Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Estimating Surface Carbon Fluxes Based on a Local Ensemble Transform Kalman Filter with a Short Assimilation Window and a Long Observation Window

[1]Yun Liu, [1]Eugenia Kalnay*, [1]Ning Zeng, [2] Ghassem Asrar, [3]Zhaohui Chen, [4]Binghao Jia

1 Dept. of Atmospheric and Oceanic Science, University of Maryland – College Park
2 Joint Global Change Research Institute/PNNL, College Park, MD
3 School of Environmental Science, University of East Anglia, Norwich, UK
4 State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

*Corresponding author

Abstract

We developed a Carbon data assimilation system to estimate the surface carbon fluxes using the Local Ensemble Transform Kalman Filter and atmospheric transfer model of GEOS-Chem driven by the MERRA-1 reanalysis of the meteorological fields based on the Goddard Earth Observing System Model, Version 5 (GEOS-5). This assimilation system is inspired by the method of Kang et al. [2011, 2012], who estimated the surface carbon fluxes in an Observing System Simulation Experiment (OSSE) mode, as evolving parameters in the assimilation of the atmospheric CO2, using a short assimilation window of 6 hours. They included the assimilation of the standard meteorological variables, so that the ensemble provided a measure of the uncertainty in the CO2 transport. After introducing new techniques such as “variable localization”, and increased observation weights near the surface, they obtained accurate carbon fluxes at grid point resolution. We developed a new version of the LETKF related to the “Running-in-Place” (RIP) method used to accelerate the spin-up of EnKF data assimilation [Kalnay and Yang, 2010; Wang et al., 2013, Yang et al., 2014]. Like RIP, the new assimilation system uses the “no-cost smoothing” algorithm for the LETKF [Kalnay et al., 2007b], which allows shifting at no cost the Kalman Filter solution forward or backward within an assimilation window. In the new scheme a long “observation window” (e.g., 7-days or longer) is used to create an LETKF ensemble at 7-days. Then, the RIP smoother is used to obtain an accurate final analysis at 1-day. This analysis has the advantage of being based on a short assimilation window, which makes it more accurate, and of having been exposed to the future 7-days observations, which accelerates the spin up. The assimilation and observation windows are then shifted forward by one day, and the process is repeated. This reduces significantly the analysis error, suggesting that this method could be used in other data assimilation problems.

Key words: Carbon Data Assimilation, Surface Carbon Flux, LETKF

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions

1. Introduction

The exchange of carbon among atmosphere, land and oceans contributes to changes in the Earth's climate, and is also sensitive to climate conditions. The CO2 concentration in the atmosphere is affected by both the natural variability of the Earth's planetary system, and anthropogenic emissions. The terrestrial and oceanic ecosystems absorb more than one-half of the atmospheric anthropogenic CO2 emission [Le Quéré *et al.*, 2016]. It is thus essential to quantify the dynamics of earth surface carbon fluxes (SCF), and the variations of carbon sources and sinks.

A common approach for estimating SCF from atmospheric CO2 measurements and atmospheric transport models is referred to as a "top-down" approach. The "top-down" methods estimate SCF through techniques such as Bayesian synthesis approach [Rödenbeck et al., 2003; Gurney et al., 2004; Enting, 2002; Bousquet et al., 1999], different types of ensemble Kalman filters (EnKF) [e.g. Peters et al., 2005, 2007; Feng et al., 2009; Zupanski et al. 2007; Lokupitiya et al., 2008], or variational data assimilation method [e.g., Baker et al., 2006, 2010; Chevallier et al., 2009].

Kang et al. [2011, 2012] developed a "top-down" carbon data assimilation system by coupling an atmospheric general circulation model (AGCM), including atmospheric CO2 concentrations, with the Local Ensemble Transform Kalman Filter (LETKF) [Hunt et al., 2007]. The meteorological variables (wind, temperature, humidity, surface pressure) and CO2 concentrations were assimilated simultaneously in order to include the uncertainties of meteorological field, and their impact on the transport of atmospheric CO2. They carried out Observing System Simulation Experiments (OSSEs), and their carbon assimilation system achieved for the first time an accurate estimation of the evolving SCF at the model grid resolution, without requiring any *a priori* information. The carbon surface fluxes were obtained from the data assimilation as "unobserved evolving parameters", by augmenting the state vector at each column with a surface carbon flux (SCF). The Local Ensemble Transform Kalman Filter (LETKF) then estimated this evolving parameter from the error covariance between the low level atmospheric CO2 and the estimated SCF, and after a spin-up of about one month, the LETKF accurately recovered the nature run seasonal surface carbon fluxes.

Kang et al., [2011, 2012] used a short 6-hour assimilation window for both atmospheric and CO2 observations because atmospheric observations are usually assimilated at this frequency, and because all Ensemble Kalman Filters require short windows to ensure that the forecast perturbations growth remains linear. Such short data assimilation, required by the LETKF, also protects the system from becoming ill conditioned [Enting, 2002, Fig. 1.3], and as a result it does not require additional *a priori* information.

We note that the use of such short assimilation window differs very much from most other "top-down" approaches for estimating SCF that use long assimilation windows varying from a few weeks to months [e.g., Baker et al., 2006, 2010; Peters et al., 2005, 2007; Michalak, 2008; Feng et al., 2009].

Although the Kang et al. methodology was successful, it is computationally expensive, requiring ensemble forecasts and data assimilation not only for the carbon variables, but also for the standard atmospheric variables, in order to estimate the uncertainties of the CO2 atmospheric transport process. In this study, we used an LETKF carbon cycle data assimilation system with a state-of-the-art atmospheric transport model, the GEOS-Chem [Bey et al., 2001; Nassar et al., 2013], which is driven by the MERRA-1 reanalysis of the Goddard Earth Observing System Model, Version 5 (GEOS5). As a result, our system, unlike Kang et al [2011, 2012] does not include an estimation of transport uncertainties related to the meteorological field.

The ultimate goal of our LETKF_C system is to estimate the grid-point SCFs, which, as in Kang et al. [2011, 2012], are treated like time-evolving parameters in the system. As mentioned before, an Ensemble Kalman Filter requires a short assimilation window in order to have the ensemble perturbations evolve linearly and remain Gaussian. On the other hand, it is well known that the training needed to estimate evolving parameters through data assimilation could be quite long, so that it benefits from having many observations. Therefore, a short assimilation window would slow down the training needed for the estimation of the SCF error covariance, and hence lengthen the spin-up time.

To address this problem, we developed a new version of the LETKF related to the "Running-in-Place" (RIP) method used to accelerate the spin-up of EnKF data assimilation [Kalnay and Yang, 2010; Wang et al., 2013; Yang et al., 2012]. Like RIP, it uses the "no-cost smoothing" algorithm for the LETKF [Kalnay et al., 2007b] that allows shifting at a negligible cost the Kalman Filter solution forward or backward within a given assimilation window. Briefly, the new scheme works like this: a long "observation window" (e.g., 7-days, containing all the observations within 7 days) is used to create a temporary LETKF ensemble analysis at 7-days. Then the RIP smoother is used to obtain a final analysis at 1-day. This analysis has the advantage of being based on a short assimilation window, which makes it more accurate, and of having been exposed to the 7-days of observations, which accelerates the spin up time. The assimilation and observation windows are then shifted forward by one day, and the process is repeated. We have tested this new method (short assimilation, long observation window) achieving a significant reduction of analysis errors, and we believe that the method could be useful in other data assimilation problems.

This paper is organized as follows: Section 2 briefly describes the LETKF_C system. Section 3 explores the effect of combing assimilation and observation windows in an OSSE framework. Section 4 presents results on the proposed methodology. A summary and discussion are presented in section 5.

2. LETKF_C data assimilation system

A data assimilation system includes a forecast model, observations, and a data assimilation method that optimally combines them. In the proposed LETKF_C data assimilation system we use the GEOS-Chem as the forecast model and LETKF as the data assimilation method. The pseudo-observations of our OSSE experiments are created at the locations of the real carbon observations retrieved from Orbiting Carbon Observatory-2 (OCO-2) satellite [Crisp et al., 2004].

