15A.3 CO_2 DATA ASSIMILATION WITH THE COUPLED ATMOSPHERE-VEGETATION MODEL USING LOCAL ENSEMBLE TRANSFORM KALMAN FILTER

Ji-Sun Kang, Eugenia Kalnay, Ning Zeng, and Junjie Liu
Department of Atmospheric and Oceanic Science, University of Maryland, College Park
Inez Fung – University of California, Berkeley

1. Introduction

Human activities result in increased carbon consumption and emission of CO_2 to the atmosphere and the impact on climate change due to global warming is serious. Understanding the carbon sources and sinks is important for the future of human life on the earth. Since the released CO_2 in the atmosphere must either go to the land, ocean, or just remain in the atmosphere, the capacity of each reservoir to absorb and store atmospheric CO_2 can give us invaluable information. In addition, the interaction and feedback between the behavior of these reservoirs and changing climate cannot be neglected so the problem is not simple.

In order to understand and predict atmospheric CO_2 in terms of climate change, it is necessary not only to consider the anthropogenic carbon use but also to understand the biogeochemical processes interacting with climate. For this purpose, the role of numerical modeling is essential. Thus, this study focuses on the development of coupled system of atmosphere-vegetation that allows simulating how these two systems interact with each other, and whether an advanced data assimilation scheme, LETKF, can be used to estimate simultaneously the standard atmospheric variables and CO_2 measurements in order to estimate surface fluxes of CO_2.

2. SPEEDY Model with Atmospheric CO_2

The SPEEDY model (Molteni, 2003) is an atmospheric general circulation model (AGCM) with simplified physical parameterization schemes that are computationally efficient, but that maintain the basic characteristics of a state-of-the-art AGCM with complex physics. It has triangular truncation T30 with 7 sigma levels.

The original version of SPEEDY has five dynamical variables including zonal (U) and meridional wind (V) components, temperature (T), specific humidity (q), and surface pressure (Ps). To use the model for this study, two variables have been added; one is an atmospheric CO_2 which has only two processes of advection and diffusion and the other is a surface flux of carbon (Cflux) which is a source and sink of the atmospheric CO_2. Chemical processes for CO_2 have been ignored.

To verify the realism of SPEEDY results, the mean atmospheric CO_2 concentration on the surface layer of our model was compared to that of the NCAR CCM [1] (Fung, 2006), a state-of-the-art model (Fig. 1). Although SPEEDY has only seven vertical layers and other simplifications, it has produced results reasonable enough to justify using it for this research. Since the horizontal and vertical resolution is much coarser in SPEEDY than NCAR CCM, the mixing of SPEEDY seems to be relatively strong compared to NCAR CCM.

3. VEGAS and SLand Models

The terrestrial carbon model VEgetation-Global-Atmosphere-Soil, VEGAS, (Zeng, 2005), which is coupled to the physical land surface model Simple-Land, SLand (Zeng et al., 2000a) is used in this research. The VEGAS simulates the dynamics of vegetation growth and competition among different plant functional types which include broadleaf tree, needleleaf tree, cold grass, and warm grass. The SLand models the first order effects relevant to climate simulation. Here, SLand is turned on if it is over the land; over oceans and ice we use the original method of the SPEEDY to calculate surface flux of heat and moisture.

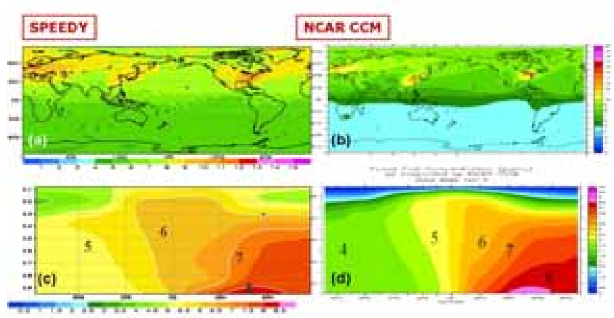

Figure 1 Annual mean for the third year of simulated atmospheric CO_2 concentration (ppmv) on the surface layer (a) by SPEEDY, (b) by NCAR CCM, and the vertical cross section of it (c) by SPEEDY, (d) by NCAR CCM

4. Coupling of SPEEDY and VEGAS

In order to account for the interaction of carbon between biosphere and atmosphere, VEGAS with SLand were coupled to SPEEDY (Fig. 2). At first, the spin-up run with only SLand-Vegas for 200 years was forced by 9-year mean of variables (heat and moisture on the surface) from SPEEDY. Then, the coupled spin-up was continued for 30 years so that the source and the sink of carbon over the land under the atmospheric environment of SPEEDY have been estimated (Fig. 3). Over ocean, the surface fluxes of carbon are given by the monthly prescribed fluxes of Takahashi et al. (2002).

