

## Final Report

DOE Award DE- FG02-07ER64436

Title: Carbon Data Assimilation with a Coupled Ensemble Kalman Filter

PI: Eugenia Kalnay, UMD

Co-PI: Inez Fung, UCB

### **Abstract:**

We proposed (and accomplished) the development of an Ensemble Kalman Filter (EnKF) approach for the estimation of surface carbon fluxes as if they were parameters, augmenting the model with them. Our system is quite different from previous approaches, such as carbon flux inversions, 4D-Var, and EnKF with approximate background error covariance (Peters et al., 2008). We showed (using observing system simulation experiments, OSSEs) that these differences lead to a more accurate estimation of the evolving surface carbon fluxes at model grid-scale resolution.

The main properties of the LETKF-C are: a) The carbon cycle LETKF is coupled with the simultaneous assimilation of the standard atmospheric variables, so that the ensemble wind transport of the CO<sub>2</sub> provides an estimation of the carbon transport uncertainty. b) The use of an assimilation window (6hr) much shorter than the months-long windows used in other methods. This avoids the inevitable “blurring” of the signal that takes place in long windows due to turbulent mixing since the CO<sub>2</sub> does not have time to mix before the next window.

In this development we introduced new, advanced techniques that have since been adopted by the EnKF community (Kang, 2009, Kang et al., 2011, Kang et al. 2012). These advances include “variable localization” that reduces sampling errors in the estimation of the forecast error covariance, more advanced adaptive multiplicative and additive inflations, and vertical localization based on the time scale of the processes.

The main result has been obtained using the LETKF-C with all these advances, and assimilating simulated atmospheric CO<sub>2</sub> observations from different observing systems (surface flask observations of CO<sub>2</sub> but no surface carbon fluxes observations, total column CO<sub>2</sub> from GoSAT/OCO-2, and upper troposphere AIRS retrievals). After a spin-up of about one month, the LETKF-C succeeded in reconstructing the true evolving surface fluxes of carbon at a model grid resolution. When applied to the CAM3.5 model, the LETKF gave very promising results as well, although only one month is available.

# Final Report: Estimation of Surface CO2 Fluxes from Atmospheric Data Assimilation

For brevity, in this final report, we follow the presentation with that title that Dr. Ji-Sun Kang (UMD, now at KIAPS, Korea) gave at the 9<sup>th</sup> International Carbon Dioxide Conference in Beijing, June 3-7, 2013 (also attached). Her co-authors are Eugenia Kalnay PI, UMD), Junjie Liu (Co-I, JPL), Inez Fung (Co-PI, UCB) and Takemasa Miyoshi (UMD, now at AICS, RIKEN, Japan).

It is remarkable that after her talk, Dr. Wouter Peters, the principal architect of the operational NOAA CarbonTracker system (Peters et al., 2005) used to keep track of carbon dioxide uptake and release at the Earth's surface over time, stood up and spent several minutes explaining to the audience why he considers that our LETKF-C approach is superior to his own CarbonTracker, and that he plans to upgrade his system following our work.

## 1. Introduction: The LETKF-C

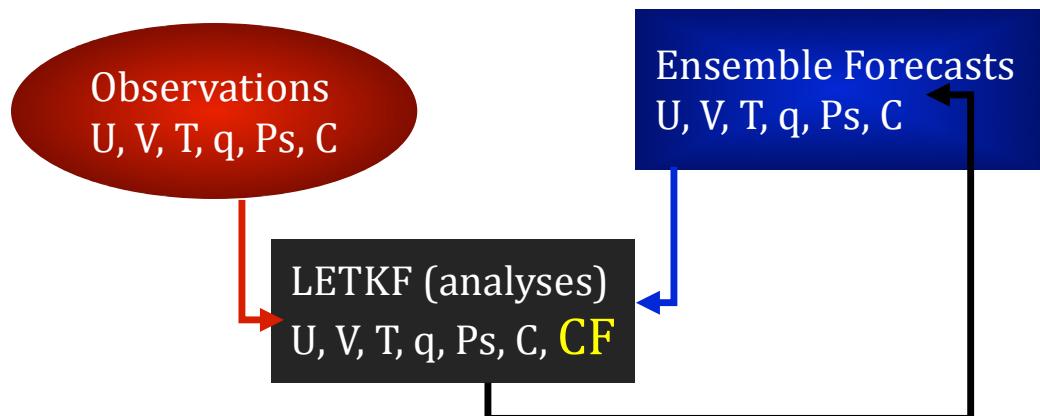


Fig. 1: This schematic shows in red the observations (in red) assimilated every 6 hours, including the atmospheric variables zonal wind U, meridional wind V, temperature T, surface pressure Ps and atmospheric CO2 (C). The forecast model has the same variables but in model space (in blue). From these two inputs, the Local Ensemble Transform Kalman Filter-Carbon (LETKF-C) generates an analysis for the model variables as well as for the augmented variable surface carbon fluxes (CF).

As shown in Figure 1, the LETKF-C consists of the atmospheric observations and the ensemble forecasts of the atmospheric variables and CO2, read by the LETKF. The carbon fluxes (CF) are appended to the model vector state, and the background error covariance estimated by the LETKF includes the terms associated with CF.

This allows the LETKF to estimate the changes of the CF as if they were “evolving parameters” (not measured).

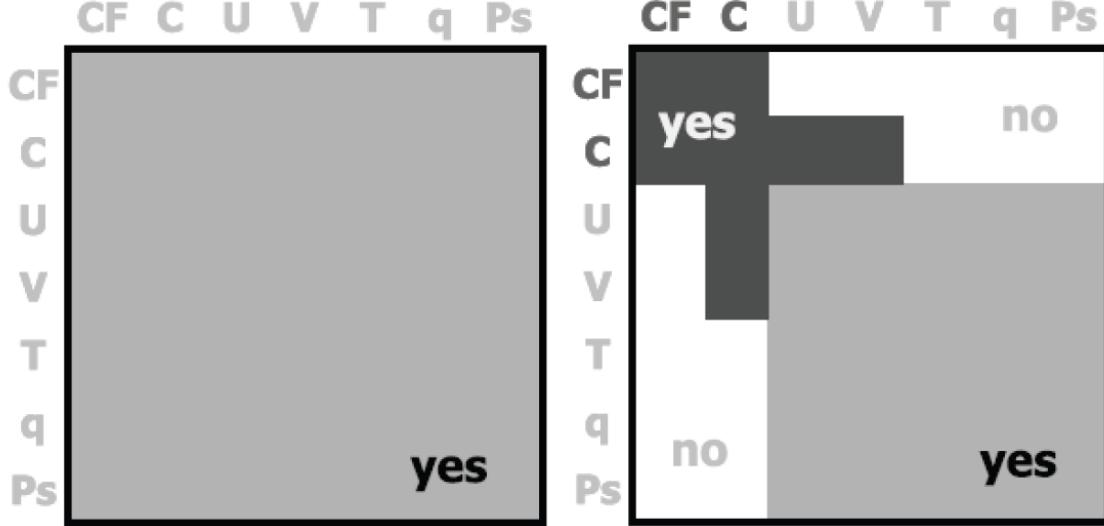


Fig. 2: Schematic of the background error covariance matrix  $P^b$ . Left: standard  $P^b$  without variable localization, in which error covariances among variables are computed from the ensemble forecast. Right: newly proposed  $P^b$  with variable localization only the error covariance between the wind and the atmospheric CO2 (C), and the covariance between the atmospheric CO2 and the surface fluxes are included. The covariance between carbon variables and other variables with which they are not physically connected is zeroed out, thus substantially reducing the sampling errors.

Fig. 2 shows how we reduce substantially sampling errors by zeroing out the error covariances between the carbon variables and and temperature, moisture and surface pressure, since these variables are not physically interacting. The standard approach, by contrast, is to compute the covariance from the ensemble forecast perturbations around the mean (as shown in the left panel). For a system with a limited number of ensemble members, the covariance will be small but not zero, and thus introduce sampling errors. We found that this approach, now accepted in the ensemble research community, improved significantly the accuracy of the estimated surface fluxes.

## 2. The model

The model used for these experiments is an adaptation of the SPEEDY model of Molteni (2003). The SPEEDY-C is a spectral atmospheric GCM with T30L7 resolution. To the original variables, Kang (2009) added atmospheric CO2 (C) as an inert tracer transported by the atmospheric wind. For the surface carbon fluxes (CF), no observations were used and the CF estimated by the analysis was persisted during the next forecast, without observations.

The observations simulated in these experiments are created from a “nature” (truth) long run, at positions and times close to the real observations, adding random errors with realistic standard deviations. The simulated observations include

- Rawinsonde observations of U, V, T, q, Ps.
- Ground-based observations of atmospheric CO<sub>2</sub>, including 18 hourly and 107 weekly data on the globe.
- Remote sensing of CO<sub>2</sub>:
  - AIRS retrievals, whose averaging kernel peaks at mid-upper troposphere.
  - GOSAT retrievals, whose averaging kernel is nearly uniform throughout the column.

The ensemble size used is 20 members, since using 30 members showed little improvement. The initial conditions are random so that, unlike other studies, we provide the system with no *a priori* information.

### **3. Results:**

We summarize the many results that we obtained with a single figure that shows that in an OSSE, using realistic observations of the current and soon to be launched observing systems, it is possible to estimate surface fluxes of carbon.

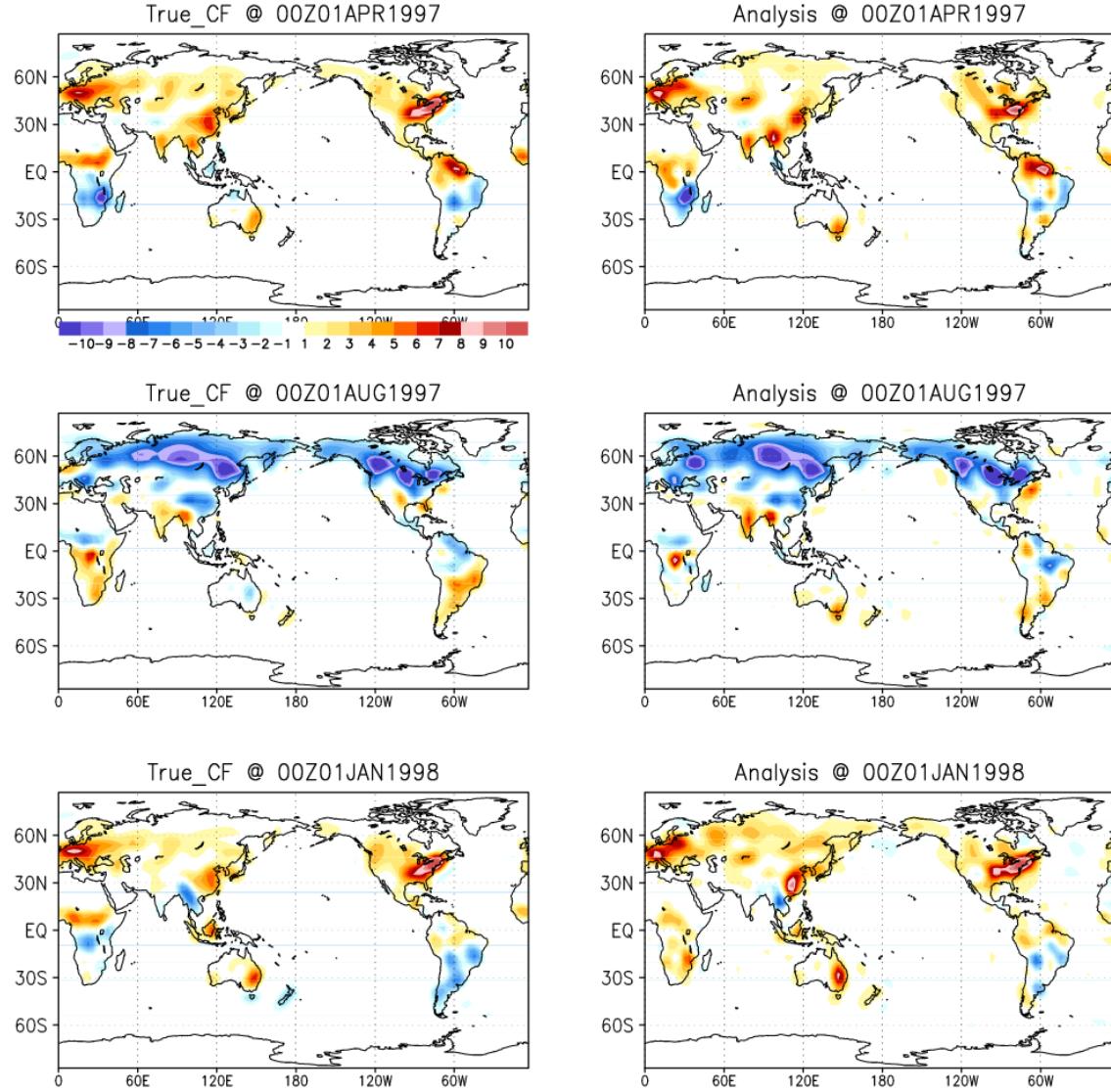


Fig. 3: Left: “True Carbon Fluxes” (True CF) from the nature run. Right: LETKF-C estimation of the CF after 3, 7 and 12 months of data assimilation. The initial conditions were random, so they had no *a priori* information.

These results are remarkable: they show that with the advanced data assimilation system LETKF-C (whose characteristics are further discussed below), a perfect model, and observations similar to the observations that will be available after the launch of OCO-2, it is indeed possible to estimate the surface carbon fluxes with fairly high accuracy and resolution equivalent to the model grid scale.

Note that the analysis is more accurate in the NH than in the SH. There are two reasons for this result: a) the atmospheric analysis was based only on the use of rawinsondes, much more abundant in the NH than in the SH; b) there are more ground based observations of CO<sub>2</sub>, especially in Europe and North America.

#### 4. Characteristics of the Carbon Cycle Data Assimilation within LETKF-C (Kang et al., JGR, 2011, 2012).

The main characteristics of our system that led to the success shown in Fig. 3 are:

- Simultaneous analysis of meteorological and carbon variables.
  - Advantages: The coupled ensemble provides an accurate estimate of the CO<sub>2</sub> uncertainty due to the atmospheric transport.
- “Localization of Variables”
  - Advantages: reduces sampling errors by zeroing out, rather than estimating from the ensemble covariances that should be zero.
- Advanced inflation methods.
  - Advantages: Adaptive multiplicative inflation (Miyoshi, 2011) was adapted to our system. For the carbon fluxes, without observations, multiplicative inflation would not work, so we included additive inflation that was equivalent to a return to the background error, (Whitaker et al., 2008), and which we found optimal when tuned.
- Vertical localization of column mixed CO<sub>2</sub> observations.
  - Since the time scales for changes in the atmospheric CO<sub>2</sub> are much shorter near the surface than in the upper troposphere, we created a vertical localization for the column average changes that was larger near the surface.
  - Advantages: the larger attribution of column total CO<sub>2</sub> changes to the layers near the surface resulted in a significant increase in the accuracy of the analysis of CF.
- Use of a short (6-hour) window. This is the most controversial characteristic of our system, so we discuss it in further detail.
  - Most of the CO<sub>2</sub> inversion groups have adopted much longer window lengths (weeks to months).
  - This started in the 1980’s when only 10’s or 100’s of ground based CO<sub>2</sub> observations were available on the globe.
  - It is also based on the idea that CO<sub>2</sub> is an inert gas, so that its memory of the surface carbon fluxes is very long.
  - But now we have satellite observations of CO<sub>2</sub> that we should consider (e.g., AIRS, GOSAT and soon OCO-2).
  - Furthermore, and very importantly, a long window can take advantage of the long memory of the impact of surface carbon fluxes on atmospheric CO<sub>2</sub> only if the model can keep track of the flow of CO<sub>2</sub> during the long window.
  - This is a very difficult problem with a long window, as pointed out by Enting (2002, fig. 1.2), who pointed out that inverting the CO<sub>2</sub> flow to obtain surface fluxes is an ill-posed problem due to turbulent mixing.
  - Fig. 4 is a schematic explaining why a short window has a much better chance to estimate the surface fluxes from the changes in atmospheric CO<sub>2</sub> than a long window.

- With a short window the analysis system can use the strong correlation between C and CF before the transport of C blurs out the essential information of CF forcing.
- We may not be able to take advantage of this correlation because this signal becomes blurred with long windows.
- We found that in our system extending the window to 3 weeks resulted in somewhat worse results, especially in regions that included neighboring sources and sinks (see Fig. 5).

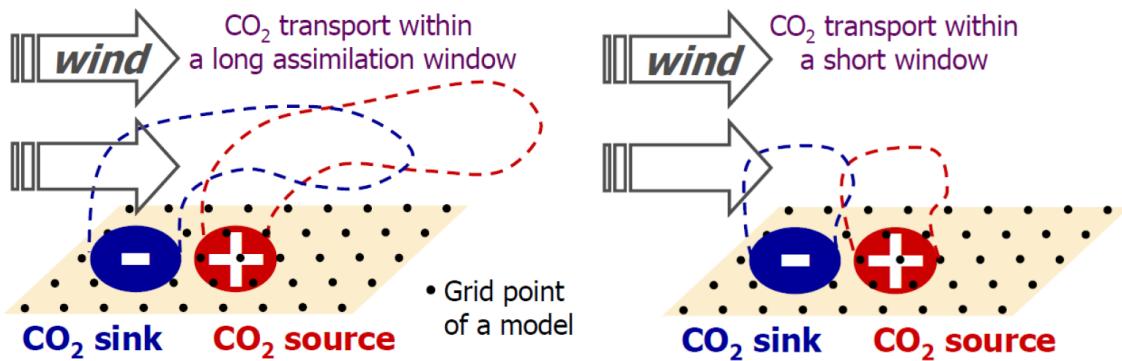


Fig. 4: Schematic showing the impact of a short or a long window on the attenuation of the observed CO<sub>2</sub> information. With long windows the information is blurred.

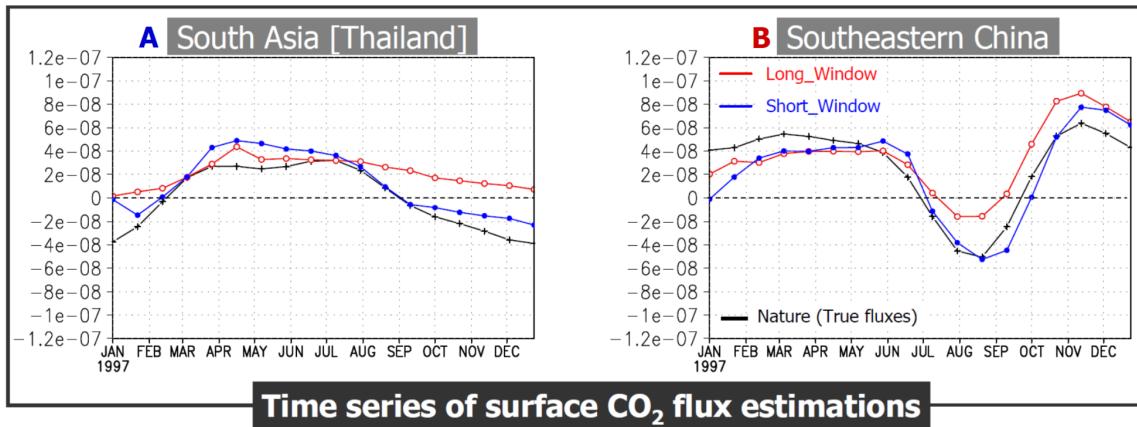


Fig. 5: Evolution of the surface carbon fluxes at two regions A; South Asia (Thailand), which is a sink from August to February in the nature run; and B: Southeastern China, which is a strong source, except from July to October. It is clear that the short window (6 hours) is more accurate, whereas the long window maintains Thailand

as a source throughout the year, and underestimates the intensity of the sink in China during the second half of the year.

## 5. Testing of the LETKF-C on the higher resolution CAM3.5 model:

In order to test whether this methodology can still be successful on a higher resolution and more realistic model such as the NCAR Community Atmospheric Model version 3.5, with  $2.5^\circ \times 1.9^\circ$  horizontal resolution and 26 vertical levels up to 0.03 hPa, we ported the LETKF-C to the CAM 3.5 model, in collaboration with Junjie Liu, of JPL. We started, as we did with the SPEEDY model, with an OSSE (perfect model) in order to test the methodology when we know the truth. Fig. 2 over Europe (a data rich region) shows excellent results confirming the robustness of our approach.