2.1 GEOS-Chem model and the "nature" run

GEOS-Chem is a global 3-D atmospheric Chemical transport model driven by the NASA reanalysis (MERRA-1) of meteorological fields from the Goddard Earth Observing System data assimilation System Version 5, by the NASA Global Modeling and Assimilation Office [Bosilovich et al., 2015]. This model has been applied worldwide to a wide range of atmospheric composition and transport studies. The

GEOS-Chem model used in this study is the version v10-01 with a resolution of 4° x 5° (latitude x longitude), and 47 hybrid pressure-sigma vertical levels for CO2 simulation [Nassar et al., 2013]. GEOS-Chem is driven by the MERRA-1 reanalysis with 72 hybrid vertical levels, extending from the surface up to 0.01 hPa. The data is provided by the GEOS-Chem support team, based at the Harvard and Dalhousie Universities with support from the NASA Earth Science Division and the Canadian National and Engineering Research Council, who re-gridded the original data of spatial resolution of 0.25° x 0.3125° into the resolution of 4° x 5°.

GEOS-Chem requires the SCFs as a set of parameters at each grid point in order to simulate the CO2 concentration in the atmosphere. It is not possible to observe the global SCFs directly. Therefore, the SCFs are created from a "bottom-up" approach (used as "truth" in our experiments) and used for the simulation of atmospheric CO2 concentration with GEOS-Chem. The "bottom-up" SCFs used in this study include the three components shown in Equation (1): 1) terrestrial carbon fluxes (Fta); 2) air-sea carbon fluxes (Foa); 3) anthropogenic fossil fuel emissions (Ffe).

$$SCF = Fta + Foa + Ffe \qquad (1)$$

The Fta values are derived from the VEgetation Global Atmosphere Soils (VEGAS) model [Zeng et al., 2004; Zeng et al., 2005], forced by the real evolving weather, as given by the GEOS-Chem. The Foa values are from Takahashi et al. [2002], a climatological seasonal cycle estimated for the 1990s, and the Ffe values are from Fossil Fuel Data Assimilation System (FFDAS) for the year 2012 [Asefi-Najafabady et al., 2014]. The air-sea carbon flux and Ffe values were scaled using the global carbon budget data of Le Quéré et al. [2015], in order to include interannual variations. A nature run for atmospheric CO2 concentration simulation is driven by the SCFs in units of $(\frac{kgC}{m^2yr})$ based on all three datasets.

In OSSEs, the nature run serves as the "truth". We assume that the true "bottom-up" carbon fluxes are not known in our data assimilation experiments, and they will be estimated using the atmospheric pseudo-observations derived from the "truth", as described in more details below. The nature run obtained by coupling GEOS-CHEM with VEGAS is fairly realistic [Liu et al., 2017], so we use it to create the pseudo OCO-2 observations for the period of January 2015- March 2016.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

2.2 Pseudo-Observations

The ultimate goal of this model-data assimilation system is to estimate the SCFs at every land grid point using real observations such as the conventional surface CO2 measurements of GlobalViewplus (GV+) flask network provided by Cooperative Global Atmospheric Data Integration Project [2016], and the carbon observations from satellites such as the Greenhouse Gases Observing Satellite (GOSAT) [Yokota et al., 2004], and the Orbiting Carbon Observatory-2 (OCO-2) [Crisp et al., 2004]. In this study, we use the actual OCO-2 observations locations to develop the pseudo-observations for the OSSE assimilation experiments.

The actual OCO-2 observations cover the entire globe once every 14 days with very high spatial resolution (i.e., ~1 km^2). The observations are the CO2 column-averaged dry air mole fractions over the entire OCO-2 pixel (defined as Xco2). The observation quality is greatly affected by conditions such as cloud cover, surface type and the solar zenith angle at the time of measurement. The OCO-2 retrieval algorithm uses a warning level (WL) between 0 and 19, to indicate the quality of measurements, where WL=0 means "most likely good", and WL=19 means "least likely good" observations. The OCO-2 observations used in this study were provided by David Baker (personal communication) who averaged the original high spatial resolution observation into the coarse spatial resolution of ~1 degree, with an average of 10-second time window, using the "good quality" observations retrieval defined by WL <= 15. We further aggregated these observations at the nearest GEOS-Chem output time of the 0, 6, 12, 18 UTC for each model day. The actual location, time and error scales of the OCO-2 observations were then used to create the pseudo-observations. The typical one-day coverage of observation of OCO-2 is shown in Figure 1. The values of Xco2 in the winter are significantly larger than those in summer of the Northern hemisphere. The OCO-2 observations are missing in the winter, for middle and high latitude regions (latitude > ~30). The pseudo-observations are then created by obtaining the "true" CO2 from the "nature" run using the location and time of the valid observation, and then adding random errors with due consideration to the scales of the corresponding real observations. These derived pseudo-observations are more realistic than the GOSAT observations used in Kang et al. [2012], because they are anchored to the original OCO-2 observations and to their quality.

2.3 The LETKF data assimilation system

The ensemble Kalman filter (EnKF) is a powerful tool for data assimilation that was first introduced by Evensen [1994]. The key element of this method is to derive the forecast uncertainties from an ensemble of integrated model simulations. A variety of ensemble Kalman filter assimilation methods have been proposed [Burgers et al., 1998; Houtekamer and Mitchell, 1998; Anderson, 2001, 2003; Bishop et al., 2001; Whitaker and Hamill, 2002; Tippett et al., 2003; Ott et al., 2004, Hunt et al., 2004]. The local ensemble transform Kalman filter (LETKF) introduced by Hunt et al. [2007] is chosen for this study.

The LETKF is an extension of the Local Ensemble Kalman Filter [Ott et al., 2004] with the implementation of the ensemble transform filter [Bishop et al., 2001; Wang and Bishop, 2003]. It is widely used for data assimilation, including several operational centers, and was also used for carbon data assimilations by Kang et al. [2011, 2012].

As discussed earlier, we follow Kang et al., [2011] in estimating the SCFs as evolving parameters, augmenting the state vector C (the prognostic variable of atmospheric CO2) with the parameter SCF, i.e., $\mathrm{X} = [\mathrm{C}, \mathrm{SCF}]^{\mathrm{T}}$. The analysis mean $\overline{\mathrm{X}}^{\mathrm{a}}$ and its ensemble perturbations X^{a} are determined by Equations (2.1 and 2.2) at every grid point, and the ensemble analysis is used as initial conditions for the ensemble forecast in the next cycle.

$$\bar{X}^a = \bar{X}^b + X^b \tilde{K}(y^o - \bar{y}^b) \qquad (2.1)$$

$$X^a = X^b[(K-1)\tilde{P}^a]^{1/2} \qquad (2.2)$$

Here $\bar{X}^b$ is the ensemble mean of the forecast (background) ensemble members; X^b is a matrix whose columns are the background perturbations of $X_k^b - \bar{X}^b$ for each ensemble member X_k^b (k=1,…,K), where K is the ensemble size); y^o is a vector of all the observations; $\bar{y}^b$ is the background ensemble mean in observation space ($\bar{y}^b = H(\bar{X}^b)$), where H is the observation forward operator that transforms values in the model space to those in the observation space); $\tilde{P}^a = \left[(Y^b)^T R^{-1} Y^b + \frac{(K-1)I}{r}\right]^{-1}$ is the analysis error

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Open Access
EGU
Discussions

covariance matrix in ensemble space, which is a function of $Y^b = HXb$, the matrix of background ensemble perturbations in the observation space, R, the observation error covariance (e.g., measurement error, aggregation error, representation error), and of r , a multiplicative inflation parameter; and $\widetilde{K} = \tilde{P}^a Y^b R^{-1}$. LETKF assimilates simultaneously all observations within a certain distance at each analysis grid point, which defines the localization scale. Hunt et al. [2004] introduced a 4-dimensional version, and Hunt et al. [2007] provide a detailed documentation of the 4D-LETKF which we are using.

