Estimation of Surface Carbon Fluxes with Data Assimilation

*University of Maryland, College Park, +NASA, JPL, #UC Berkeley

Thanks to: Members of Weather-Chaos Group @ UMD, and Dr. Ning Zeng @ AOSC, UMD

AOSC department seminar
Oct. 20, 2011
Outline

- Introduction
- Objectives of our study
- Carbon Cycle Data Assimilation System
  - Localization of variables – (1)
  - Advanced variance inflation methods – (2)
  - Vertical localization of column mixed observations – (3)
  - Observation impacts – (4)
- Summary of Carbon Cycle Data Assimilation Experiments
- Application to moisture/heat flux estimations
- Future Plans
Introduction

- Substantial increase of atmospheric CO$_2$ caused by human activities
- About half of anthropogenic CO$_2$ emission has sunk into land and ocean
  - The capacity of the land and ocean CO$_2$ uptake varies substantially with time and space, and is strongly dependent on climate anomalies (e.g. El Nino-drought and fire, changes in balance between plant growth and death, etc)

http://earthobservatory.nasa.gov/Library/CarbonCycle/carbon_cycle4.html
Introduction

- In order to understand the carbon cycle and its impact on climate change, we need to quantify the temporal and spatial CO₂ sources and sinks at the Earth’s surface

- **Difficulties:**
  - Lack of the spatial coverage of direct carbon flux measurements
  - Lack of the sophistication of numerical models of the carbon cycle
  - Atmospheric CO₂ mixing ratio measurements are used for estimating surface CO₂ fluxes: “top-down approach”
Inversion Problem: top-down approach

- Inversion modeling (e.g. Gurney et al., 2004; Rodenbeck et al., 2003)
  - in-situ & flask atmospheric CO₂ observations
  - Sub-continental and sub-seasonal scales
  - Inverse of transport model: difficult and ill-posed
  - Computationally impractical for high-density data

- Data Assimilation (DA) (e.g. Peters et al., 2007; Baker et al., 2010; Feng et al., 2009)
  - Satellite CO₂ observations in addition to in-situ & flask data
  - Model-grid scale and weekly estimates (e.g. Carbon Tracker)
Inversion Problem: Issues

- A prior estimate of surface CO₂ fluxes
  - It should be pre-calculated by independent observations or another model simulation
- Transport errors (for several weeks window)
  - It is one of critical factors to degrade the flux estimation (e.g. Stephens et al.; 2007; Miyazaki, 2009; Liu et al., 2011)
Inversion Problem: Issues

- **A prior estimate of surface CO₂ fluxes**
  - It should be pre-calculated by independent observations or another model simulation

- **Transport errors** (for several weeks window)
  - It is one of critical factors to degrade the flux estimation (e.g. Stephens et al.; 2007; Miyazaki, 2009; Liu et al., 2011)
Inversion Problem: Issues

Minimizing the difference between the simulated CO_2 concentration and the observed CO_2, prior errors of CO_2 variables

\[ J = \frac{1}{2} \left[ \frac{(M(S) - O)^2}{\sigma_o^2} + \frac{(S - S_o)^2}{\sigma_{S_o}^2} \right] \]

- **A priori fields of surface CO_2 fluxes, S_0**
- **Atmospheric CO_2 observation, O**
- **Transport model, M**
- **a-posteriori estimate, S**

**A prior estimate of surface CO_2 fluxes**
- It should be pre-calculated by independent observations or another model simulation

**Transport errors** (for several weeks window)
- It is one of critical factors to degrade the flux estimation (e.g. Stephens et al.; 2007; Miyazaki, 2009; Liu et al., 2011)

- **Require a prior estimate of surface CO_2 flux fields**
- **Don’t account for the transport errors explicitly**
Objectives of our study

- Explore the feasibility of estimating surface CO₂ fluxes at the model-grid scale by assimilating atmospheric variables (U, V, T, q, Ps) and atmospheric CO₂ simultaneously
  - Consider transport errors in analyzing CO₂ variables
  - No a-priori information for CO₂

"Carbon Data Assimilation with a Coupled Ensemble Kalman Filter"
Supported by Climate Change Prediction Program in Department of Energy

Simulation mode (SPEEDY)
Develop new methodologies
University of Maryland
Prof. Eugenia Kalnay

Realistic System (CAM/CLM)
Assimilating real observation of GOSAT & AIRS
UC Berkeley
Prof. Inez Fung
Local Ensemble Transform Kalman Filter  
(Hunt et al., 2007)