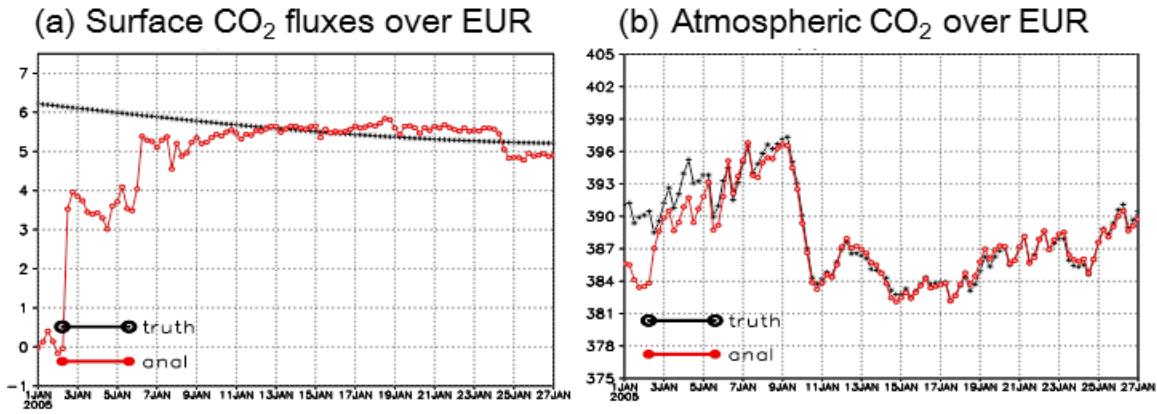


Fig. 6: Comparison of the true surface carbon fluxes and near surface atmospheric CO<sub>2</sub> with the LETKF-C corresponding analyses over Europe, showing excellent agreement in this data rich region, and a spin-up of only about 10 days, even though the initial conditions were zero for the carbon fluxes and random for the atmospheric CO<sub>2</sub>.

## 6. Contributions of Co-I Liu (JPL)

- Liu et al (2012) published a paper on assimilating (real) AIRS CO<sub>2</sub> observations along with meteorology observations with Ensemble Kalman filter using the CAM3.5 model, in *J. Geophys. Res* [1]. The paper concluded that assimilating AIRS CO<sub>2</sub> observations improves the accuracy of CO<sub>2</sub> vertical profiles.
- We investigated the sensitivity of AIRS CO<sub>2</sub> observations to the surface CO<sub>2</sub> flux with GEOS-Chem adjoint model and presented the work in the AIRS science team meeting in April 2012 in Pasadena.
- Liu investigated the sensitivity of AIRS CO<sub>2</sub> observations to the surface CO<sub>2</sub> flux with GEOS-Chem adjoint model and presented the work in the AIRS science team meeting in April 2012 in Pasadena.
- Liu investigated the accuracy of ACOS-GOSAT v2.9 CO<sub>2</sub> retrievals with the ensemble CO<sub>2</sub> analyses generated by assimilating AIRS CO<sub>2</sub> along with surface flask and TCCON CO<sub>2</sub> observations. We presented this work in the OCO-2 science team meeting held in May 2012.

- Liu coupled an ensemble Kalman filter with the latest Community Earth System Model (CESM) v1.0, and studied the impact of uncertainty in meteorology fields on the carbon flux simulation in the model, which was presented in the MODIS science team meeting held in Silver Spring.
- Liu started building the data assimilation system to assimilate leaf area index into the land model of the CESM

## 7. Future plans:

As we have shown in this progress report, we have succeeded in showing that with an advanced data assimilation system like the LETKF-C that we developed, it is indeed possible to estimate accurately and with high resolution the evolution of the surface carbon fluxes. We should include the caveats that we assumed a perfect model in creating the nature run, and we did not include a diurnal cycle in the SPEEDY model, but still this is an extremely encouraging result. As indicated before, Dr. Wouter Peters, who created the CarbonTracker system in NOAA, publicly indicated that our system is better in many ways than his own, and that he plans to upgrade it following our methodology. We also note that a similar methodology can be applied to the estimation of surface fluxes of heat, moisture and momentum, and that preliminary experiments gave promising results (Kang et al., 2013, *in preparation*). This would be a major breakthrough, since we don't have currently good estimates of these very important fluxes.