2.4 Choosing the long observation window (OW) and the short assimilation window (AW)

Like other data assimilation methods, LETKF proceeds in analysis cycles that consist of two steps, a forecast step and an analysis step. In the analysis step, the model forecast (also called prior or background), and the observations, are optimally combined to produce the analysis (also called the posterior), which is the best estimate of the current state of the system under study. In the forecast step, the model is then advanced in time with the analysis as the initial condition and its result becomes the forecast for the next analysis cycle. The assimilation time window for a regular 4D-LETKF is the length of the forecast ~~step~~. All observations within the assimilation time window are used to constrain the state at the end of the assimilation window.

The focus in this study is on the estimation of SCFs that are time varying parameters in GEOS-Chem. As discussed earlier, a preliminary LETKF analysis, which provides the weights for each ensemble perturbation, is performed over a longer window (e.g., 7 days with observations starting at time t). Then, the "No-Cost" smoothing [Kalnay et al, 2007b, Kalnay and Yang, 2010] is applied, using the same analysis weights obtained at the end of the long observation window (e.g., 7 days) for each ensemble member, but combining the ensemble perturbations at the end of the corresponding short assimilation window (e.g., 1-day). This creates the final 1-day analysis (at time t+AW), which benefits from the information from all the observations made throughout the long OW (7 days), and from the linearity of the perturbations in the short AW of 1 day, which is required for accuracy. At this time the procedure is repeated starting at t+AW, one day later.

In this new approach, we have the flexibility to combine a short assimilation window (AW) of length m, with a long observation window (OW) of length n, to improve the estimation of SCF. In the forecast step, the model is integrated from t to t+n, to produce the forecast corresponding to the observations within the OW. In the analysis step, the observations and corresponding forecasts within the OW are used by the LETKF to estimate optimal weights for the ensemble members. The "No-Cost" smoother applies these optimal weights to determine the analysis of the model state and the SCF parameter at $t + m$. The resulting analysis is then used as the initial conditions for the next analysis cycle starting from time $t + m$.

2.5 Experimental setup

In our experiments we used an ensemble size of 20 members, which was optimal in the sense that increasing the number of ensemble members did not significantly improve the results. The initial ensemble is created by random selection of the state and flux values from the model-based "nature" run for both SCF and atmospheric CO2 concentration. Therefore, the initial uncertainties of fluxes and CO2 values are equivalent to their "natural" variability. Based on a sensitivity analysis, we found a horizontal localization radius of 150 km is optimal for our system. Following Kang el al. [2012], a vertical localization is also applied assigning a larger weight to the CO2 updating on surface layers to reflect the expected dominance of layers near the ground in the change of the column total CO2 measured by OCO-2.

2.6 Additive Inflation Method

The inflation is very important for our LETKF_C data assimilation system. The LETKF uses the forecast ensemble spread to represent forecast uncertainties. All EnKFs tend to underestimate the uncertainty in their state estimate because of nonlinearities and limited number of ensemble members (Whitaker and Hamill, 2002). Underestimating the uncertainty (ensemble spread) leads to overconfidence in the background state estimate, and less confidence in the observations, which will eventually lead the EnKF to ignore the observations and result in filter divergence. This is also true for our carbon-LEKF

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

data assimilation system. The ensemble spread of CO2 in GEOS-Chem model decreases during model integration when the ensemble members are using the same meteorological forcing and SCF values, which is very different from the system with prognostic meteorological fields with the ensemble spread of model state increasing during model integration (not shown). The ensemble spread of SCFs also does not increase during model integration because the SCFs are predicted using persistence. However, the LETKF decreases the ensemble spreads for both SCFs and CO2 during analysis steps. Therefore, without inflation, the ensemble spread of the CO2 and SCFs would be continuously decreasing during data assimilation, and soon would become too small for LETK to accept any observations, and hence, cause filter divergence.

There are different types of inflation methods that address the problem of overconfidence: multiplicative inflation, relaxation to prior, and additive inflation [e.g. Anderson and Anderson, 1999; Mitchell and Houtekamer, 2000; Zhang et al., 2004; Whitaker et al., 2008; Miyoshi, 2011]. For this study, we chose additive inflation, which adds random fields to the analysis before the ensemble forecast of the next analysis cycle. Additive inflation has some advantages compared to multiplicative inflation because it prevents the effective ensemble dimension from collapsing toward the dominant directions of error growth [Whitaker et al., 2008; Kalnay et al., 2007a]. We applied additive inflation to the ensemble of atmospheric CO2 and SCF to increase perturbations in the initial conditions, for the next time step. Following Kang et al [2012], the added field for each ensemble member is selected randomly from the nature run. Pairs of atmospheric CO2 and surface CO2 flux fields are chosen randomly within one year before the analysis time and then scaled to a magnitude corresponding to 30% of their seasonal variance.

3. Sensitivity analysis for AW and OW length

We tested the new version of the LETKF with short AW and long OW, described in previous sections. To test this new technique, we conducted two sets of experiments using the LETKF_C system in an OSSE framework with OCO-2 like observations. The first set of experiments used the regular 4D-LETKF settings (with a single window length AW=OW) to investigate the effect of the length of AW for estimating SCF. In the second set of experiments, we investigated the optimal OW length choosing the best AW from the first set of experiments. The assimilation period for all experiments was 1 January 2015 to 1 March 2016. The annual mean RMSEs differences are calculated by removing from the simulation results the spin-up period of first two months. The details of experimental settings are shown in Table 1.

Table 1. Lengths of Assimilation Window (AW), and Observation Window (OW), and the resulting mean RMSEs for different experiments. The first four experiments use regular 4D-LETKF, with AW=OW. The last four experiments use AW=1 day, found to be optimal, and different OWs.

	EXP1	EXP2	EXP3	EXP4	EXP5	EXP6	EXP7	EXP8
AW	6 hours	1 day	3 days	7 days	1 day	1 day	1 day	1 day
OW	6 hours	1 day	3 days	7 days	2 days	8 days	15 days	30 days
RMSE $(\frac{kgC}{m^2yr})$	0.077	**0.059**	0.068	0.074	0.053	**0.041**	**0.040**	0.050

Sensitivity analysis for different assimilation windows

The sensitivity of SCF estimates to the length of AW was investigated based on the first set of experiments (EXP1-EXP4) with regular 4D-LETKF settings, where the length of OW is the same as that of the AW. All experiments used the same observations and initial conditions. Since the coverage of OCO-2 observation network is too sparse for our LETKF_C assimilation system to estimate the SCF signal in a short time scales, here we focus mainly on the estimation of SCF in the seasonal and longer time scale.

Figure 2 shows the estimated global total surface fluxes from the first set of experiments. The "true" global total surface fluxes show a clear seasonal cycle with very large carbon uptake during the growing season of Northern Hemisphere (NH), from May to August, and carbon release during other seasons with the peak release during November. All experiments reproduced fairly well the seasonal cycle of SCF.

When the AW is very short (6 hours), there are large magnitude and high frequency noise overlaying the seasonal cycle. The magnitude of high frequency errors of SCF estimation in EXP1 is comparable with the seasonal variability of SCF (Figure 2a). When the AW=7 days is selected, the high frequency errors of estimation decay, but the deviations of estimates from the "truth" increase. The EXP2 with AW= 1 day produced the best estimation of SCF among all four experiments with equal observation and assimilation windows (Figure 2).

The advantage of AW=1 day (EXP2) is clearly seen from the smaller average global root mean square error (RMSE) and from Figure 2c. The RMSE of surface carbon flux is calculated as

$$RMSE(t) = \sqrt{E_x\left(\left(F^a(x,t) - F^n(x,t)\right)^2\right)} \quad (3)$$

where x and t are space and time location; F^a and F^n indicate the analysis and the " true" SCF from nature run, respectively. E_x is spatial average. The estimations from experiments with long AW (3 days and 7 days) have the smaller RMSE for the first 3 months (January to March), when the "truth" had very little variation because the long AWs enhances the signal and smoothes the high-frequency noise. The experiments with long AW could miss the fine-scale signals of SCF variation and fail to catch its variation with time. Therefore, the estimations with long AW showed large RMSE during the period when the SCF had larger variations. The estimation with AW of 6 hour showed very large RMSE because of the overwhelming high frequency noise. The estimation with AW of 1 day had the smallest RMSE among all the experiments with regular 4D-LETKF.

The yearly mean RMSEs of SCFs showed very similar spatial patterns, but different amplitudes for different experiments (Figure 3). The large RMSEs of SCF estimation located in Southeast American, Southeast of China and Russia, and resembled that of the SCF variance (not shown). The regions of higher variance indicate more information is needed to resolve such large variance by observations, which is hard to achieve. As expected, the SCF RMSE of 0.059 from EXP2 with AW of 1 day is significantly smaller than the RMSE from EXP1 with a short AW of 6 hour (0.077 $\frac{kgC}{m^2yr}$), and EXP3 and EXP4 with longer AWs of 3 days (0.068 $\frac{kgC}{m^2yr}$) and 7 days (0.074 $\frac{kgC}{m^2yr}$) respectively.

Our results show that the preferred AW for estimating SCF is 1 day. This is distinctly different from previously published studies that indicate either a very short AW (6 hours) [Kang et al 2011, 2012], or a very long AW (longer than a few weeks) [e.g., Baker et al. , 2006, 2010; Peters et al., 2005, 2007; Michalak, 2008; Feng et al., 2009]. A short AW can better constrain the model state and therefore produce a better parameter estimation. It is worth mentioning that a very short AW of 6 hours can degrade the SCF estimation with high frequency noise, in our LETKF-C system. We postulate that the high frequency noise is related to the sampling errors in the CO2-SCF covariance that has smaller signal-noise ratio compared to those in experiments with longer AWs.

Sensitivity analysis for different observation windows (OW)

The results presented earlier and associated discussion suggest that parameter estimation through data assimilation benefits from long training time and having sufficient number of observations, implying that the length of OW is critical for the estimation of desired parameter(s). We investigated this sensitivity to find the suitable length of OW for estimating SCF in the second set of experiments (EXP5-EXP8), all based on the optimum AW=1 day that was identified from the first set of experiments, with different OW lengths.

The estimated global total SCFs in the second set of experiments show a clear seasonal cycle matching the “truth” (Figure 4a). Compared with EXP2 (OW=1) shown with the green line in Figure 2a), EXP5 (OW=2days) reduced the high frequency noise significantly when the OW length was increased from 1 day to 2 days. There is still some high frequency noise in the SCF estimation for EXP5, because the observations for 2 days are not sufficient to smooth out the high frequent noise introduced into the estimation through data assimilation. The estimated global total SCFs for EXP6 (OW=8days), EXP7 (OW=15), EXP8 (OW=30) are much smoother than that of EXP5 (OW=1day), because of their longer OW. However, the estimation for OW of 30 days shows a clear time shift compared with “truth”, especially during the transient period with the majority of plants switching from dormant phase in the winter to the growing phase in the spring. The surface carbon fluxes change rapidly during this period. The time shift can also be seen in the estimations for these experiments with OW of 15 days, but it

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

is less pronounced. In the LETKF technique, most of observations in a long OW are introduced at a time later than the assimilation time. Since the SCFs are temporal evolving parameters, the information (variation) of future surface fluxes is brought into the estimation of current time when the future observations are included in the OW. Therefore, the estimated SCF with a very long OW tend to shift towards its future value. The estimated SCF with moderate OW=8 days and 15 days (EXP6 and EXP7) are more accurate than those with a short OW of 2 days (EXP5) and very long OW of 30 days (EXP8), by avoiding the significant high frequent noise observed in EXP5 (OW=2 days) and severe time shift present in EXP8 (OW=30 days). The global mean RMSEs of estimated SCF from OW=8 and 15 days (EXP6 and EXP7) are significantly smaller than those from OW=2 and 30 days, i.e., EXP5 and EXP8 (Figure 4c).

The spatial pattern of annual average RMSE of SCF for EXP5 (OW=2 days; Figure 5) is similar to those in the first set of experiments, which had short AW=OW (Figure 3). The regions with large RMSE in EXP5 (OW=2 days) disappear with OW=7 and 15 days in EXP6 and EXP7, because the long OWs enhance the signals for SFC estimation. The large RMSE in SCF estimates for EXP8 (OW=30 days) are primarily in the Northern Hemisphere mid-latitudes, because of the time shift in estimations with OW=30 days. The mean RMSEs of experiments with moderate OWs of 8 and 15 days are 0.041 $\frac{kgC}{m^2yr}$ and 0.040 $\frac{kgC}{m^2yr}$, respectively, which is significantly smaller than those from experiments with OWs of 2 days (0.053 $\frac{kgC}{m^2yr}$) and 30 days (0.050 $\frac{kgC}{m^2yr}$).

A longer OW requires a longer forecast period for each forecast step, which results in additional computational time/cost. For example, EXP7 with OW of 8 days used 8-time more computational time for its estimation compared to EXP2. Furthermore, the length of OW is also constrained by the time scale of estimation parameters. A long OW tends to generate a time shift for its estimation. For seasonal and longer time scales, OW(s) in moderate range of 8~15 days appear to be most suitable for the LETKF_C estimates of the SCF. EXP7 and EXP8 show almost the same quality of SCF estimation, but EXP7 has higher computational efficiency. The best configuration thus appears to be EXP7 with an OW of 8 days and AW of 1 day, referred as the “benchmark” experiment hereafter.

We note that the high frequency noise in EXP1 with a short AW of 6 hours can be smoothed out by a long OW (i.e. 8-15 days). We speculate that an experiment with AW of 6 hours and OW 8 days will produce similarly realistic estimations as the "benchmark" experiment; however, it requires much more computational time.

5 Evaluating estimated fluxes from the "benchmark" experiment

With the moderate long observation and short assimilation windows, we obtained best estimates of surface carbon fluxes, and their seasonal cycle. This section describes the SCF estimates from the "benchmark" experiment. Figure 6 shows a comparison of surface carbon fluxes based on the "benchmark" assimilation experiment and nature ("truth") run for Northern Hemisphere Summer and Winter seasons. The "bottom-up" carbon fluxes used in the "nature" run show a very strong seasonal cycle over the continents, except Antarctica. The North Hemisphere mid-latitude areas are very large carbon sinks in the Summer, and carbon sources in the Winter. The strong seasonal cycle of surface fluxes mainly related to the variability of terrestrial ecosystems that absorbs large amount of CO2 during the growing season (Spring and Summer) and release carbon back to the atmosphere during dormant seasons (Fall and Winter). The estimated surface fluxes in the seasonal time scale follow closely the "truth". The benchmark assimilation experiment closely reproduces the spatial pattern of surface fluxes globally, for different seasons. The difference between the benchmark estimation and "truth" shown in Figures 6 d & f are very small. There are positive carbon fluxes over Northern Hemisphere mid-latitudes in the Winter, thus a positive bias in estimated atmospheric CO2 concentration is expected.

A successful estimation of surface fluxes requires a good assimilation of atmospheric CO2, and a good estimation of surface flux parameters helps the model-assimilation system to produce a good analysis of atmosphere CO2. Figure 7 shows the comparison of surface atmospheric CO2 concentrations between the benchmark assimilation experiment and nature ("truth") run for Northern Hemisphere Summer and Winter. The spatial pattern of assimilated CO2 matches the "truth" very well. The analysis successfully reproduced the seasonal cycle of CO2 over Northern Hemisphere mid-latitudes, with low CO2 concentration in Summer (Figures 7a-c) and high CO2 in Winter (Figures 7b-d),

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions

consistent with seasonal cycle of CO2 absorption and release by terrestrial ecosystems. There are positive CO2 concentrations located at high latitudes of North America and far East Asia regions during Winter 2016 (Figure 7f), due to the positive bias in estimated SCF (Figure 6f).

The consistency of annual mean estimated SCF for both benchmark experiment and "truth" is a very important feature for our LETKF_C assimilation system (Figure 8a). In EnKF assimilation the ensemble spread is considered as a good representation of uncertainties associated with both parameters and model state [e.g., Evensen 2007, Liu et al. 2014]. The surface carbon fluxes are special parameters that vary with time and it is very hard to quantify their uncertainty during assimilation. When the ensemble spread of parameters are too small to drive model with a robust response, the estimation fails. The additive inflation with 30% of nature variability is used to maintain the amplitude of parameters ensemble spread. Although the ensemble spread of the global total surface flux, in our experiments, is bigger than its error (Figure 8a), we still estimated very well the global total surface CO2 fluxes (ensemble mean), and their seasonal variability. This is consistent with findings of Liu el al [2014], that parameter estimation can tolerate some inconsistency between parameter ensemble spread and parameter error.

The global mean RMSE of SCF decrease from an initial value of ~0.1 $kg\ C\ m^{-2}y^{-1}$ to ~ 0.04 $kg\ C\ m^{-2}y^{-1}$ in just a few analysis cycles (Figure 8b). It does not further decrease during assimilation because the SCF values vary temporally. The signals added by observations are mainly used to reproduce the temporal variation of SCF.

It is very important for a SCF estimation to reproduce the spatial distribution of the annual mean of the SCF, since it identifies the carbon sources and sinks in the Earth System. Though the amplitude of annual mean SCF is much smaller than the seasonal cycle of SCF, the estimated spatial pattern of annual mean SCF in the benchmark experiment (Eq. 4) is generally consistent with the "truth" (Figure 9).

$$\Delta F(x) = E_t\big(F^a(x,t)\big) - E_t\big(F^n(x,t)\big) \quad (4)$$

In summary, we found that the OSSE experiments using long observation windows and short assimilation windows resulted in the best estimates of SCF.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

6 Summary and Conclusions

We have developed a LETKF-GEOS-Chem carbon data assimilation (LETKF_C) system to estimate the surface carbon fluxes (SCF). The GEOS-Chem is run by the single realization of reanalysis meteorology fields driven by MERRA. The SCF vary spatially and temporally in the system.

The LETKF requires a short assimilation window to avoid an ill-posed condition caused by the nonlinear processes in the forecast model with a long forecast time. The parameter estimation favors a long training period and many observations. Based on these features, we developed a new method to accurately estimate the SCF. The new scheme separates original assimilation time window into observation (OW) and assimilation (AW) windows, allowing the flexibility to apply an OW that is different than the AW. Like RIP, the new technique takes advantage of the "no-cost smoothing" algorithm developed for the LETKF by Kalnay et al. [2007b] that allows to translate the Kalman Filter solution forward or backward within the observation window.

The new method was applied to the LETKF_C system in the OSSE mode using a dataset developed based on the OCO-2 observation characteristics. The sensitivity experiments for this model-assimilation system demonstrated that the new technique, with a short AW and long OW, significantly improves the SCF estimation as compared to regular 4D-LETKF with identical observation and assimilation windows. The best AW for SCF estimation is 1 day, which is different from the typical AW of 6 hours used in the meteorological assimilations. An OW in the range of 8-15 days is required to estimate the surface carbon fluxes for seasonal and longer time scales. The benchmark experiment with AW of 1 day and the OW of 8 days successfully reproduced the mean seasonal and annual SCF.

References:

Anderson JL. (2001), An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev., 129, 2884–2903.

Anderson JL. (2003), A local least squares framework for ensemble filtering. Mon. Wea. Rev., 131, 634–642.

Anderson, J. L., and S. L. Anderson (1999), A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, Mon. Weather Rev., 127, 2741–2758, doi:10.1175/15200493(1999)127<2741:AMCIOT>2.0.CO;2.

Asefi-Najafabady, S., P. J. Rayner, K. R. Gurney, A. McRobert, Y. Song, K. Coltin, J. Huang, C. Elvidge, and K. Baugh (2014), A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results, J. Geophys. Res. Atmos., 119, 10,213–10,231, doi:10.1002/2013JD021296.

Banks, H.T. (1992a), Control and estimation in distributed parameter systems. In: H.T. Banks, Editor, Frontiers in Applied Mathematics vol. 11, SIAM, Philadelphia, pp 227.

Banks, H.T. (1992b), Computational issues in parameter estimation and feedback control problems for partial differential equation systems. Physica D 60, 226-238.

Baker, D. F., S. C. Doney, and D. S. Schimel (2006), Variational data assimilation for atmospheric CO2, Tellus, Ser. B, 58, 359–365, doi:10.1111/j.1600-0889.2006.00218.x.

Baker, D. F., H. Bösch, S. C. Doney, D. O'Brien, and D. S. Schimel (2010), Carbon source/sink information provided by column CO2 measurements from the Orbiting Carbon Observatory, Atmos. Chem. Phys.,10, 4145–4165, doi:10.5194/acp-10-4145-2010.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. Field, A. M. Fiore, Q. Li, H. Liu, L. J. Mickley, and M. Schultz (2001), Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23,073-23,096.

Nassar, R., L. Napier-Linton, K.R. Gurney, R.J. Andres, T. Oda, F.R. Vogel, and F. Deng (2013), Improving the temporal and spatial distribution of CO2 emissions from global fossil fuel emission data sets, J. Geophys. Res. Atmos., 118, 917-933, doi:10.1029/2012JD018196.

Bishop CH, Etherton BJ, and Majumdar SJ. (2001), Adaptive sampling with the ensemble transformation kalman filter. Part i: theoretical aspects. Mon. Wea. Rev., 129, 420–436

Bosilovich MG, Akella S, Coy L et al. (2015) MERRA-2: Initial evaluation of the climate. Series on Global Modeling and Data Assimilation, NASA/TM, 104606.

Burgers G, Van Leeuwen P, Evensen G. (1998), Analysis scheme in the ensemble Kalman filter. Mon. Wea. Rev., 126, 1719–1724.

Chevallier, F., R. J. Engelen, C. Carouge, T. J. Conway, P. Peylin, C. Pickett-Heaps, M. Ramonet, P. J. Rayner, and I. Xueref-Remy (2009), AIRS-based versus flask-based estimation of carbon surface fluxes, J. Geophys. Res., 114, D20303, doi:10.1029/2009JD012311.

Cooperative Global Atmospheric Data Integration Project (2016), Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2015; obspack_co2_1_GLOBALVIEWplus_v2.1_2016_09_02; NOAA Earth System Research Laboratory, Global Monitoring Division. http://dx.doi.org/10.15138/G3059Z

Crisp, D., et al. (2004), The Orbiting Carbon Observatory (OCO) mission, Adv. Space Res., 34, 700–709, doi:10.1016/j.asr.2003.08.062.

Evensen G. (1994), Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., 99(C5), 10143–10162.

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

Feng, L., P. I. Palmer, H. Bösch, and S. Dance (2009), Estimating surface CO2 fluxes from space-borne CO2 dry air mole fraction observations using an ensemble Kalman filter, Atmos. Chem. Phys., 9, 2619–2633, doi:10.5194/acp-9-2619-2009.

HARLIM, J. and HUNT, B. R. (2007), Four-dimensional local ensemble transform Kalman filter: numerical experiments with a global circulation model. Tellus A, 59: 731–748. doi:10.1111/j.1600-0870.2007.00255.x

Houtekamer PL, Mitchell HL. (1998), Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev., 126, 796–811.

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, doi:10.1016/j.physd.2006.11.008.

Liu Y, Liu Z, Zhang S, Jacob R, Lu F, Rong X, Wu S (2014), Ensemble-Based Parameter Estimation in a Coupled General Circulation Model. Journal of climate, 27, 7151–7162.

Le Quéré, C., Moriarty, R., Andrew, et al. (2015), Global carbon budget 2014, Earth

Syst. Sci. Data, 7, 47-85, doi:10.5194/essd-7-47-2015.

Le Quéré C, Andrew RM, Canadell JG et al. (2016), Global Carbon Budget 2016, Earth Syst. Sci. Data, 8, 605-649, doi:10.5194/essd-8-605-2016.

Lokupitiya, R. S., D. Zupanski, A. S. Denning, S. R. Kawa, K. R. Gurney, and M. Zupanski (2008), Estimation of global CO2 fluxes at regional scale using the maximum likelihood ensemble filter, J. Geophys. Res., 113, D20110, doi:10.1029/2007JD009679.