5. Implementation of LETKF on SPEEDY Model with CO_2 and CFlux

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

LETKF (Hunt, 2005, Hunt et al., 2007) is an advanced ensemble Kalman filter data assimilation scheme. It is a square-root ensemble filter in which the observations are assimilated simultaneously to update the ensemble mean while the ensemble perturbations are updated by transforming the forecast perturbations through a transform matrix term as Bishop et al. (2001) introduced. The analysis is done independently at every grid point using observations from a local region.

The analysis in this study has seven variables, U, V, T, q, Ps, CO_2, Cflux and we use 20 ensemble members for the forecast run.

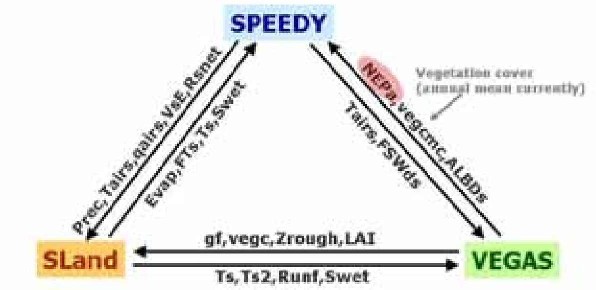

Figure 2 Concept of coupling: variables in the interfaces among three systems

6. Experimental Design

While the coupled system (SPEEDY-VEGAS with SLand) generates the truth or nature run, from which observations are derived, the SPEEDY with six prognostic variables including CO_2 is used as the forecast model in which the Cflux over land is updated only by the analysis, which means the forecast model does not have a dynamical forecast equation for the surface flux (the forecast of Cflux over land during the analysis cycle is persistence). Like the nature run, the forecast model also uses over ocean the monthly-prescribed carbon flux of Takahashi et al. (2002).

From the truth, the observations of variables were generated by adding Gaussian random perturbation. The standard deviation of those random perturbations depends on the scale of each variable; 1 m/s for U and V, 1 K for T, 0.1 g/kg for q, 1 hPa for Ps, 1 ppmv for CO_2, 10^{-9} kg/m^2/s for Cflux. The observations of the atmospheric variables are located at the distribution of rawinsonde of which coverage is about 5.7 % globally while those of CO_2 and Cflux are uniformly distributed at every other grid so that the coverage is about 25%. The initial condition for the ensemble forecast at 00Z01JAN2002 has been made by adding 20 random perturbations to the truth which were also at randomly chosen time.

In order to see how the background error covariance of the atmospheric variables effects on that of carbon related variables (CO_2 and Cflux), we designed two types of experiments: One is the **univariate** data assimilation in which CO_2 and Cflux are updated by these two and not affected by other atmospheric variables. The other is the **multivariate** data assimilation so that all the dynamical variables are included in one vector.

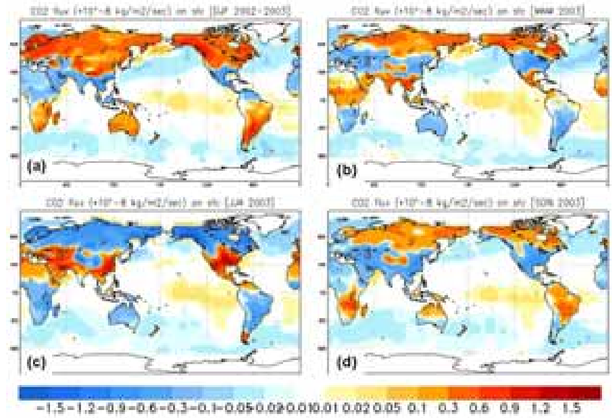

Figure 3 Seasonal pattern of the simulated surface CO_2 flux (kg/m2/s) by the coupled system; (a) DJF, (b) MAM, (c) JJA, and (d) SON.

