ACP-2014-337 (Editor – Ning Zeng)

Response to Reviewer 1

The authors thank the reviewer 1 for a thoughtful review of the manuscript. The responses for the reviewer's specific comments are as follows.

Comment:

This study evaluates the influence of CO_2 observations on the analysis of CO_2 surface fluxes. The influence matrix concept, which is routinely used within the NWP community, has been employed to assess the benefit of different surface observation sites within the CarbonTracker framework. The novelty of this study is in its application to carbon science since the specific tools/methods discussed here are well established.

General Comment:

It seems that a few choices and assumptions (and accordingly the final results) are very much tied to the CarbonTracker setup that the authors have used. Hence, the conclusions may not be reflective of the performance of a generic ensemble Kalman Filter in which the lag window size, localization, inflation parameters etc. can be tuned. In fact there is no discussion of inflation in Section 2.2. The following comments are intended to provide the authors with a few starting points that can make the study more appealing to the general carbon data assimilation, and not just the CarbonTracker, community.

Author's response: The purpose of this work is to estimate the effect of CO_2 observations on the analysis of surface CO_2 flux in the globe. Until now, no studies have investigated how CO_2 observations are used to optimize the surface CO_2 flux using the influence matrix analysis in the real carbon data assimilation. We think that this work is the first step to diagnose the impact of specific CO_2 observations to the estimated CO_2 flux using any CO_2 inversion technique.

The lag window was set to 5 weeks because several previous studies have already shown that the 5 weeks of window are appropriate to estimate the surface CO_2 flux for the globe in CarbonTracker. Recently, we are investigating which window size and localization are appropriate for the analysis of surface CO_2 flux in Asia using CarbonTracker. The results will be presented in another paper. Again, this study is to investigate the impact of CO_2 observations in the globe not just in local region (e.g., Asia). Therefore, it is reasonable to use 5 weeks of lag window because the 5 weeks lag window have been used and found to be appropriate for the Globe, North America, and Europe in previous studies.

In addition, we have added a discussion of inflation in Section 2.2 as follows.

"Many inflation techniques (e.g., Wang and Bishop, 2003; Bowler et al., 2008; Whitaker et al., 2008; Li et al., 2009; Anderson, 2009; Miyoshi, 2011; Kang et al., 2012) have been used to maintain proper ensemble spread and to improve the performance of EnKF data assimilation. Although the EnSRF in CarbonTracker does not use the inflation method, Kim et al. (2012) demonstrated that the ensemble spread measured by rank histograms is maintained properly."

Specific Comments:

1) By the authors' own admission, a lag window of 5 weeks may not be sufficient to optimize the surface CO_2 flux in Asia (Section 3.3.3). This raises two main questions: a) Why didn't the authors use a lag window of more than 5 weeks? Bruhwiler et al. [2005] (Figure 1 in their paper) showed that for some of the remote sites, the lag window might need to be in the order of months. 5 weeks is suboptimal in that respect, and may very well be the reason why the SH (and a few of the MBL) sites seem to provide little to no information (Figure 8). Can the authors show some sensitivity tests when the lag window is increased beyond 5 weeks? Or is this not feasible given the CarbonTracker setup? If the latter assumption is true, then this drawback needs to be clarified early in Section 1. b) The authors repeatedly claim that the cumulative impact over five weeks would be greater than the average self-sensitivity of 4.8%, which is calculated over the most recent assimilation cycle (i.e., one week). But no quantitative value is provided for this 'cumulative impact'. In general, an ensemble Kalman filter is designed to propagate the covariances in time, and hence the cumulative impact can be calculated over the entire analysis period and not just the most recent assimilation cycle. Again if this is an artifact of the Carbon Tracker setup, then this needs to be clearly stated. Or else the authors need to provide magnitudes for the cumulative impact of the observations.

Author's response: Specific answers for the reviewer's questions are as follows.

a) In CarbonTracker framework, it's possible to change the length of assimilation lag window. We have used a lag window of 5 weeks not because it cannot be changed in CarbonTracker, but because several previous studies using CarbonTracker have reported the lag window of 5 weeks is appropriate to estimate the surface CO_2 flux. Both CarbonTracker North America (Peters et al. 2007) and CarbonTracker Europe (Peters et al. 2010) have used the lag window of 5 weeks. Peters et al. (2007) mentioned that the lag window of 5 weeks is appropriate for North America. In addition, Kim et al. (2012, 2014) and Zhang et al. (2014a and b) have shown that the lag window of 5 weeks could produce realistic surface carbon fluxes in Asia and the globe in CarbonTracker. Because of these many previous studies, we have used the lag window of 5 weeks. As mentioned earlier, the purpose of this study is to estimate the impact of individual observations on a particular CO_2 flux analysis in the globe

using CarbonTracker. Therefore it was necessary to use an appropriate length of lag window for the entire globe not just for Asia.

In previous Section 3.3.3 (3.2.3 in the revised manuscript), we mentioned "In addition, the five-week assimilation lag is effective in optimizing the surface CO_2 flux in this region." Therefore we mentioned the effectiveness of the five-week lag window in Asia, but at the same time we were interested in some possibilities to use a longer lag window for Asia. To investigate which lag window is more appropriate for Asia, we are now testing several different assimilation parameters (e.g., ensemble size, length of lag window, etc.) for Asia using CarbonTracker. Therefore, we have added the following texts at the end of previous Section 3.3.3 (3.2.3 in the revised manuscript).

"A study on the effect of various assimilation window and ensemble size on the estimation of the surface CO_2 flux in Asia is under way to investigate which lag window and ensemble size are appropriate for Asia in CarbonTracker."

In addition, a discussion on the lag window of MBL sites is shown in the response for the reviewer question 2).

b) First of all, the impact can be calculated either in the most recent assimilation cycle or in the length of lag window (or even in the entire analysis period). It is not related with CarbonTracker setup, but related with which one is more appropriate.

Even though Liu et al. (2009) has used an ensemble Kalman Filter, Liu et al. (2009) has calculated the observation impact at each assimilation cycle because there was no lag window in Liu et al. (2009) which is associated with NWP. For the same reason, Cardinali et al. (2004) has calculated the observation impact at each assimilation cycle. There have been no studies on the cumulative observation impact yet. Because we have applied the influence matrix concept to carbon science for the first time and the lagged assimilation window is used in CarbonTracker, we had to consider the cumulative impact as well as the impact in the most recent assimilation cycle.

Following the reviewer's suggestion, we have provided magnitudes for the cumulative impact of the observations in the abstract and Section 3.2.1 as follows.

"Because the surface CO_2 flux in each week is optimized by five weeks of observations, the cumulative impact over five weeks is 19.1%, much greater than 4.8%."

The cumulative impact considers the previous observation effect which is included in the previous analysis. Therefore the forecast from the previous analysis already includes some percentage of previous observation impact. This kind of concept can also be applied to the observation impact calculation for NWP which does not use the lagged assimilation cycle. Because the cumulative observation impact is used for the first time in this study, we have added a schematic (Fig. 2) and texts in Section 2.3 as

follows.

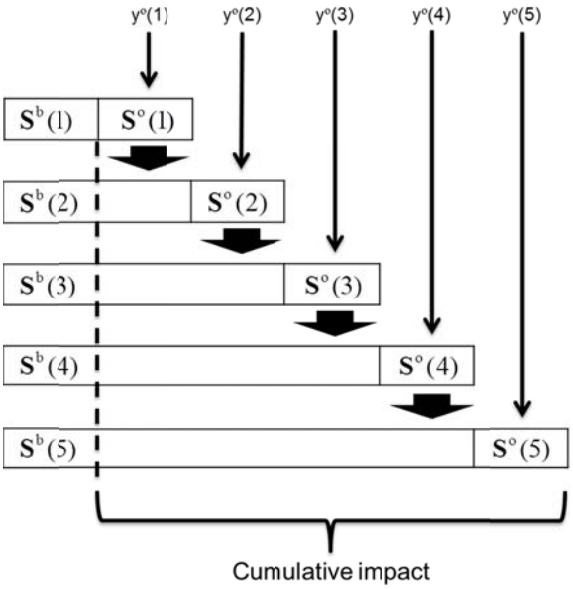


Figure 2. Schematic diagram of calculating cumulative impact in CarbonTraker. $\mathbf{S}^{b}(\bullet)$ indicates the analysis sensitivity to background at each analysis cycle within five weeks of lag, where $\bullet$ denotes each week from 1 to 5. $\mathbf{S}^{o}(\bullet)$ indicates the analysis sensitivity to observation at each analysis cycle.

"The influence matrix is calculated for the most recent week of each cycle because the background at the most recent week of each cycle is updated once by observations.

The cumulative impact of the influence matrix for the five weeks of lag can be calculated because the background in the lagged window already includes the effect from previous observations. For example, Fig. 2 shows that $\mathbf{S}^{b}(5)$ is affected by $\mathbf{S}^{o}(1)$, $\mathbf{S}^{o}(2)$, $\mathbf{S}^{o}(3)$, and $\mathbf{S}^{o}(4)$, where the number inside of parenthesis represent the week of the five-week assimilation lag. If $\mathbf{S}^{o}(\bullet)$ has a value between 0 and 1, $\mathbf{S}^{b}(1)$, the analysis sensitivity to background at the first week, represents an information from previous analysis cycle and is calculated as

$$\mathbf{S}^{b}(1) = (1-\mathbf{S}^{o}(1))(1-\mathbf{S}^{o}(2))(1-\mathbf{S}^{o}(3))(1-\mathbf{S}^{o}(4))(1-\mathbf{S}^{o}(5)), \quad (17)$$

Using Eq. (13), the cumulative impact of the influence matrix is

$$\mathbf{S}^{o}_{cum} = 1 - \mathbf{S}^{b}(1) = 1 - (1 - \mathbf{S}^{o}(1))(1 - \mathbf{S}^{o}(2))(1 - \mathbf{S}^{o}(3))(1 - \mathbf{S}^{o}(4))(1 - \mathbf{S}^{o}(5)), \quad (18)$$

where $\mathbf{S}^{o}_{cum}$ is the cumulative impact of observations during the lagged window. The cumulative impact was defined within the five-week assimilation lag and calculated when $\mathbf{S}^{o}(5)$ exists."

2) Figure 4a – it is particularly curious that the self-sensitivity of the MBL sites are the same as the self-sensitivity of the Difficult sites. In Section 3.2.1, the authors argue that the spread of the analysis CO_2 concentrations is small at the MBL sites. But they have to be an order of magnitude lower to compensate for the fact that the model-data mismatch values at the MBL sites are 10 times lower than the model-data mismatch values at the Difficult sites (based on Table 2). Can the authors show a time-series of how the spread in the analysis CO_2 concentrations compare between these two sets of sites? Are the spread in the analysis CO_2 concentrations that different during the NH winter months? Or is it because that the assimilation system is unable to use the information from the MBL sites, given the constraints on the lag window size?

Author's response: The time-series of the spread of analysis CO_2 concentration [ppm^2] for MBL and Difficult site are shown in Fig. rev_1. As we have denoted in Section 3.2.1, the spread of the analysis CO_2 concentration is much smaller at the MBL sites than that at the Difficult sites.

As denoted in Section 2.2, the observations at MBL sites affects globally because they are considered to include information on large footprints of flux signals as mentioned in Peters et al. (2007). Therefore, regardless of the lag window size, the information from the MBL sites is used well in CarbonTracker. The small spread of analysis CO_2 concentration in the MBL sites are caused by small CO_2 flux spread in Antartica because most MBL sites are located in Antartica.

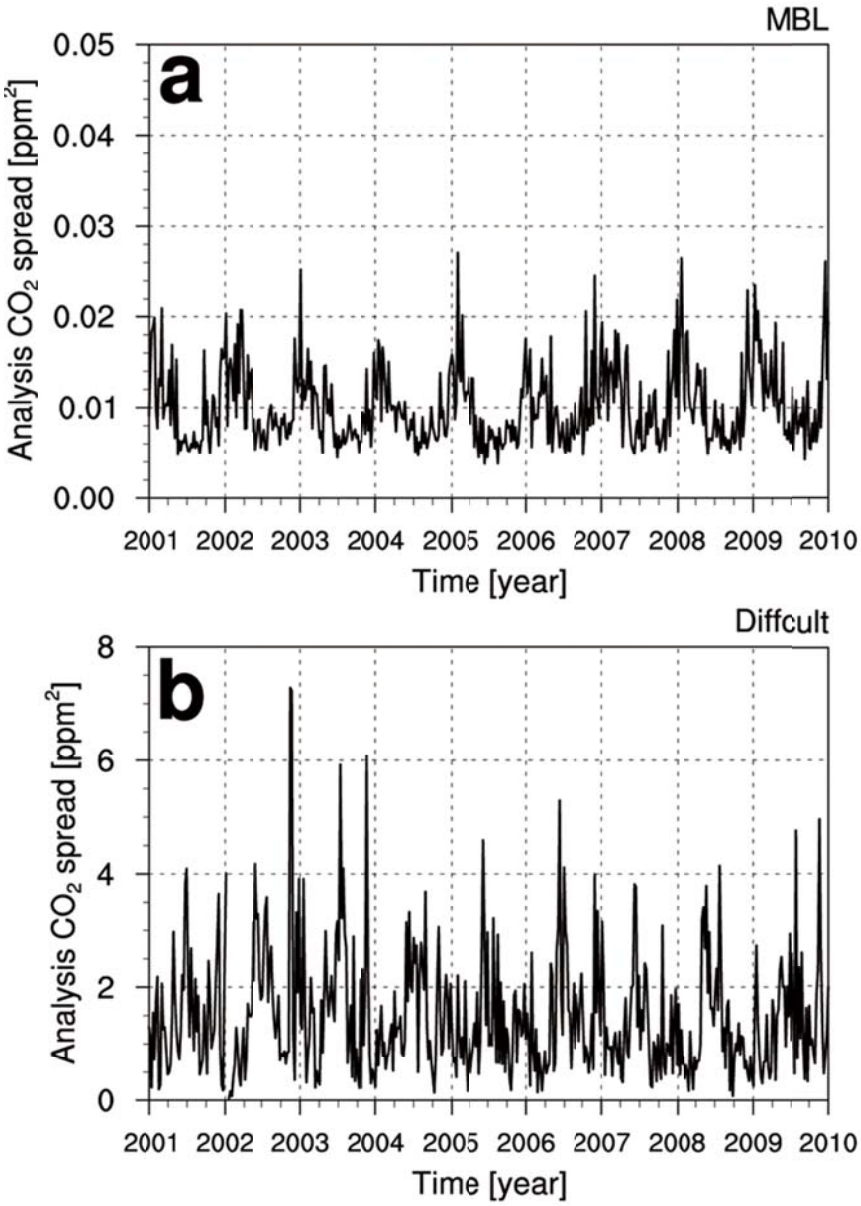


Figure rev_1. The spread of analysis CO_2 concentration [ppm^2] for (a) MBL and (b) Difficult category.

3) Does the calculation of the influence matrix take into account systematic errors in the observations? Or are the authors implicitly assuming that the observations do not have systematic errors? If so, why is this a valid assumption?

Author's response: Including Cardinali et al. (2004) and Liu et al. (2009), no previous observation impact studies have taken into account systematic errors in the observations. The observations used in this study are the same as the observations used in previous studies on CarbonTracker. As shown in Table 1, these observations are collected and managed by NOAA-ESRL. Therefore, we have considered that the quality of the observations is good. These observations are disseminated as ObsPACK (Observation Package; http:://www.esrl.noaa.gov/gmd/ccgg/obspack) after CT2011_oi which is the current release of CarbonTracker.

4) Section 2.3 – This section mirrors Section 2 in Liu et al. [2009] very closely. But it skips an important assumption, i.e., Equations 16 and 17 assume that observation errors are not correlated. This needs to be added in the text.

Author's response: Following the reviewer's suggestion, we have revised the texts to read, "More specifically, if the observation errors are not correlated, the diagonal elements of the influence matrix (i.e., self-sensitivity) are calculated as~".

5) Section 3.1 is called 'validation' but it is unclear what is being 'validated' in this sections. Liu et al. [2009] had a similar section titled 'validation' but in that study different data-denial experiments were proposed. Have the authors considered data-denial experiments to better demonstrate the applicability/utility of this influence matrix approach for the carbon flux estimation problem? The authors should show some sensitivity experiments using the data-denial approach, especially to bring out the value of MBL vs Difficult sites.

Author's response: We agree with the reviewer's opinion on the unclear title of previous Section 3.1. Therefore we have removed the previous Section 3.1 and have moved the content in the previous Section 3.1 to the first paragraph in new Section 3.1.1. While Liu et al. (2009) used an ideal model (Lorenz 40-variable model) to perform data-denial experiments, our study applied the influence matrix analysis in the real carbon data assimilation using CarbonTracker and real CO_2 observations, as the reviewer has indicated. The computational cost of this study is much expensive compared with Liu et al. (2009). Therefore, we think that this work is the first step to diagnose the impact of specific CO_2 observations to the estimated CO_2 flux. The data-denial experiments are out of scope of this study and would be considered in the future. In addition, the value of MBL and Difficult sites are already shown in the above response to the reviewer question 2).

6) Section 3.1 – Why do the authors claim that the self-sensitivity in EnKF should have a value less than one? Can the authors justify this statement? Further, Lines 15-18 need to be rephrased as it currently gives the impression that when the analysis error covariance in

4DVAR is calculated using the inverse of the Hessian matrix of the cost function, then this being an approximate method will result in self-sensitivity values greater than one.

Author's response: Cardinali et al. (2004) demonstrated that the self-sensitivity is theoretically between 0 and 1 if observations are not correlated. Liu et al. (2009) also mentioned that the calculation of the self-sensitivity requires no approximations when the observation errors are not correlated, so that the self-sensitivity satisfies the theoretical limits between 0 and 1. Even though there is the theoretical limit, the calculation of the analysis error covariance in 4D-VAR can introduce spurious values larger than 1 because it is based on a truncated eigenvector expansion with the vectors obtained through the Lanczos algorithm, as denoted by Cardinali et al. (2004). Therefore, we have revised the sentences of previous Section 3.1 as follows. In the revised manuscript, the sentences are in Section 3.1.1.

"Cardinali et al. (2004) demonstrated that the self-sensitivity is theoretically between 0 and 1 if observations are not correlated. In 4D-VAR, Cardinali et al. (2004) denoted that analysis error covariance based on the Hessian representation with truncated eigenvector expansion can introduce the self-sensitivities greater than 1 for only a small percentage of the cases. In contrast, the self-sensitivity in EnKF theoretically has a value lesser than 1 (Liu et al. 2009). Nevertheless, the self-sensitivity in this study shows a value greater than one because the sparse observations cause insufficient reduction of the background and observation operator used has nonlinearity in calculating the transport of CO_2 concentrations. In this study, 13 observations from the total of 76,801 observations used for data assimilation present a value greater than one. This is only 0.02% of the total number of observations, which implies that the calculated self-sensitivity is generally valid."

7) My biggest disappointment is that the quality of the optimized CO_2 fluxes has not been assessed. Some robust ways of evaluating the posterior CO_2 fluxes (i.e., comparison to biosphere model output, comparison of posterior CO_2 concentrations to independent datasets like aircraft observations etc.) would have been beneficial for the reader. Only the uncertainty reductions are presented in Figure 7. Additionally, the color bar should be different for JJA and DJF to bring out the uncertainty reductions for DJF. The same recommendation applies for Figure 12.

Author's response: Kim et al. (2012) already showed the high quality of the optimized CO_2 flux compared with independent aircraft observations using CarbonTracker. To avoid redundancy, we have not shown the quality of the optimized CO_2 flux.

However, following the reviewer's suggestion, we have compared the optimized CO_2 flux of this study with that of other previous studies (Tables rev_1 and rev_2). Results from CT2010 (CarbonTracker 2010) used in this study, CT2013 (CarbonTracker-NOAA; http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/), CTE2013 (CarbonTracker-EU;

http://www.carbontracker.eu) (Peylin et al., 2013), and Zhang et al. (2014b) using CONTRAIL data (CarbonTracker-China; http://www.carbontracker.net) were compared. Saeki et al. (2013) used NIES-TM transport model and NOAA observations and JR-STATION data. Niwa et al. (2012) used NICAM-TM transport model and GLOBALVIEW-CO2 observations and CONTRAIL data. In contrast, Pan et al. (2011) used bottom-up method to estimate global net forest sink from forest inventory data and long term ecosystem studies.

We have compared biosphere, ocean, and biomass burning emission (fire flux in CarbonTracker) except fossil fuel emission because biosphere, ocean, and biomass burning emission were used in previous studies. Because the study period of each study is different, we have shown two tables which are for a longer period (Table rev_1) and for a shorter period (Table rev_2). As shown in Tables, the optimized CO_2 flux of this study is similar to those of other previous studies in the globe, land, and ocean. Therefore, we think the quality of the optimized CO_2 flux of this study is good enough to investigate the purpose of this study which is estimating the effect of CO_2 observations on the analysis of surface CO_2 flux in the globe.

Table rev_1. Global annual average optimized CO_2 fluxes (including biomass burning emission and without fossil fuel emission) of each study for globe, land, and ocean. Unit is P g C yr^{-1}.

	This study	CT2010	CT2013	CTE2013 Peylin et al. (2013)	Saeki et al. (2013)	Pan et al. (2011)
Period	2001-2009				2000-2009	2000-2007
Globe	-3.71	-3.68	-3.82	-3.59	-3.51	
Land	-1.59	-1.78	-1.78	-1.85	-1.9	-1.2
Ocean	-2.12	-1.9	-2.01	-1.74	-1.61	

Table rev_2. Global annual average optimized CO2 fluxes (including biomass burning emission and without fossil fuel emission) of each study for globe, land, and ocean. Unit is P g C yr^{-1}.

	This study	CT2010	CT2013	CTE2013 Peylin et al. (2013)	CT-China Zhang et al. (2014b)	Niwa et al. (2012)
Period	2006-2009				2006-2010	2006-2008
Globe	-4.49	-3.68	-4.69	-4.44	-4.5	-4.46
Land	-2.09	-1.78	-2.63	-2.52	-2.43	-2.67
Ocean	-2.4	-1.9	-2.07	-1.93	-2.08	-1.79

As mentioned in the manuscript, we set the same color bar for both JJA and DJF in Figs.

7 and 12 because we compared the seasonal and regional characteristics of the uncertainty reduction. If we use different color bar for JJA and DJF, it is difficult to compare how the uncertainty reduction and root mean square difference are different in JJA and DJF and how they different in Asia and North America. For the original purpose of two Figures, we kept the original Figures.

References

Anderson, J. I.: Spatially and temporally varying adaptive covariance inflation for ensemble filters, Tellus, 61A, 72-83, doi:10.1111/j.1600-0870.2008.00361.x, 2009.

Boweler, N. E, Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E.: The MOGREPS short-range ensemble prediction system, Q. J. R. Meteorol. Soc., 134, 703-722, 2008.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmos. Chem. Phys., 5, 2691-2702, 2005.

Cardinali, C., Pezzulli S., and Andersson E.: Influence-matrix diagnostic of a data assimilation system, Q. J. R. Meteorol. Soc., 130, 2767-2786, 2004.

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

Kim, J., Kim, H. M., and Cho, C.-H.: Application of Carbon Tracking System based on ensemble Kalman Filter on the diagnosis of Carbon Cycle in Asia, Atmosphere, 22(4), 415-447, 2012. (in Korean with English abstract)

Kim, J., Kim, H. M., and Cho, C.-H.: The effect of optimization and the nesting domain on carbon flux analyses in Asia using a carbon tracking system based on the ensemble Kalman filter, Asia-Pacific J. Atmos. Sci., 50, 327–344, doi:10.1007/s13143-014-0020-y, 2014.

Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, Q. J. R. Meteorol. Soc., 135, 523-533, 2009.

Liu, J, Kalnay, E., Miyoshi, T., and Cardinali, C.: Analysis sensitivity calculation in an ensemble Kalman filter, Q. J. R. Meteorol. Soc., 135, 1842-1851, 2009.

Miyoshi, T.: 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, 2011.

Niwa, Y., Machida, T., Sawa, Y., Matsueda, H., Schuck, T. J., Brenninkmeijer, C. A. M., Imasu, R., and Satoh, M.: Imposing strong constraints on tropical terrestrial CO_2 fluxes using passenger aircraft based measurements, J. Geophys. Res., 117, D11303, doi:10.1029/2012JD017474, 2012.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., and Hayes, D.: A Large and Persistent Carbon Sink in the World's Forests, Science, 333, 988–993, 2011.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proc. Nat. Acd. Sci. U.S.A., 104, 18925-18930, 2007.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R. Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P. Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Global Change Biol., 16, 1317-1337, doi:10.1111/j.1365-2486.2009.02078.x, 2010.

Saeki, T., Maksyutov, S., Sasakawa, M., Machida, T., Arshinov, M., Tans, P., Conway, T. J., Saito, M., Valsala, V., Oda, T., Andres, R. J., and Belikov, D., Carbon flux estimation for Siberia by inverse modeling constrained by aircraft and tower CO_2 measurements, J. Geophys. Res. Atmos., 118, 1100-1122, doi:10.1002/jgrd.50127, 2013.

Wang, X., and Bishop, C. H.: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, J. Atmos. Sci., 60, 1140-1158, 2003.

Zhang, H. F., Chen, B. Z., van der Laan-Luijkx, I. T., Chen, J., Xu, G., Yan, J. W., Zhou, L. X., Fukuyama, Y., Tans, P. P., and Peters, W.: Net terrestrial CO_2 exchange over China during 2001-2010 estimated with an ensemble data assimilation system for atmospheric CO_2, J. Geophys. Res. Atmos., 119, 3500-3515, doi:10.1002/2013JD021297, 2014a.

Zhang, H. F., Chen, B. Z. , van der Laan-Luijkx, I. T., Machida, T., Matsueda, H., Sawa, Y., Fukuyama, Y., Langenfelds, R., van der Schoot, M., Xu, G., Yan, J. W., Cheng, M. L., Zhou, X., Tans, P. P., and Peters, W.: Estimating Asian terrestrial carbon fluxes from CONTRAIL aircraft and surface CO_2 observations for the period 2006-2010, Atmos. Chem. Phys., 14, 5807-5824, doi:10.5194/acp-14-5807-2014, 2014b.

ACP-2014-337 (Editor – Ning Zeng)

Response to Reviewer 2

The authors thank the reviewer 2 for a thoughtful review of the manuscript. We agree with many of the reviewer's points and have made the necessary changes. The responses for the reviewer's specific comments are as follows.

Comment:

The manuscript has examined the contribution of CO_2 observations to the optimized CO_2 flux within Carbon Tracker EnKF assimilation system. Quantitative analysis of which observation data give more correction to the prior is indeed very interesting trial, and would give an essential feedback to the community of data providers. Especially in Carbon Tracker, it would not be very easy due to a special assimilation window, which is still remained as the non-resolved problem to the authors. However, based on the methodology introduced by Liu et al. (2009), authors have made very useful tool to estimate observation impact on the analyzed CO_2 fluxes. Although the paper was not written in a very exciting way, a revision focusing on the presentation would bring this manuscript qualified to the publication.

Specific Comment:

1) Abstract of the manuscript needs serious revision. Major reason may be because authors use several terminologies (e.g. self-sensitivity, analysis sensitivity, information content) which need their explanations or definitions, for general readers. The abstract of the manuscript contains too much detailed results that may not be appropriate for a general abstract. Thus, the current abstract does not concisely deliver what exactly you have done. This referee suggests to emphasize important findings of your research as a discerning summary.

Author's response: Following the reviewer's suggestion, we have revised the abstract. We have added definitions of some terminologies and tried to emphasize important findings.

2) Isn't there any way to estimate the cumulative impact? Any idea? As the authors pointed out, the posterior flux seems to be determined mostly by the prior flux, not by the assimilation of the observation based on the analysis of self-sensitivity. However, there is just a statement saying that the cumulative impact would be greater. Can you "prove" it? Figure 12 lets us guess roughly how the cumulative impact would be though. Still, the first week seems to give the largest correction to the prior, doesn't it? It would be quite important message for Carbon

Tracker users.

Author's response: Following the reviewer's suggestion, we have provided magnitudes for the cumulative impact of the observations in the abstract and Section 3.2.1 as follows.

"Because the surface CO_2 flux in each week is optimized by five weeks of observations, the cumulative impact over five weeks is 19.1%, much greater than 4.8%."

The cumulative impact considers the previous observation effect which is included in the previous analysis. Therefore the forecast from the previous analysis already has some percentage of observation impact. This kind of concept can also be applied to the observation impact for NWP which does not use the lagged assimilation cycle. Because the cumulative observation impact is used for the first time in this study, we have added a schematic (Fig. 2) and texts in Section 2.3 as follows.

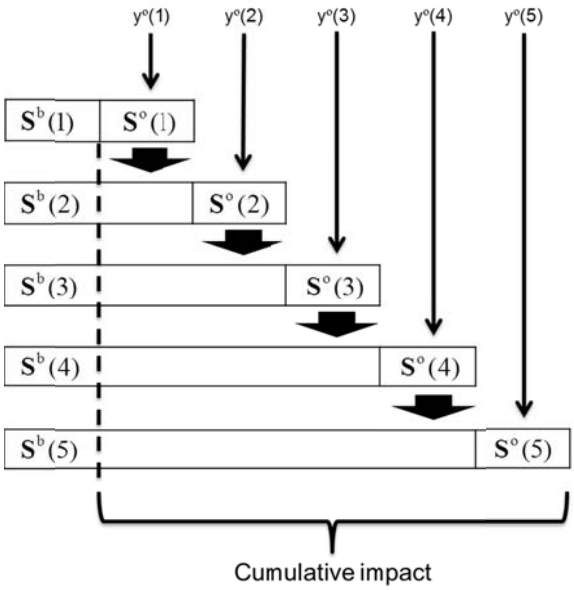


Figure 2. Schematic diagram of calculating cumulative impact in CarbonTraker. $\mathbf{S}^{b}(\bullet)$ indicates the analysis sensitivity to background at each analysis cycle within five weeks of lag, where $\bullet$ denotes each week from 1 to 5. $\mathbf{S}^{o}(\bullet)$ indicates the analysis sensitivity to observation at each analysis cycle.

"The influence matrix is calculated for the most recent week of each cycle because the background at the most recent week of each cycle is updated once by observations.

The cumulative impact of the influence matrix for the five weeks of lag can be calculated because the background in the lagged window already includes the effect from previous observations. For example, Fig. 2 shows that $\mathbf{S}^{b}(5)$ is affected by $\mathbf{S}^{o}(1)$, $\mathbf{S}^{o}(2)$, $\mathbf{S}^{o}(3)$, and $\mathbf{S}^{o}(4)$, where the number inside of parenthesis represent the week of the five-week assimilation lag. If $\mathbf{S}^{o}(\bullet)$ has a value between 0 and 1, $\mathbf{S}^{b}(1)$, the analysis sensitivity to background at the first week, represents an information from previous analysis cycle and is calculated as

$$\mathbf{S}^{b}(1) = (1-\mathbf{S}^{o}(1))(1-\mathbf{S}^{o}(2))(1-\mathbf{S}^{o}(3))(1-\mathbf{S}^{o}(4))(1-\mathbf{S}^{o}(5)), \quad (17)$$

Using Eq. (13), the cumulative impact of the influence matrix is

$$\mathbf{S}^{o}_{cum} = 1-\mathbf{S}^{b}(1) = 1-(1-\mathbf{S}^{o}(1))(1-\mathbf{S}^{o}(2))(1-\mathbf{S}^{o}(3))(1-\mathbf{S}^{o}(4))(1-\mathbf{S}^{o}(5)), \quad (18)$$

where $\mathbf{S}^{o}_{cum}$ is the cumulative impact of observations during the lagged window. The cumulative impact was defined within the five-week assimilation lag and calculated when $\mathbf{S}^{o}(5)$ exists."

3) Lines 19-22 of p.13568: What about the computational cost of this process? Are the authors doing this process at every analysis step?

Author's response: We did the process at every analysis step. The computational cost of this process is not much. Most of the computational cost in CarbonTracker is used for TM5 transport model run to calculate model CO_2 concentration. Compared to the TM5 model run, the computational cost for the analysis procedure in CarbonTracker is much smaller.

4) Equation (16) is just a case of l=j in Equation (17). Any reason to write exactly same equation twice? Unnecessary repetition makes the manuscript a little boring.

Author's response: Following the reviewer's opinion, we have deleted Eq. (17).

5) Lines 19-22 of p.13572: Isn't there any possible link with the prescribed Pb in EnKF of Carbon Tracker?

Author's response: The self-sensitivity value greater than 1 may be associated with the prescribed Pb in EnKF of CarbonTracker. However, we found that the greater self-sensitivity is more directly related with the sparse observations. We found that 7 of the

total 13 cases were occurred in Eurasian Boreal region with very sparse observations. Regardless of the prescribed Pb or non-prescribed Pb, the $\mathbf{HP}^{a}\mathbf{H}^{T}$ becomes large if there are few observations. In a different study, we found that the greater self-sensitivity in Eurasian Boreal region decreases from 7 to 2 cases when the additional observations in the region were assimilated.

The other reason of the greater self-sensitivity is associated with transport model mentioned already in the manuscript.

Therefore, we have revised the text to read, "Nevertheless, the self-sensitivity in this study shows a value greater than one because the sparse observations cause insufficient reduction of the background and observation operator used has nonlinearity in calculating the transport of CO_2 concentrations."

6) The reason why the inverse relationship between the average self-sensitivity and the number of observations is not shown was not explained. Authors said that is due to the insufficient number of observations. It does not make sense. It is just denying the statement of inverse relationship, because it is not valid when the number of observation is few. Thus, please find another reasonable reason within your experimental settings.

Author's response: As the number of observations increases, the average self-sensitivity decreases. The inverse relationship between the average self-sensitivity and the number of observations is not shown in the SH because the increase of the number of observations is not enough to cause the decrease of the average-self sensitivity. Therefore we have revised the text to read, "In the SH, an inverse relationship between the average self-sensitivity and the number of observations is not clearly shown (Fig. 6d), which is due to the insufficient increase of the number of observations assimilated in the SH compared with the other regions. However, the seasonal variability of the average self-sensitivity appears clearly in the SH. Therefore the inverse relationship is distinctly shown when the increase of the number of observations is enough to cause the decrease of the average self-sensitivity.".

7) Explaining Figure 6, authors continue to mention the inversely proportional relationship between the number of observations and self-sensitivity even though the results do not show it consistently. It seems to this referee that it is just visible in Figure 6(d), because the increase rate of the number of Continuous observations is remarkable.

Author's response: As the reviewer indicated, we have mentioned the inverse relationship only for Continuous site category when explaining the previous Fig. 6 (Fig. 7 in the revised manuscript).

8) Figure 7: it would be better to plot the reduction of self-sensitivity rather than the reduced ones.

Author's response: The self-sensitivity is calculated in observation space. Therefore the reduction of self-sensitivity is also calculated in observation space and cannot be shown as the form in previous Fig. 7 (Fig. 8 in the revised manuscript). In addition, the self-sensitivity was not calculated in every week in the 5-weeks of assimilation window. Instead of the reduction of self-sensitivity, the average standard deviations of background and posterior CO_2 fluxes in one- and five- week were shown to investigate the influence of the surface CO_2 flux uncertainties on the seasonal and regional characteristics of the self-sensitivities. In addition, the cumulative impact implies the overall observation impact during the lagged assimilation window.

9) Lines 27-28 of p.13576: "and the seasonal variability of the surface . . . variation of the self-sensitivities" seems unnecessary repetition.

Author's response: Following the reviewer's opinion, we have revised the text to read, "Therefore, the surface CO_2 flux uncertainty is one of the components to determine the magnitude and seasonal variation of the self-sensitivities."

10) Lines 8-9 of p.13577: Do the authors indicate the temporal resolution of the station? Please rephrase it.

Author's response: The number of observations at one station depends on the temporal resolution, missing rate, and total period of observations. Therefore we have revised the text to read, "Because the magnitude of the information content at one observation site is proportional to the self-sensitivity and the number of observations, the observation sites with a high average self-sensitivity or a large number of observations show high information content. The number of observations at one station depends on the temporal resolution, missing rate, and total period of observations. Therefore, the observation sites located in North America and Asia generally show high average information content."

11) Some statements are so trivial, not worth pointing out: e.g. lines 21-22 of p.13577, lines 10-12 of p.13578.

Author's response: For the lines 21-22 of p. 13577, even though they are trivial, we need a text to explain the previous Fig. 9b (Fig. 10b in the revised manuscript). Therefore we have revised the text to read, "As in the globe, the Continuous site category is the most informative in the NH (Fig. 10b)".

For the lines 10-12 of p. 13578, we have deleted the texts following the reviewer's

opinion.

12) Line 3 of p.13579: have the authors really assimilated only surface CO_2 concentration data? What's the criterion of surface layer? Carbon Tracker assimilates observations which are located up in the air either.

Author's response: To clarify the surface observations used in this study, at the first paragraph in Section 2.4, we have revised the text to read, "The observations used in this study are surface CO_2 mole fraction data observed at sites distributed around the globe (Table 1 and Fig. 3). As in Peters et al. (2007), the surface CO_2 mole fraction data used in this study includes surface air samples collected around the globe and from tall towers."

13) Lines 26-27 of p.13580: while statement needs to be rephrased.

Author's response: Following the reviewer's opinion, we have rephrased the text to read, "The self-sensitivity and spatial coverage of the observation sites are inversely correlated in the NH, whereas these factors are not apparently related in the Tropics and SH.".

14) At the end of line 20 of p.13581, it would be better to mention a possible advanced data assimilation method which allows considering high-resolution data, because Carbon Tracker may not be able to assimilate those high-resolution data (such as remote sensing data) easily with the current algorithm.

Author's response: The sentence in line 20 of p. 13581 implies that new observation sites are necessary in regions with a low spatial density of observation sites (e.g., Asia) to obtain the beneficial effect of additional observations on the surface CO_2 flux analysis in the current CarbonTracker framework. We have not mention the reviewer's suggestion because we have mentioned that the use of high-resolution data (e.g., CONTRAIL, GOSAT etc.) in CarbonTracker is the future work at the last paragraph of Section 4. In fact, we plan to assimilate the GOSAT data in CarbonTracker in the future.

15) Unit of MDM should be presented in Table 1 rather than Table 2.

Author's response: We have revised the Tables following the reviewer's opinion.

16) When explaining Figure 9 (section 3.3.1), please make sure there is no Continuous data in SH.

Author's response: At the end of previous Section 3.3.1 (Section 3.2.1 in the revised manuscript), we have added the text to read, "In addition, the information from the Continuous site category is zero because there is no Continuous data in the SH."

References

Liu, J, Kalnay, E., Miyoshi, T., and Cardinali, C.: Analysis sensitivity calculation in an ensemble Kalman filter, Q. J. R. Meteorol. Soc., 135, 1842-1851, 2009.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proc. Nat. Acd. Sci. U.S.A., 104, 18925-18930, 2007.

Influence of CO_2 observations on the optimized CO_2 flux in an ensemble Kalman filter

서식 있음

JinWoong Kim[1], Hyun Mee Kim[1] and Chun-Ho Cho[2]

[1]{Atmospheric Predictability and Data Assimilation Laboratory

Department of Atmospheric Sciences, Yonsei University, Seoul, Republic of Korea}

[2]{National Institute of Meteorological Research, Jeju, Republic of Korea}

Correspondence to: Hyun Mee Kim (khm@yonsei.ac.kr)

Abstract

~~Various data assimilation schemes have been applied in studies on atmospheric CO_2 inversion. An influence matrix based on the linear statistical analysis scheme can diagnose the impact of individual observations on a particular analysis.~~ In this study, ~~to estimate~~ the effect of CO_2 observations on an analysis of surface CO_2 flux~~, both the analysis sensitivity and the information content~~ were calculated using an~~the~~ influence matrix in the CarbonTracker, which is an inverse modeling system for estimating surface CO_2 flux based on an ensemble Kalman filter. The influence matrix represents a sensitivity of the analysis to observations. The experimental period was from January 2000 to December 2009. The diagonal element of the influence matrix (i.e., analysis sensitivity) ~~global average self-sensitivity~~ is globally 4.8% on average, which implies that the analysis extracts 4.8% of the information from the observations and 95.2% from the background each assimilation cycle. Because the surface CO_2 flux in each week is optimized by five weeks of observations, the cumulative impact over five weeks is 19.1%, much greater than~~would be greater than~~ 4.8%. The analysis ~~analysis~~ sensitivity is inversely proportional to the number of observations used in the assimilation, which is distinctly apparent in continuous observation categories with a sufficient number of observations. The time series of the globally averaged analysis ~~analysis~~ sensitivities shows seasonal variations, with greater sensitivities in summer and lower sensitivities in winter, which is attributed to the surface CO_2 flux uncertainty. The time-averaged analysis sensitivities in the Northern Hemisphere are greater than those in the

Tropics and the Southern Hemisphere. The trace of the influence matrix (i.e., information content) is a measure of the total information extracted from the observations. The information content indicates an imbalance between the observation coverage in North America and that in other regions. Approximately half of the total observational information is provided by continuous observations, mainly from North America, which indicates that continuous observations are the most informative and that comprehensive coverage of additional observations in other regions is necessary to estimate the surface CO_2 flux in these areas as accurately as in North America. ~~In addition, the uncertainty of the surface CO_2 flux in Asia, where observations are sparse, is reduced by assimilating five weeks of observations as opposed to one week of observations in North America, which indicates that a longer assimilation window with a lag may beis necessary to optimize the surface CO_2 flux in Asia.~~

1 Introduction

Atmospheric CO_2 observations can be used to quantitatively estimate the sources and sinks of surface carbon fluxes. Thus, atmospheric CO_2 inversion studies using atmospheric CO_2 observations have been conducted (Gurney et al., 2002; Ciais et al., 2010; Peylin et al., 2013). Various studies applying state-of-the-art data assimilation methods have been carried out to estimate the surface carbon cycle at global and regional scales. The methods employed for the atmospheric CO_2 inversion studies include variational data assimilation methods (Chevallier et al., 2005, 2009a, 2009b; Baker et al., 2006, 2010; Basu et al., 2013), ensemble Kalman filter (EnKF) (Peters et al., 2005, 2007, 2010; Feng et al., 2009; Miyazaki et al., 2011; Kang et al., 2011, 2012; Chatterjee et al., 2012; Kim et al, 2012, 2014), and maximum likelihood ensemble filter (Zupanski et al., 2007; Lokupitiya et al., 2008). These studies have applied the data assimilation method used in numerical weather prediction (NWP) to estimate surface CO_2 fluxes.

Recent studies on atmospheric CO_2 inversion have focused on analyzing the difference between prior and optimized surface CO_2 fluxes obtained by using new inversion methods or observations (Chevallier et al., 2009a; Basu et al., 2013), as well as the carbon cycle based on optimized surface CO_2 fluxes. By contrast, the impact of various atmospheric CO_2 observations on the estimation of surface CO_2 fluxes has rarely been studied. One method employed to evaluate the impact of observations on atmospheric CO_2 inversion is the calculation of the uncertainty reduction (Peters et al., 2005; Meirink et al., 2008; Chevallier et al., 2009b; Feng et al., 2009), which is a ratio

between the variances of the prior and posterior state vectors. A large uncertainty reduction implies that observations have a large impact on the estimation of surface CO_2 fluxes. However, the uncertainty reduction cannot measure the impact of individual observations on the estimated (i.e., analyzed) surface CO_2 fluxes. Another method for assessing the impact of observations is to calculate the information content, which is the amount of information obtained from observations (Rodgers, 2000). Engelen and Stephen (2004) calculated the information content of infrared satellite sounding observations on atmospheric CO_2 concentrations. To estimate the impact of simulated CO_2 observations on surface flux analysis, Zupanski et al. (2007) calculated the information content using the information matrix in the ensemble subspace. However, similar to the uncertainty reduction, these methods calculate the impact of all observations, rather than calculating the impact of individual observations on surface CO_2 flux analysis.

Data assimilation algorithms are fundamentally based on a linear statistical assumption (Talagrand, 1997). Both sequential and variational algorithms combine background and observation information to estimate parameters based on the linear assumption. According to the linear assumption, the influence matrix that measures the impact of individual observations on estimated parameters can be calculated in the observation space. