is coupled to the physical land surface model SLand (Zeng et al., 2000a), to SPEEDY-C. With this coupled atmosphere-vegetation-soil model, we can calculate the CO_2 fluxes over land with time, but the climatology of the SPEEDY-VEGAS-SLand is significantly different from SPEEDY-C. [1]

3. Three types of data assimilation: LETKF

Using LETKF (Hunt et al., 2007), we have tried three types of data assimilation: One is an **uncoupled (univariate) data assimilation** in which atmospheric CO_2 concentration and surface CO_2 flux are updated by CO_2 observations and not affected by other atmospheric variables. Another is an **one-way multivariate data assimilation** in which the atmospheric CO_2 concentration and surface CO_2 flux are updated by these two variables as well as the wind fields, while the wind field, in addition to other atmospheric variables such as specific humidity and temperature, is not affected by these two carbon-related variables. The other is a **multivariate data assimilation** so that all the dynamical variables are included in one vector.

Here, the dynamic variables of analysis are (U, V, T, q, Ps, CO_2, CF). We do not have any observation of CO_2 fluxes (CF). Thus, CF can be updated by only background error covariance.

4. Experimental design: OSSEs

4.1. Perfect model simulation

We used the same model, SPEEDY-C, for both of nature and forecast, so model error could be ignored. Furthermore, we included only fossil fuel emission as the carbon fluxes forcing which is constant with time and 6 PgC/yr (Andres et al., 1996). We tested three types of analysis introduced above.

The observations of the atmospheric variables are located at the rawinsonde distribution of which coverage is about 9% in horizontal, while those of atmospheric CO_2 are uniformly distributed at every other grid so that the coverage is about 25%. The observation error for each variable is as follow; 1m/s for U and V, 1K for T, 0.1g/kg for q, 1hPa for Ps, 1ppmv for atmospheric CO_2.

Analysis has been done every 6 hour; we have been used 20 ensemble members, and 8% of multiplicative inflation. All the results are the 2-month analysis. The initial condition of CF are generated by randomly choosing 20 CF from the nature run and adding small random perturbation so it does not use any a priori information at all.

4.2. Imperfect model simulation

SPEEDY-VEGAS-SLand was used for nature run. This model has a climatology rather different than the SPEEDY-C which we use as the forecast model so that we need to consider model error, and only the one-way multivariate data assimilation was

[1] *Corresponding author address, 4337 CSS, University of Maryland, College Park, MD 20742. E-mail: jskang@atmos.umd.edu*

examined for imperfect model simulation. We have the time-varying CF over land from the coupled model and the monthly prescribed CF over ocean given by Takahashi et al. (2002).

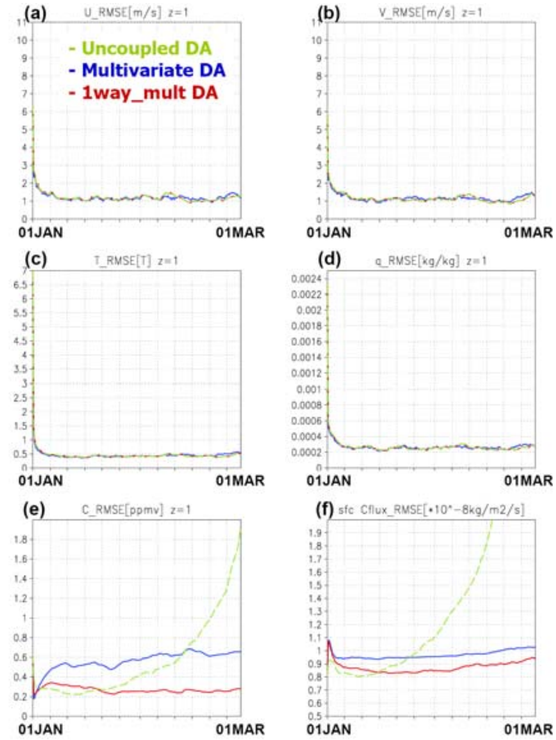

Figure 1. RMS error of analysis from three types of data assimilation: uncoupled(green),multivariate(blue), and one-way multivariate(red) data assimilation for (a) U, (b) V, (c) T, (d) q, (e) atmospheric CO_2 on the lowest layer of model, and (f) surface CO_2 fluxes

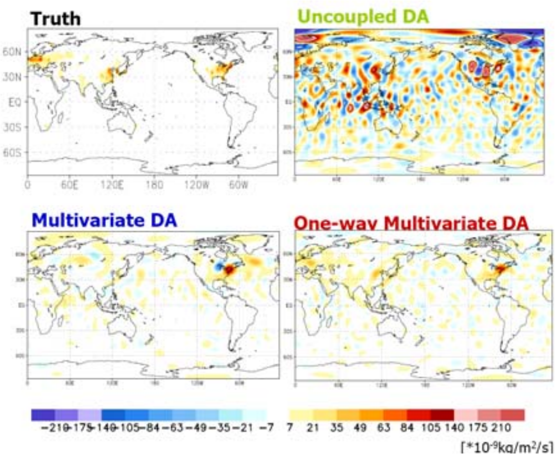

Figure 2. True state of surface CO_2 fluxes after two months of analysis, and the resultant analysis of it from three types of analysis.

10% of multiplicative inflation has been used except for the different inflation experiments (section 6.2). Rest of experimental settings is same as that of perfect model simulation.

5. Perfect model simulation

Both multivariate systems performed well in terms of RMS error (Fig. 1 and 2). The one-way multivariate assimilation resulted in the optimal performance for the CO_2 variables because it minimizes the sampling in the feedback from carbon variables to the atmospheric variables. By contrast, the univariate assimilation of carbon had larger errors and diverged. Using one-way multivariate data assimilation technique, we also tested the experiment which has the observations of CO_2 concentration only in the lowest layer and got the comparably good results (not shown). Moreover, experiment with daily observations of CO_2 still estimated CF reasonably well (not shown).

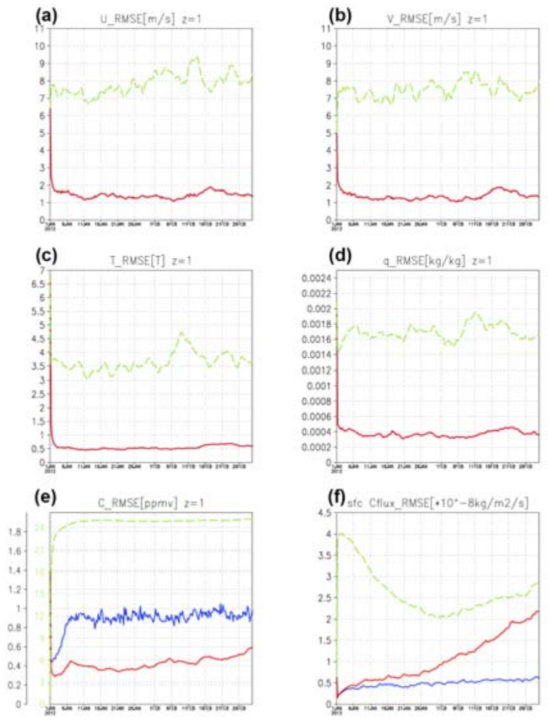

Figure 3. RMS error of analysis from imperfect model simulations: no bias correction (green), bias correction(red), and bias correction + different inflation(blue) for (a) U, (b) V, (c) T, (d) q, (e) atmospheric CO_2 on the lowest layer of model, and (f) surface CO_2 fluxes

6. Imperfect model simulation

Since we confirmed that one-way multivariate data assimilation has optimal performance through the perfect model simulation, we applied only this technique for imperfect model case. From the figure of RMS error (Fig.3: green), we found that the forecast was significantly different from the nature so that the ensemble system does not represent the true state well.

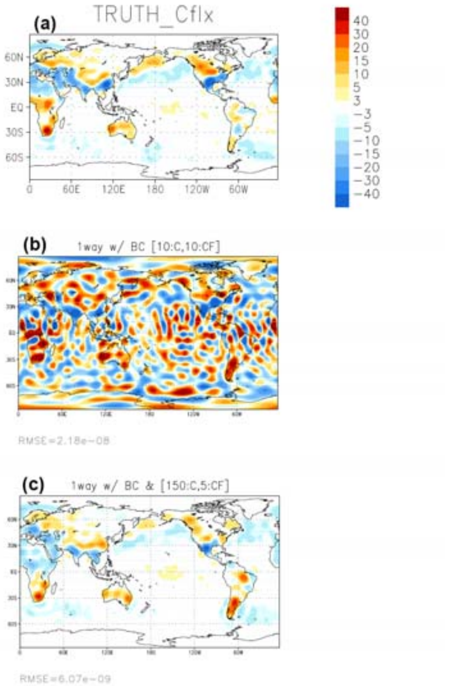

Figure 4. (a) True state of surface CO_2 fluxes after two months of analysis, and the resultant analysis of it from the experiment (b) with bias correction, (c) with bias correction + different inflation factors for CO_2 variables.

6.1. Model bias correction

We tried to estimate and fix the model bias we saw the previous experiment. Similar to Danforth et al. in 2007, we averaged the distance between the nature and the six-hour forecast which started from the true state for two months. Then, we subtracted this estimated model bias from the ensemble forecast before the analysis step.

With this simple method, we could get remarkable improvement in the analysis. Especially for the atmospheric variables and atmospheric CO_2 which have the observations, the current ensemble system worked very well. The analysis of CF, however, got diverged with time (Fig.3: red, Fig. 4(b) and 5(b)).

6.2. Inflation for atmospheric CO_2 and surface CO_2 fluxes

We found that the ensemble spread of CO_2 is not enough to estimate CF well so that we made a test run which has different inflation factor for CO_2 and CF. With large inflation (=1.50) for CO_2 and small inflation (=0.05) for CF, we obtained an excellent performance of CF (Fig.3: blue). Analysis of CO_2 got a little worse, but not bad if we consider 1ppmv of the observation error (Fig. 3, 4, and 5).

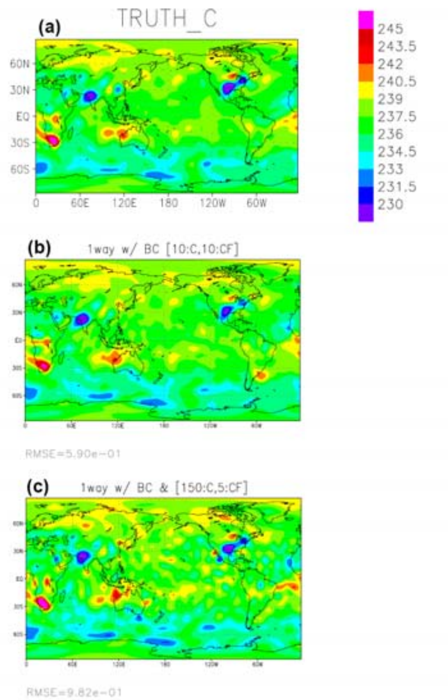

Figure 5. (a) True state of atmospheric CO_2 on the lowest layer after two months of analysis, and the resultant analysis of it from the experiment (b) with bias correction, (c) with bias correction + different inflation factors for CO_2 variables.

7. Summary and discussion

This study is a component of a project for CO_2 data assimilation with LETKF/CAM3.5 system (PIs: Eugenia Kalnay/Inez Fung). The CAM3.5 is very complex and expensive whereas the SPEEDY and the VEGAS are simple but realistic, having only an intermediate complexity. The coupled system of SPEEDY-VEGAS has produced reasonable results and LETKF has been also implemented successfully.

Through the perfect model simulation with three types of data assimilation, we can conclude that the multivariate CO_2 data assimilation experiments were performed for the first time, and the results indicate that multivariate EnKF assimilation is much more effective in estimating both atmospheric CO_2 and surface CO_2 fluxes, even in the absence of observations or prior estimations of surface fluxes.

Under the imperfect model assumption, we could estimated and remove the model bias and then get encouraging results and further improve the results with different inflation factor for atmospheric CO_2 and surface CO_2 fluxes. In order to find the optimal values for the inflation, we plan to calculate adaptive inflation and observation error (Li, 2008) for atmospheric variables as well as atmospheric CO_2.

References

Andres, R. J., et al. 1996: A 1° x 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990, *Global Biogeochem. Cycles*, **10,** 419-429.

Danforth, CM, E Kalnay and T Miyoshi, 2007: Estimating and correcting global weather model error. Mon. Wea. Rev. 135, 281-299. DI 10.1175/MWR3289.1|ISSN 0027-0644

Li, H. 2008: Local Ensemble Transform Kalman Filter with Realistic Observations, PhD Thesis at the University of Maryland.

Hunt, B. R., E. Kostelich, and I. Szunyogh 2007: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter, *Physica D*, **230**, 112-126.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. *Climate Dyn.*, **20**, 175-191.

Takahashi, T., S. C. Sutherland, C. Sweeney, A. Poisson, N. Metzl, B. Tilbrook, N. Bates, R. Wanninkhof, R. A. Feely, C. Sabine, J. Olafsson, Y. Nojiri, 2002: Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep-Sea Research II*, **49**, 1601-1622.

Zeng, N., A. Mariotti, and P. Wetzel, 2005: Terrestrial mechanisms of interannual CO_2 variability, *Global Biogeochemical Cycles*, **19**, GB1016, doi:10.1029/2004GB002273

Zeng, N., J. D. Neelin, and C. Chou, 2000a: A quasi-equilibrium tropical circulation model-Implementation and simulation, *J. Atmos. Sci*., **57**, 1767-1796