We would like to thank referee 1# Rayner Peter for his comprehensive review and detailed suggestions concerning our manuscript. Those comments are all valuable and very helpful for revising and improving our paper. We have studied these comments carefully and will be able to make the suggested revisions. Our response to the reviewer's comments is given below.

Referee 1#:

General comments:

This paper presents an optimization of 8 years of CO_2 fluxes from the terrestrial biosphere and ocean using a method the authors describe as a dual optimization. I am still a bit unclear on several methodological details of the paper so some of what I'm going to say in the following review is probably wrong. The authors should take notice of my misunderstandings though because they indicate places where the paper should be clearer. The most striking example of this is the elements in the control vector of the optimization. I think this vector contains a series of multipliers (λ) for patterns of terrestrial and ocean fluxes plus one global offset which is used to adjust the atmospheric concentration. The atmospheric concentration is adjusted once per assimilation window (six weeks) but I am unclear about the time resolution of λ. If it is also six weeks then the method seems to be an ensemble version of a classic synthesis inversion (e.g. Enting et al., 1995; Rayner et al., 1999)) with the time windowing technique suggested by Law (2004). If this is correct then some of the claimed advantages of the method don't apply. For example, the authors claim (P14291) that the 1x1 degree resolution of the model avoids the aggregation problem described by Kaminski et al. (2001). In fact the aggregation problem concerns flux patterns in the real world which are outside the subspace spanned by the control vector. The resolution of the flux patterns themselves (i.e. the transport model) doesn't help this problem.

Response: Thank you for these comments. For the time resolution of λ, we actually use a similar strategy to that in Peters et al., (2005). The weekly multiplier at each time *t* in this study is estimated six times and the final result is used as the estimates of λ_t. We will add these descriptions in Sec 2.2 in the revised paper. Moreover, the method in this study is different from an ensemble version of a classic synthesis inversion from two main aspects. Firstly, we use a newly proposed method DOM (Dual Optimization Method, Zheng et al., 2014) to estimate the multipliers (λ) and gridded flux simultaneously in each window instead of using the classical Bayesian synthesis method. In a classical synthesis inversion method, researchers either use the Bayesian Synthesis Inversion Method (shorted for BSIM in my manuscript) to estimate the flux only (e.g. Enting et al., 1995) or use an ensemble version to estimate multipliers (e.g. Peters et al., 2007). The DOM uses the information of CO_2 concentration observations to obtain the optimized multipliers and fluxes simultaneously based on a statistical model (See Zheng et al., 2014 for the details). Secondly, the previous ensemble method usually involves a forecast model for multipliers (e.g. Lokupitiya et al., 2008, Peters et al., 2007, 2010). This is different from our system which does not need a forecast model for the multipliers.

For the aggregation error, we agree that the aggregation problem concerns the internal shape of the flux pattern. We will correct this description in the revised paper. But

GCAS-DOM can still show its superiority. Kaminski et al., (2001) pointed that the aggregation errors for large regions may be of the same order of magnitude as the fluxes themselves and hence inverted fluxes should be cautiously used when answering practical issues. On the other hand, a global inversion in fine grids with long periods based on a batch synthesis inversion method may often be computationally prohibitive. GCAS-DOM incorporates a moving-window method to do global assimilation in a high-resolution grid, and therefore it would significantly advance our understanding of regional carbon cycles.

Referee: I also don't quite understand the computational burden of the problem. As I understand it, the authors solve for approximately 250 fluxes each six week window (weeks here defined like GlobalView with 48 weeks in a year). That's about 2000 unknowns per year or approximately 16000 for the whole period. That's not an immense problem even using the analytic matrix methods. There might be other reasons for the windowing technique, e.g. an effective weak constraint on transport but I don't accept the primary reason is computational.

Response: Thank you for the comment. In fact, the computational amount of the system depends on two aspects: the number of multipliers and fluxes. For each six week window, we solve for about 250 multipliers. However, the resolution of flux is 1 degree, which generates 64800(180*360) elements each week. Therefore, the computational burden mainly depends on the dimension of the flux rather than the multipliers. That's about 3,130,400 (64800*48) unknowns per year and the relevant cost of matrix operations (see Eq. (11)) will be at least 3130400^2 which is an immense problem.