7. Results

7.1 RMSE

In general, the multivariate data assimilation has less root mean square errors (RMSE) than the univariate one and the predominance is significant for CO_2 (Fig. 4). This means that the better information of the atmospheric variables can improve the CO_2 forecast either. Cflux in the multivariate case also has better result than that of the univariate one till the middle of February. However, we found that the RMSE of most variables are growing after one month. That is because our nature and forecast does not have the same physics for calculating the surface flux; the nature run has used SLand because of the coupling with VEGAS but the forecast model does use the intrinsic scheme of the SPEEDY instead of SLand. Thus, the difference between two runs gets too large to assume a perfect model experiment.

7.2 Model bias

In order to solve the problem due to the model bias shown in the previous section, we first calculated the model bias (Fig. 5) which is subtracting the forecast from the truth. The bias is especially apparent over the land.

8. 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. Thus, we have developed a tool good enough to do experiments testing methods for CO_2 data assimilation with a relatively inexpensive simulation system.

Preliminary results are encouraging, but in order to proceed, we plan to handle the model atmospheric biases that arise from the difference between the nature and the forecast model following the method that Li (2008) introduced, and use realistic observations of CO_2 and fluxes.

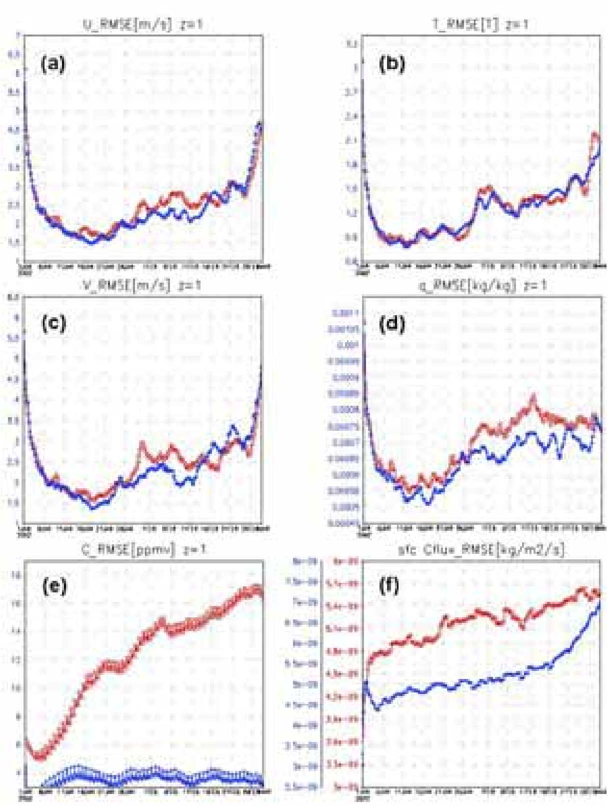

Figure 4 RMSE of (a) U, (b) T, (c) V, (d) q, (e) CO_2, (d) Cflux on the lowest layer of analysis, where red symbols indicate the results of the univariate data assimilation and blue the multivariate one.

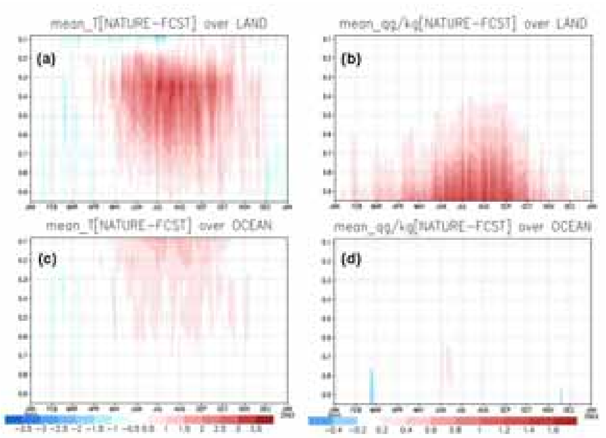

Figure 5 The model bias of (a) T, (b) q over the land, and (c), (d) same variables over the ocean.

References

Bishop, C. H., Etherton, B. J., and Majumdar, S.,2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.

Fung, I., 2006: Lecture Note on 'IMAGe/MSRI Summer Graduate Workshop on Data Assimilation for the Carbon Cycle'

Hunt, B. R., 2005: An efficient implementation of the Local Ensemble Kalman Filter, http://arxiv.org/abs/physics/0511236.

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.

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

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