Kalnay E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy (2007a), 4-D-Var or ensemble Kalman filter?. Tellus, Ser. A, 59, 758–773, doi:10.1111/j.1600-0870.2007.00261.x.

Kalnay E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy (2007b), Response to the discussion on "4-D-Var or EnKF?" by Nils Gustafsson. Tellus, Ser. A, 59, 778-780, doi: 10.1111/j.1600-0870.2007.00263.x.

Kalnay, E. and Yang, S.-C. (2010), Accelerating the spin-up of Ensemble Kalman Filtering. Q.J.R. Meteorol. Soc., 136: 1644–1651. doi:10.1002/qj.652

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., Kalnay E, Miyoshi T, Liu J, Fung I (2012), Estimation of surface carbon fluxes with an advanced data assimilation methodology: SURFACE CO 2 FLUX ESTIMATION. Journal of geophysical research, 117.

Mitchell, H. L., and P. L. Houtekamer, (2000), An adaptive ensemble Kalman filter. Mon. Wea. Rev., 128, 416–433.

Michalak, A. M. (2008), Adapting a fixed-lag Kalman smoother to a geostatistical atmospheric inversion framework, Atmos. Chem. Phys., 8, 6789–6799.

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

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.

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski,

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics Discussions
Open Access

L. Bruhwiler, and P. P. Tans (2005), An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304, doi:10.1029/2005JD006157.

Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: Carbon tracker, Proc. Natl. Acad. Sci. U. S. A., 104, 18,925–18,930, doi:10.1073/pnas.0708986104.

Tippett M, Anderson JL, Bishop CH, Hamill TM, Whitaker JS. (2003), Ensemble square root filters. Mon. Wea. Rev., 131, 1485–1490.

Wang, S., M. Xue, A. D. Schenkman, and J. Min (2013), An iterative ensemble square root filter and tests with simulated radar data for storm scale data assimilation. Quart. J. Roy. Meteor. Soc., 139, 1888-1903.

Whitaker JS, Hamill TM. (2002), Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913–1924.

Whitaker JS, X. Wei, Y. Song, and Z. Toth (2008), Ensemble data assimilation with the NCEP global forecast system. Mon. Wea. Rev., 136, 463–482.

Yang, S., E. Kalnay, and T. Miyoshi (2012), Accelerating the EnKF Spinup for Typhoon Assimilation and Prediction. Wea. Forecasting, 27, 878–897, https://doi.org/10.1175/WAF-D-11-00153.1

Yokota, T., H. Oguma, I. Morino, and G. Inoue (2004), A nadir looking SWIR FTS to monitor CO2 column density for Japanese GOSAT project, in Proceedings of the Twenty-fourth International Symposium on Space Technology and Science (Selected Papers), pp. 887–889, Jpn. Soc. for Aeronaut. and Space Sci., Tokyo.

Zeng, N., A. Mariotti, and P. Wetzel (2005), Terrestrial mechanisms of interannual CO2 variability, Global Biogeochemical Cycles, 19, GB1016, doi:10.1029/2004GB002273.

Zeng, N., H. Qian, E. Munoz, and R. Iacono (2004), How strong is carbon cycle-climate feedback under global warming? Geophys. Res. Lett., 31 L20203, doi:10.1029/2004GL020904.]

Zhang, F., C. Snyder, and J. Sun (2004), Impacts of initial estimate and observation availability on convective-scale data assimi- lation with an ensemble Kalman filter. Mon. Wea. Rev., 132, 1238–1253

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Zupanski, D., A. S. Denning, M. Uliasz, M. Zupanski, A. E. Schuh, P. J. Rayner, W. Peters, and K. D. Corbin (2007), Carbon flux bias estimation employing Maximum Likelihood Ensemble Filter (MLEF), J. Geophys. Res., 112, D17107, doi:10.1029/2006JD008371.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions

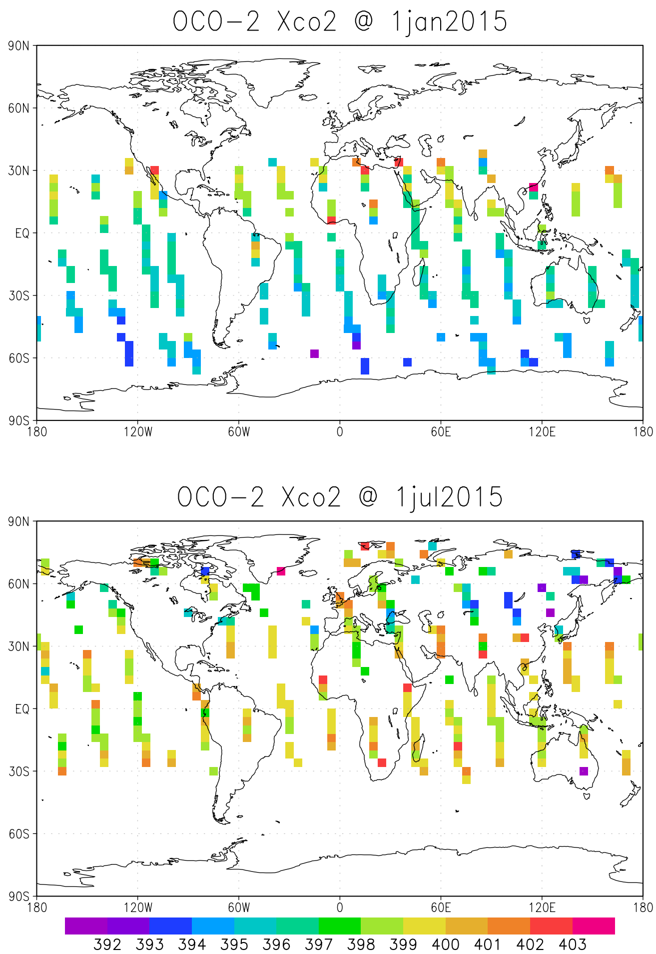


Figure1 The 10-seconds average of good quality OCO-2 Xco2 observations (Warning Level <=15), obtained from David Baker for 1 January 2015 (top panel) and 1 July 2015 (lower panel).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions

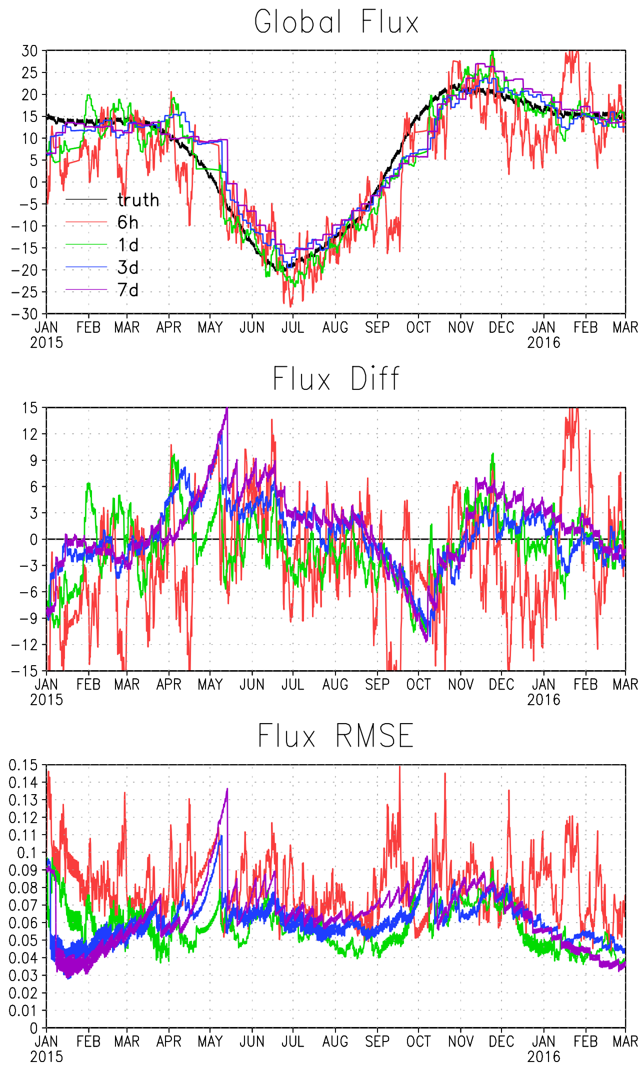


Figure 2 Upper panel: the global total SCF from nature run ("truth", black line) and from the estimations of the first set of experiments with different AW. Middle Panel: the difference of global total SCF between the estimations from the experiments with different AW and the nature run ("truth"). Lower panel: the global average RMSE of the estimated SCFs from the experiments with different AW.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions
Open Access