- **Analysis** = $(1-K) \times \text{background} + K \times \text{obs}$
  - $K$ (Kalman gain) is determined by the error statistics of ensemble forecast (background) and observations
  - EnKF provides **background** and **analysis** uncertainty estimation in every analysis step ($P^b$, $P^a$)
  - LETKF assimilates the observations **locally**

$y^0$, $R$, $x^{b(i)}$, $P^{b}$, $x^{a(i)}$, $P^{a}$
UMD-Berkeley LETKF-C System

Parameter estimation: state vector augmentation
- Append \( \text{CF} \) (surface CO\(_2\) fluxes)
- Update \( \text{CF} \) as part of the data assimilation process

Simultaneous analysis of carbon and meteorological variables
- Multivariate analysis with a localization of the variables (Kang et al., 2011)
- Update all variables at every six hours

\[
X^b = \begin{bmatrix}
X \\
\text{CF}
\end{bmatrix}
\quad : \quad \text{model state vector (U, V, T, q, Ps, C)}
\]

\[
\text{CF}
\quad : \quad \text{surface CO}_2 \text{ flux}
\]
(1) Localization of Variables

If variables in the state vector are not physically correlated each other, error covariance between those variables can introduce a sampling error into the analysis system.

Zeroing out the background error covariance between those variables improves the result of the analysis.

(Kang et al., 2011, JGR)
Observing System Simulation Experiments (OSSEs)

- We assume that *we know the true state!*
  - True state (nature run) is generated by a simulation of the model
  - Observations are simulated from the true state
  - Forecast starts from perturbed/random initial guess

  ➔ *See if our new data assimilation techniques or new datasets improve the analysis compared with the truth*

- Test of three data assimilation techniques
  - Localization of variables – (1)
  - Advanced inflation methods – (2)
  - Vertical localization of column mixed CO₂ data – (3)

- Test of impact of CO₂ observations on surface CO₂ flux estimation
  - In-situ & flasks, GOSAT (OCO-2) and AIRS – (4)
OSSEs - (1)

- **Nature run**: assumed true state in the experiments
  - SPEEDY-C: the modified version of SPEEDY (Molteni, 2003)
    - AGCM with a tracer gas of atmospheric CO₂ (C)
    - Spectral model with T30L7
    - Prognostic variables: U, V, T, q, Ps, C
    - No diurnal cycle
  - True CO₂ fluxes (true CF)
    - A constant fossil fuel emission (Andres et al., 1996)

- **Forecast model**
  - SPEEDY-C with persistence forecast of surface CO₂ fluxes (CF)
    - CF is updated only by the data assimilation
  - Initial condition: random (no a-priori information)
    - Ensemble mean of initial CF is close to zero
Simulated Observations - (1)

- Meteorological variables (U, V, T, q, Ps)
  - Rawinsonde
    - Every six hours
- Atmospheric CO\textsubscript{2} concentrations
  - in-situ & flask observations
    - Weekly records: black dots (107)
    - Hourly records: gray dots (18)
  - Satellite data from GOSAT
    - GOSAT provides column mixed CO\textsubscript{2}
      information which has a sensitivity near the
      surface: gray squares
- No direct measurement of surface CO\textsubscript{2}
  fluxes
Results: Impact of variable localization

- The experiments *only with a fossil fuel emission*
  - Variable localization significantly reduce sampling errors

Analysis of CO$_2$ fluxes
- with variable localization
- without variable localization

AOSG department seminar, Oct. 20, 2011
Nature with “evolving” CF

In order to estimate surface CO$_2$ fluxes evolving in time, we need more advanced data assimilation techniques.

- Advanced inflation methods
- Vertical localization of column mixing CO$_2$ data
(2) Inflation Methods

- **Background uncertainty** tends to be underestimated with a limited ensemble size due to the imperfection of the model and nonlinearity of the system.
  - Underestimation of background uncertainty is more serious over the observation-rich area.
  - **EnKF needs “inflation”**

<table>
<thead>
<tr>
<th>Multiplicative inflation</th>
<th>Additive inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiply (1.0+$\alpha$) to the background variance</td>
<td>Add perturbations to the background/analysis state</td>
</tr>
</tbody>
</table>

- **The choice of inflation parameter**
  - $\alpha$ for the multiplicative inflation
  - Scaling factor for the additive perturbation in additive inflation
  - **Manual tuning**: very expensive or often infeasible!
Experiments for inflation methods

- **Fixed multiplicative inflation** *(FixedM)*
  - Standard method
  - Fixed multiplicative inflation parameter (*α*) in time and space

- **FixedM + Additive inflation** *(FixedM + Addi)*
  - Add perturbations to analysis of CO₂ variables