Referee: Of course it's possible I'm completely misunderstanding the approach. The authors may solve for large-scale patterns plus deviations from these, in the spirit of the geostatistical methods pioneered by Michalak and colleagues. If so, please disregard the above but the authors should discuss the relationship with these techniques.

Response: Thank you for the suggestion. We solve for global-scale patterns plus the optimization of model parameters. We will add the discussion about the relationship with geostatistical methods led by Michalak et al. (2004) in the revised paper.

Referee: another concern is independent of the flux resolution and concerns the treatment of the initial condition for each window. Quite reasonably, the prior estimate for this is the result of the simulation of the previous window. The concentration is then corrected by a global offset to minimise the difference with surface values at the end of the window. This updated concentration distribution is used, without correction, as the initial condition for the next window. I see three problems with this:

1. The adjustment to match concentrations introduces a change of CO_2 mass in the atmosphere that is not associated with any fluxes. If this correction has a consistent sign it will lead to a flux series that is inconsistent with the change of CO_2 concentration over the whole timeseries, the aspect of atmospheric CO_2 of which we are most sure.

2. Why correct only the mean concentration? Peylin et al. (2005)) showed a method for improving those aspects of the 3-dimensional concentration distribution observable by the con-

centration measurement network.

Response to problems 1 and 2: Thank you for these comments. There may be some misunderstanding about the treatment of the concentrations for each window. In fact, we not only update the concentrations based on its 3-dimensional distribution, but also added an extra correction according to the mean of observations in situ. These can be described as follows. In a classic batch inversion, the measurement M could be explicitly expressed as (Peylin et al., 2005):

$$M = H_0 C_{|t_0} + H_s S$$

where $C_{|t_0}$ represents the initial condition and S represents the unknown sources during the study period. Now suppose that we have obtained the optimized flux at time *t*, $\hat{s}_t$, and aim to estimate the fluxes at time *t*+1, s_{t+1}. In an assimilation system, the initial condition in the current window includes the impacts not only from the initial condition but also from fluxes before time *t*+1 , which is denoted as *C*(t) (See Fig 1 P14304). Naturally, we run the transport model using the previously stored quantities *C*(t-1) forced by the optimized fluxes $\hat{s}_t$ to get a 3-dimensional distribution of *C*(t). A previous assimilation system e.g. Peters et al., (2005) just used the above 3-dimensional distribution as the initial condition for the window. In this study, we hold the opinion that the spatial pattern of optimized flux $\hat{s}_t$ is rational and therefore can be used to simulate the *C*(t). However, the comparison between simulated and observed concentrations indicates that there exists a system error in the simulated concentration. So we add adjustment of the average for the 3-D concentration to correct this error but keeping the spatial pattern of the 3-D concentration.

Nevertheless, this extra correction only use the average information as the referee said. We have to admit it is a simple strategy but has good effects. To test these effects, we run the system under the same configuration but without the extra correction (similar to that by Peters et al., 2005). The initial concentration prescribed here is set as the site averaged concentration in the last week of 1999. The following figure compares the total annual fluxes of the globe for the two measures with and without extra correction. The system without this correction generates a -8.931 Pg C/year carbon budget for the first year 2000 which may be strongly impacted by the initial concentration, while the system with this correction quickly stabilizes and produces a relatively reasonable annual carbon budget for the spin-up period. The difference of total carbon budget between the two measures diminishes in 2002. Further research on the system bias would be done in a following paper.

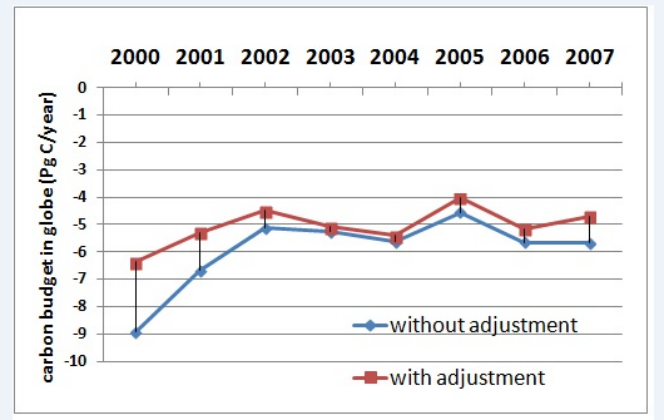

Figure: The total annual fluxes in globe for with and without extra correction

3. No account seems to be taken of uncertainty in the initial concentration field when calculating fluxes. This is a pretty direct consequence of leaving the 3-d concentration out of the state vector. Peylin et al. (2005) also showed that errors in the initial field could affect the model-measurement mismatch for 20 days i.e. about half the assimilation window so it would seem to be important to deal with this.

Response: Thank you for the comment. We agree that the error in initial concentration field can affect the mismatch for a relatively short time (about 20 days). Therefore, the assimilation system usually needs a spin-up period. Normally, we suggest discarding the estimates of the first one to two years in the data assimilation period (Nassar et al., 2011; Deng et al., 2011; Jiang et al., 2013). In our study, we have performed several tests by using different initial concentration values and the results showed that the optimized values after 1 year is reliable and hardly impacted by the initial concentration field. Therefore, our system was initialized with a globally-uniform 3-D CO_2 field of site averaged concentration in the last week of 1999. The estimate from 2001 became stable, and the results from 2001 to 2007 are reasonably used for analysis. We will discard the results in 2000 in the revised paper.

Referee: I also question the use of multipliers for flux patterns themselves rather than the more conventional use of separate multipliers for GPP and respiration. The problem arises from the diurnal cycle. I'm not clear whether the authors retain the diurnal cycle of fluxes from BEPS. If they do then a change of sign of the flux will also change the sign of the diurnal cycle. Since many of the observations in GLOBALVIEW represent particular times of the day this could affect the model-data mismatch at the heart of the inversion.

Response: Thank you for the comment. In fact, we tried to estimate multipliers for GPP and ER, separately. As both ER and GPP are much larger than the NEP fluxes by approximately one order of magnitude, they are sensitive to multipliers. Moreover, the strong correlation between ER and GPP could result in poor performance in stability. Therefore, we decided to adjust the outcome of their differences rather than their separate influences. Moreover, we use the weekly average fluxes from BEPS and corresponding concentration in situ at weekly interval to avoid the problems of the diurnal cycle. In addition, the multipliers are constrained to

be greater than zero by the uncertainties on the λ parameters and hence they would not change the sign of the flux. We will add this explanation in the revised paper.

Referee: My other general concern is prior uncertainties. These are handled via uncertainties on the λ parameters. If I understand correctly these are set at 0.1% for regions outside China and 1% for China. These uncertainties are not arbitrary, they should represent the statistics of differences between simulations of the model used for the prior and the true fluxes. See Chevallier et al. (2006, 2012) for details on how they can be calculated and some indicative numbers from a different model. The uncertainties used in this study seem very low. For example they approach 0 in the transition season as the net flux approaches 0 although the uncertainty should not. This has consequences for the results. The relatively small changes in λ are a likely consequence of these very small uncertainties. I suggest this choice should be justified.

Response: Thank you for these comments. This issue is a key problem in GCAS-DOM. The constrained variance for parameters (P14278 L11) is used to obtain an unique solution for λ, rather than to determine the prior uncertainties of fluxes. The prior uncertainty used in this study is the model error matrix Q for the prior flux (P14281 L17, P14286 L6-10). As we described in our manuscript (P14286L6-10), we use uncertainties of 1.98 Pg C/year and 0.93 Pg C/year for the land and oceanic fluxes, respectively, which are not low compared to previous studies (e.g. Deng et a.,2007, 2011, Gurney et al.,2003,2004).

On the other hand, the flux in a grid consists of six components for different PFTs and the similar climate condition in a grid will lead to strong correlation of these components. So we propose an adaptive version of DOM by adding additional regularization of parameters (P14277 L27-14288 L11). For the range of λ, we think that the value should be around 1 and set an initial interval of [0.7, 1.3], as the preferences of BEPS are basically reasonable. According to the 3 sigma principle, the standard deviation (SD) of parameters is set to be 0.1 (i.e. variance of 0.01). However, the results of regions excluding China (e.g. Europe and North America) under this circumstance are irrational compared to previous studies. This may be caused by the larger error in soil carbon estimate of China in BEPS. Therefore we try to reduce the SD for the other regions and test the values of 0.0707 (i.e. variance of 0.005) and 0.0316 (i.e. variance of 0.001). The results indicate that the setting of 0.1% for regions outside China and 1% for China can get a more reasonable pattern of flux. Our next work will focus on the optimal decision of constrained variance for parameters by the criterion of fit to the observation concentrations at sites. We will add these explanations in the revised paper.

Specific Comments:

P14271L10 note that we don't calculate the PDF by minimizing differences, that's for calculating the maximum likelihood estimate

Response: Thanks for this comment. We will correct the description in the revised paper.

P14276-7 I am confused about the time windows here. Is there perhaps an error?
e.g. We hear that the system is run from time t-1 over l steps but the observations

listed are at t+1, t+2 ... t+l-1, should this be t-1?

Response: Thanks for this comment. This is not an error. The observation concentration c_t represents the concentration at the end of *t*th week, while the flux s_t represents the weekly average flux during the *t*th week. The c_t includes the influence from initial condition at *t*-1 and the tracer emission s_t, called the initial condition term and the source term, respectively. If we want to estimate the flux s_t, we need to run the system *l* steps forward starting from time *t*-1 (See Fig.1) to get the responses at *t*, *t*+1,..., *t*+*l*-1to the initial concentration. Then the differences between these responses and observation concentrations are caused by the emission fluxes $\{s_t, s_{t+1}, \dots, s_{t-l+1}\}$ and hence can be used to estimate them.

P14281 You note that fire and fossil fluxes are not perfectly known and are excluded from the optimization. You need, then, to include their uncertainty in the observational error you use.

Response: Thanks for this suggestion. We agree with your idea that the observational error actually includes the uncertainty of fire and fossil flux emission. Neglecting their uncertainties will increase the error of optimized fluxes. Therefore, we included an extra contribution of $(0.175\text{ppm})^2$ to the observational error (See Eq.(13)). .

P14289 The comparison of posterior simulation and observations is a good idea but highlights some of the problems

Response: Thanks for this comment. Yes, the comparison of concentrations indicates some problems e.g. the inability to capture the peak and the valley and seasonal cycle identified in the residual series. We also explained in this manuscript that this may be caused by the inability of BEPS to simulate the large summer sinks (P14291 L5-15). Therefore we will make more efforts to improve the prior flux in terms of seasonal sink pattern in the future work. These problems help us identify gaps in GCAS-DOM as well as BEPS and provide useful directions for further development.

References:

Deng, F., Chen, J. M., Ishizawa, M., Yuan, C. W., Mo, G., Higughi, K., Chan, D., and Maksyutov, S.: Global monthly CO_2 flux inversion with a focus over North America, Tellus B, 59, 179–190, 2007.
Deng, F. and Chen, J. M.: Recent global CO_2 flux inferred from atmospheric CO_2 observations and its regional analyses, Biogeosciences, 8, 3263–3281, doi:10.5194/bg-8-3263-2011, 2011.

Gurney, K. R., Law, R. M., Denning, A. et al.: TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, 2003.

Gurney, K. R., Law, R. M., Denning, A. S., et al.: Transcom 3 inversion intercomparison: model mean results for the estimation of seasonal carbon sources and sinks, Global Biogeochem. Cy., 18, GB1010,

doi:10.1029/2003GB002111, 2004.

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

Nassar, R., Jones, D. B. A., Kulawik, S. S., et al.: Inverse modeling of CO2 sources and sinks using satellite observations of CO2 from TES and surface flask measurements, Atmos. Chem. Phys., 11, 6029-6047, 2011.

Peters, W., Jacobson, A. R., Sweeney, C. et al.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, 2007.
Peters, W., Krol, M. C., Van Der Werf, G. R. et al.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Glob. Change Biol., 16, 1317– 1337, 2010.

Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., and AEROCARB Contributors: Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmospheric Chemistry and Physics, 5, 3173–3186, doi:10.5194/acp-5-3173-2005, http://www.atmos-chem-phys.net/5/3173/2005/, 2005.

Zheng, H., Li, Y., Chen, J. M., Wang, T., Huang, Q., and Sheng, Y.: Applying a dual optimization method to quantify carbon fluxes: recent progress in carbon flux inversion, Chinese Sci. Bull., 59, 222–226, doi:10.1007/s11434-013-0016-5, 2014