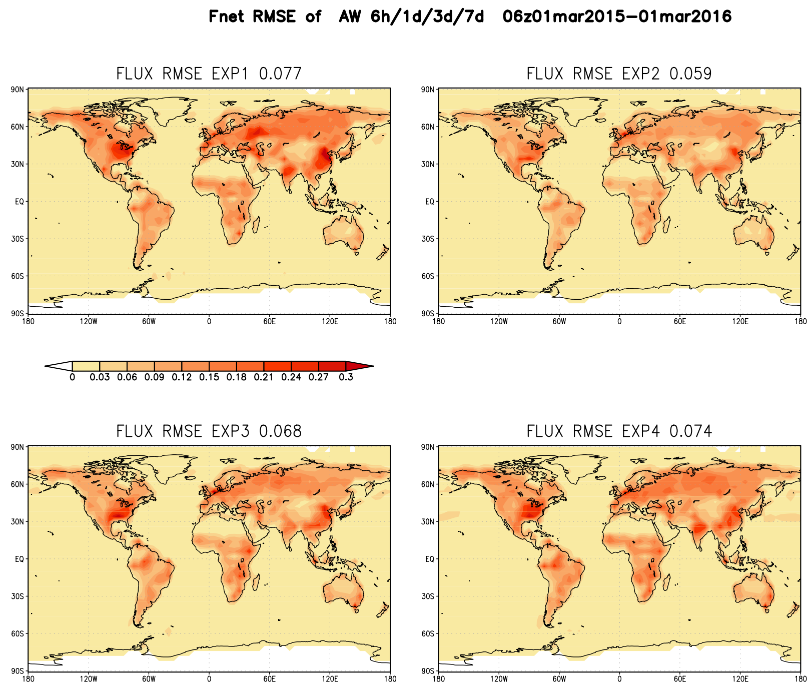


Figure 3 The spatial pattern of the annual mean RMSE of estimated SCF from the experiments with different AW (EXP1-4).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

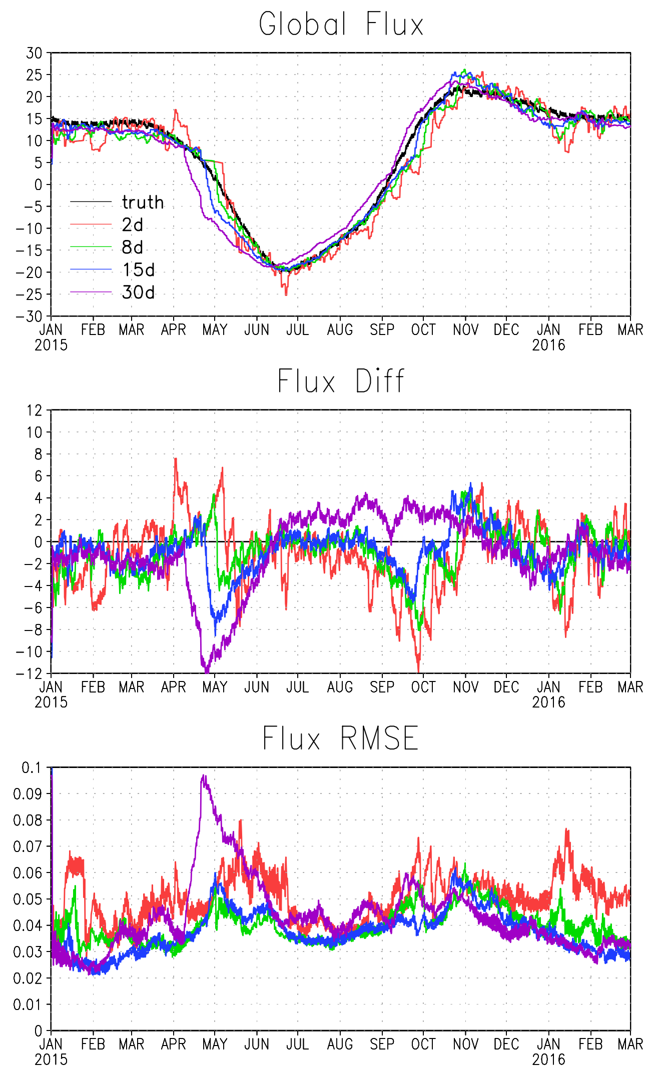


Figure 4 Same as Figure 2, except for the second set of experiments with different OW, but same AW of 1 day.

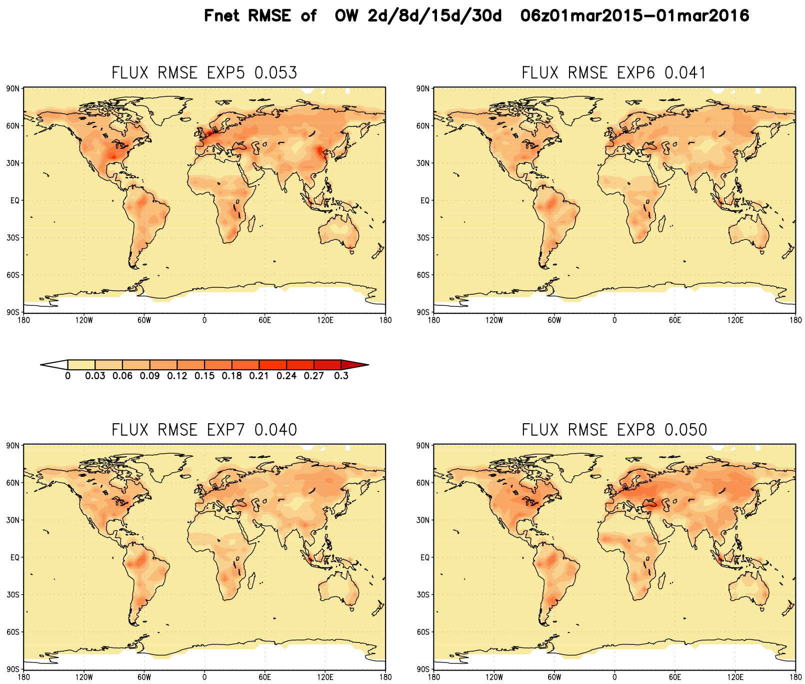


Figure 5 Same as Figure 3, except for the second set of experiments with different OW, but similar AW of 1 day.

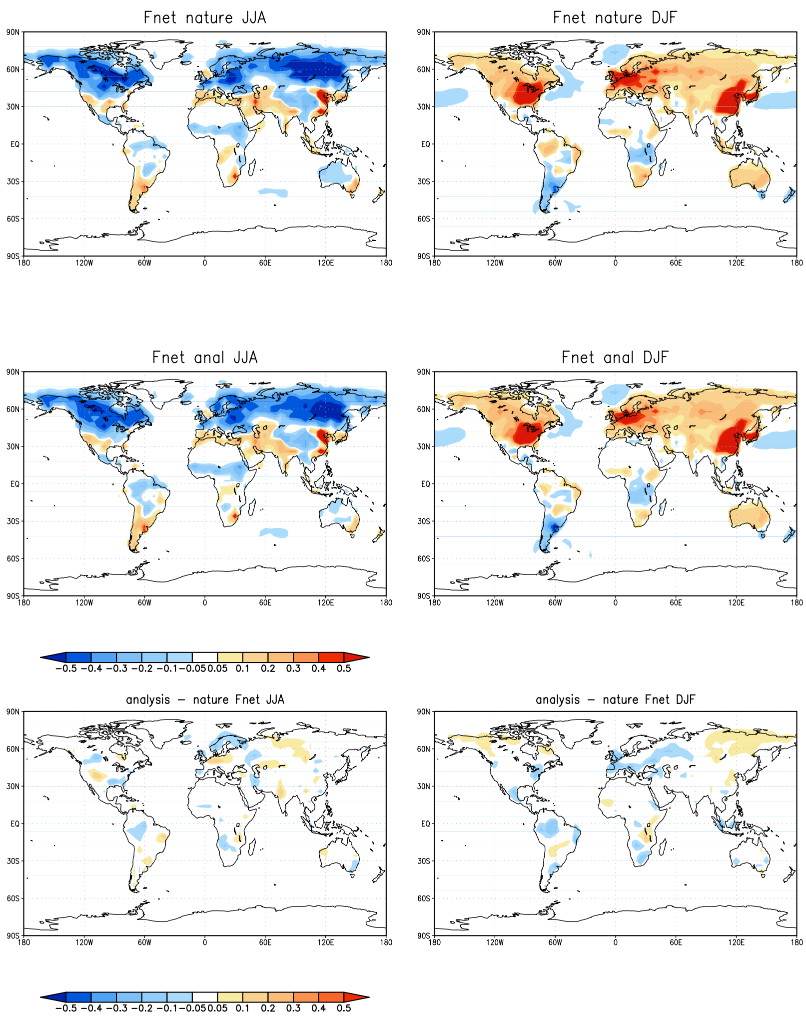
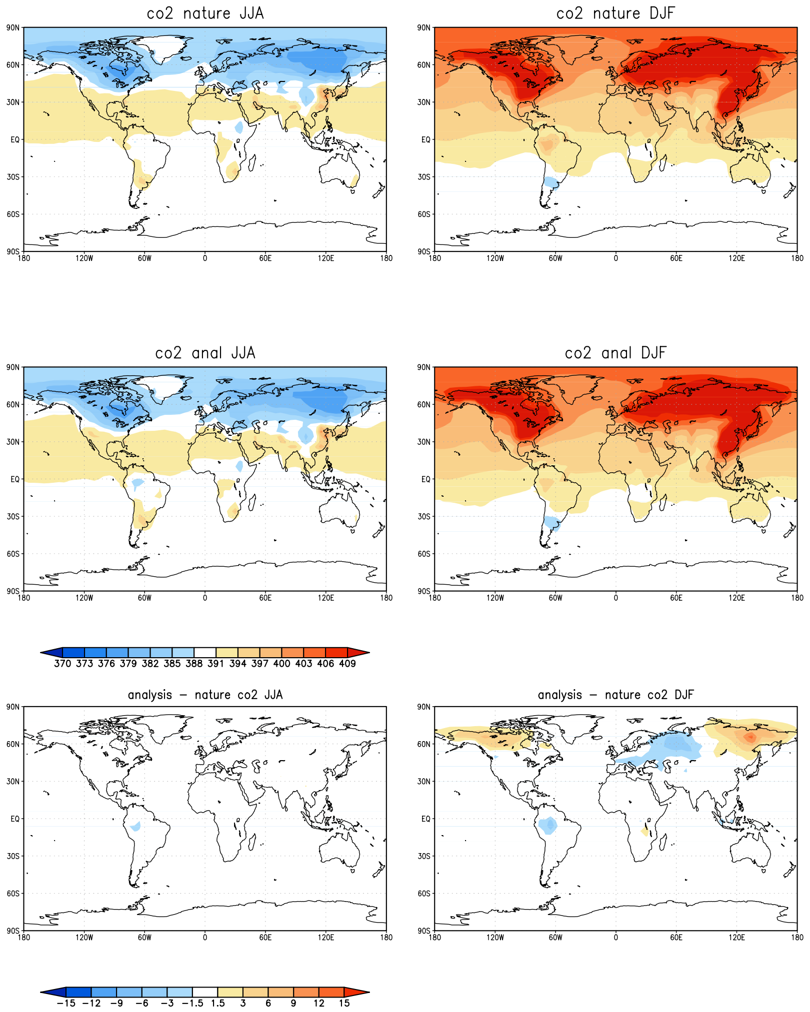


Figure 6: The SCF of “nature” run and estimation from benchmark experiment for Northern Hemisphere Summer (left panels), and Winter (right panels).The top panels are the “truth” from the “nature” run; the middle panels are the estimates from benchmark experiment; and the lower panels are the difference between estimation and “truth”.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics Discussions

1



2
3 Figure 7. Same as Figure 6, except for surface concentrations of CO2
4

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Discussions

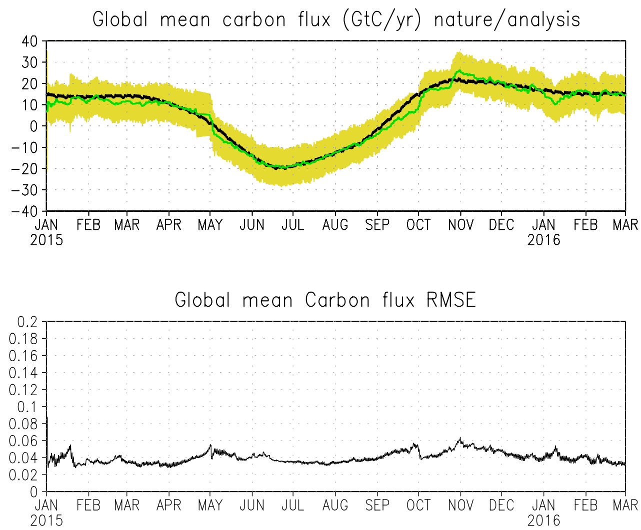
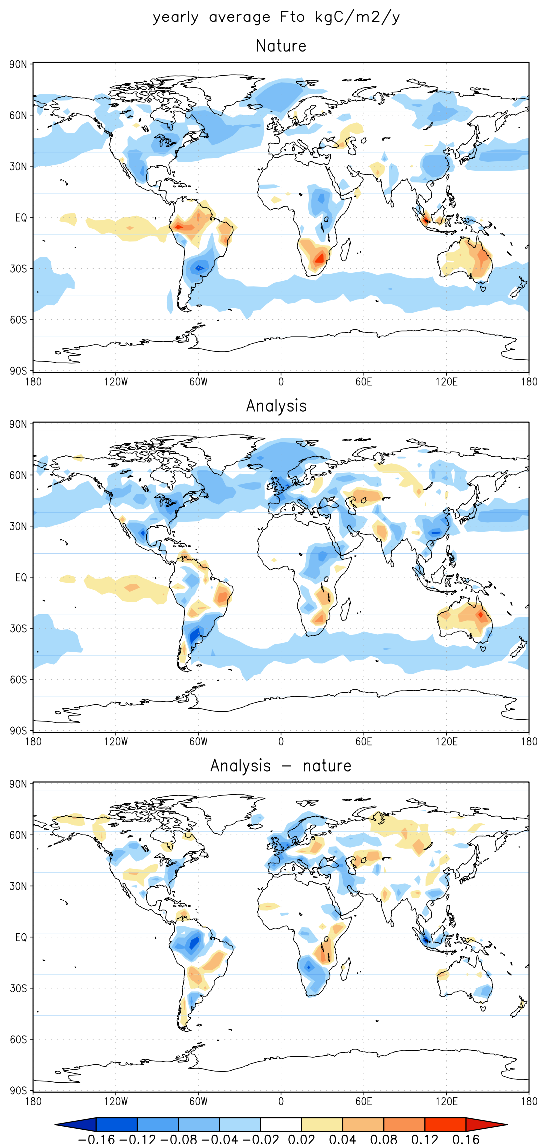


Figure 8. The global total SCF of "truth" and estimation from the benchmark experiment (upper panel);the black line is the truth, green line is the ensemble mean of the estimation, and yellow shading is the ensemble spread. The global mean RMSE of the estimated SCF from the benchmark experiment is presented in the lower panel.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-888
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 16 October 2017

Atmospheric Chemistry and Physics
Open Access
Discussions

1



2
3 Figure 9. The annual mean of SCF (with the FFE removed) for "nature" run (upper
4 panel); the annual mean of estimated SCF (with the FFE removed) from benchmark
5 experiment (middle panel); and their differences (lower panel)