Given our success and promising results, we would like to submit a follow-on proposal that would involve the collaboration of Dr. Ji-Sun Kang (now at the Korea Institute of Atmospheric Prediction Systems), with Profs Ning Zeng and Kalnay, from UMD, Prof. Inez Fung, from UCB, and Dr. Junjie Liu, from JPL.

## References:

Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, Cambridge Univ. Press, New York, doi:10.1017/CBO9780511535741.

Kang, J.-S. (2009), Carbon cycle data assimilation using a coupled atmosphere-vegetation model and the local ensemble transform Kalman filter, PhD thesis, Univ. of Md., College Park.

Kang, J.-S., E. Kalnay, J. Liu, I. Fung, T. Miyoshi, and K. Ide (2011), "Variable localization" in an ensemble Kalman filter: Application to the carbon cycle data assimilation, *J. Geophys. Res.*, 116, D09110, doi:10.1029/2010JD014673.

Kang, J.-S., E. Kalnay, T. Miyoshi, J. Liu, and I. Fung (2012), Estimation of surface carbon fluxes with an advanced data assimilation methodology, *J. Geophys. Res.*, 117, D24101, doi:10.1029/2012JD018259.

Kang, J.-S., E. Kalnay (2013), Estimating surface fluxes within ensemble Kalman filter data assimilation of atmospheric observations, *in prep.*

Liu, J., I. Fung, E. Kalnay, J.-S. Kang, E. T. Olsen, and L. Chen (2012), Simultaneous assimilation of AIRS Xco<sub>2</sub> and meteorological observations in a carbon

climate model with an ensemble Kalman filter, *J. Geophys. Res.*, 117, D05309, doi:10.1029/2011JD016642.

Liu, J., I. Fung, E. Kalnay, and J.-S. Kang (2011), CO<sub>2</sub> transport uncertainties from the uncertainties in meteorological fields, *Geophys. Res. Lett.*, 38, L12808, doi:10.1029/2011GL047213.

Miyoshi, T. (2011), The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter, *Mon. Weather Rev.*, 139, 1519–1535, doi:10.1175/2010MWR3570.1.

Molteni, F. (2003), Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multidecadal experiments, *Clim. Dyn.*, 20, 175–191.

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans, 2005: “An ensemble data assimilation system to estimate CO<sub>2</sub> surface fluxes from atmospheric trace gas observations”. *JGR*.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth (2008), Ensemble data assimilation with the NCEP Global Forecast System, *Mon. Weather Rev.*, 136, 463–482, doi:10.1175/2007MWR2018.1.

## Selected presentations

Fung, Inez, 2012. Several invited presentations on CO<sub>2</sub> and Weather Data Assimilation and prediction, including NOAA/ESRL Nov 11 2012 and NCAR Nov 12 2012.

Kang, J.-S., E. Kalnay, J. Liu, I. Fung, and T. Miyoshi, Estimation of Surface CO<sub>2</sub> Fluxes from Atmospheric Data Assimilation, 9th International Carbon Dioxide Conference, Beijing, China, 3-7 June, 2013

Kang, J.-S., E. Kalnay, T. Miyoshi, J. Liu, and I. Fung, Simultaneous data assimilation of CO<sub>2</sub> and meteorological variables within LETKF coupled with NCAR CAM model, 2012 Fall Meeting, AGU, San Francisco, CA, 3-7, December, 2012 (presented by E Kalnay).

Kalnay, E., 2013: Parameter Estimation in EnKF: Surface Fluxes of Carbon, Heat, Moisture and Momentum. Invited seminar at GFDL, January 31, 2013.

Liu, J., I. Fung, E. Kalnay, and J.-S. Kang, (2012), Simultaneous assimilation of AIRS and GOSAT CO<sub>2</sub> observations with ensemble Kalman filter. AGU annual meeting, San Francisco, 3-7 December, 2012.