- **Adaptive multiplicative inflation** + Additive inflation *(AdaptM + Addi)*
  - Estimates multiplicative inflation parameter at each grid point at every analysis step adaptively (Miyoshi, 2011)
OSSEs - (2) & (3)

- **Nature run**: assumed true state in the experiments
  - SPEEDY-C: the modified version of SPEEDY (Molteni, 2003)
    - AGCM with a tracer gas of atmospheric CO$_2$ (C)
    - Spectral model with T30L7
    - Prognostic variables: U, V, T, q, Ps, C
    - No diurnal cycle
  - “True” CO$_2$ fluxes (true CF)
    - A constant fossil fuel emission (Andres et al., 1996)
    - CASA terrestrial CO$_2$ fluxes (Gurney et al., 2004)
    - Oceanic CO$_2$ fluxes (Takahashi et al., 2002)

- **Forecast model**
  - SPEEDY-C with persistence forecast of surface CO$_2$ fluxes (CF)
    - CF is updated only by the data assimilation
Simulated Observations - (2) & (3)

- **Meteorological variables (U, V, T, q, Ps)**
  - Conventional data
    - U, V, T, q: black dots (every 12 hours)
    - Ps: gray boxes (every 6 hours)

- **Atmospheric CO₂ concentrations**
  - in-situ & flask observations
    - Weekly records: black dots (107)
    - Hourly records: gray dots (18)
  - Satellite data from GOSAT
    - GOSAT provides column mixed CO₂ information which has a sensitivity near the surface: gray boxes

- No direct measurement of surface CO₂ fluxes
Initial Conditions for Carbon Variables

True CO₂ fluxes @ initial time

Initial condition of surface CO₂ fluxes

No a-priori information!
Results: Impact of the inflation methods

- Fixed multiplicative inflation fails to estimate seasonal changes of CO$_2$ fluxes due to a serious underestimation of background uncertainty.
Results: Impact of the inflation methods

- Fixed multiplicative inflation fails to estimate seasonal changes of CO₂ fluxes due to a serious underestimation of background uncertainty.
Results: Impact of the inflation methods

- Fixed multiplicative inflation fails to estimate seasonal changes of CO₂ fluxes due to a serious underestimation of background uncertainty.
- Additive inflation and adaptive inflation improve the representation of background uncertainty significantly so that the analysis maintains the quality till the end of one-year data assimilation.

AOSC department seminar, Oct. 20, 2011
Results: Impact of the inflation methods

- Global maps of surface CO₂ fluxes in different seasons

A: True fluxes

D: FixedM
Results: Impact of the inflation methods

- Global maps of surface CO₂ fluxes in different seasons

A: True fluxes
C: FixedM+Addi
D: FixedM
Results: Impact of the inflation methods

- Global maps of surface CO$_2$ fluxes in different seasons

A: True fluxes  
B: AdaptM+Addi  
C: FixedM+Addi  
D: FixedM

April

August

January
Impact of the inflation methods on errors

- Analysis errors of atmospheric CO₂ near the surface at the end of one-year DA
  - Adaptive and additive inflations reduce the atmospheric CO₂ errors caused by the imperfection of CF forecast
(3) Vertical Localization

- **Vertical localization** of column mixing CO₂ observation from remote sensing (e.g. GOSAT, OCO-2)
  - Averaging kernel is nearly uniform in the vertical, although the forcing term (our ultimate estimate) is at the surface
  - We have **localized the column CO₂ data, updating only lower atmospheric CO₂** rather than a full column of CO₂
    \[ y_{i}^{b} = h(x_{i,k}^{b}) = \sum_{k=1}^{n_{lev}} a_{k} S(x_{i,k}^{b}) \]
    - Calculating innovation based on the averaging kernel

(Kang et al., in prep.)

Forcing is at the surface

(Wang et al., 2009)

AOSC department seminar, Oct. 20, 2011
Results: Impact of vertical localization

- This method improves analysis of CF mainly over where there are few observations and where there is strong variability of CF
  - We need a careful localization on column mixing CO₂ observations
(4) Observation Impact on errors

- Impact of CO₂ observations on surface CO₂ flux estimation
  - SFC: in-situ & flask data
  - SFC + AIRS
  - SFC + GOSAT
  - SFC + GOSAT + AIRS

![Graph showing the impact of CO₂ observations on surface CO₂ flux estimation](image)
(4) Observation Impact on errors

- Impact of CO₂ observations on surface CO₂ flux estimation
  - **SFC**: in-situ & flask data
  - **SFC + AIRS**
  - **SFC + GOSAT**
  - **SFC + GOSAT + AIRS**

RMSE of surface-layer atmospheric CO₂ over globe (ppmv)

RMSE of surface CO₂ fluxes over globe (gC/m²/yr)
Results: Observation Impacts

- Global maps of surface CO$_2$ fluxes in different seasons

A: True fluxes

D: SFC
Results: Observation Impacts

- Global maps of surface CO$_2$ fluxes in different seasons

A: True fluxes

C: SFC+GOSAT

D: SFC
Results: Observation Impacts

- Global maps of surface CO₂ fluxes in different seasons
  
  A: True fluxes
  B: SFC+GOSAT+AIRS
  C: SFC+GOSAT
  D: SFC

April

August

January
Summary of Carbon Cycle DA

- We succeed in estimating surface $\text{CO}_2$ fluxes with the advanced LETKF-C system, even without *a-priori* information (OSSEs)
  - Localization of variables
    - reduces sampling errors from the correlation between the variables which are not physically correlated
  - Advanced inflation methods
    - represents background uncertainty well
  - Vertical localization of column mixing $\text{CO}_2$ data
    - better estimate surface $\text{CO}_2$ flux changes rather than updating full column of $\text{CO}_2$
- Dedicated $\text{CO}_2$ monitoring satellite (GOSAT) contributes to the surface $\text{CO}_2$ flux estimation significantly
- AIRS $\text{CO}_2$ retrievals help $\text{CO}_2$ flux estimation due to better analysis of atmospheric $\text{CO}_2$ circulation
How about heat/moisture fluxes?

- Can we estimate **surface moisture/heat fluxes** by assimilating atmospheric moisture/temperature observations? *We can use the same methodology!*

- **OSSEs**
  - Nature: SPEEDY
  - Forecast model: SPEEDY with **persistence forecast of Sensible/Latent heat fluxes (SHF/LHF)**
  - Observations: conventional observations of (U, V, T, q, Ps) and AIRS retrievals of (T, q)
  - Analysis: U, V, T, q, Ps + **SHF & LHF**

- Fully multivariate data assimilation
- Adaptive multiplicative inflation + additive inflation
- Initial conditions: random (*no a-priori information*)
Result: Analysis of SHF

True SHF in FEB

True SHF in JUL

True SHF in DEC

Analysis of SHF in FEB

W/m²
Result: Analysis of SHF

True SHF in FEB

True SHF in JUL

True SHF in DEC

Analysis of SHF in FEB

Analysis of SHF in JUL

W/m²
Result: Analysis of SHF

True SHF in FEB

True SHF in JUL

True SHF in DEC

Analysis of SHF in FEB

Analysis of SHF in JUL

Analysis of SHF in DEC
Result: Analysis of LHF

True LHF in FEB

True LHF in JUL

True LHF in DEC

Analysis of LHF in FEB
Result: Analysis of LHF

True LHF in FEB

True LHF in JUL

True LHF in DEC

Analysis of LHF in FEB

Analysis of LHF in JUL
Result: Analysis of LHF

True LHF in FEB

True LHF in JUL

True LHF in DEC

Analysis of LHF in FEB

Analysis of LHF in JUL

Analysis of LHF in DEC
Time series of LHF/SHF

- Black: nature
- Color: analysis of LHF (blue)/SHF(red)

Recall that LHF & SHF are updated only by the data assimilation here!

Promising results from the estimation of “evolving parameters” with data assimilation
Future Plans

- More CO$_2$ datasets
  - HIAPER Pole to Pole Observations (HI PPO) data (Wofsy, 2011)
  - Orbiting Carbon Observatory (OCO-2) data

- The advanced LETKF + CAM3.5 or CAM5 model with **real** observations
  - On-going project

- Imperfect model experiments for both CO$_2$ fluxes and SHF/LHF
  - Impact of model error on flux estimation
  - Bias estimation and correction
Future Plans

- More CO\textsubscript{2} datasets
  - HIAPER Pole to Pole Observations (HIPPO) data (Wofsy, 2011)
  - Orbiting Carbon Observatory (OCO-2) data
- The advanced LETKF + CAM3.5 or CAM5 model with real observations
  - On-going project
- Imperfect model experiments for both CO\textsubscript{2} fluxes and SHF/LHF
  - Impact of model error on flux estimation
  - Bias estimation and correction

The End

Thank you for your attention!
Observation error of CO$_2$
3.0 ppmv for GOSAT
2.0 ppmv for AIRS
**LETKF with SHF/ LHF**

\[
X^b = \begin{bmatrix} X \\ F \end{bmatrix}
\]

: model state vector (U, V, T, q, Ps)

: sensible & latent heat fluxes (SHF, LHF)

- No *a-priori* information ➔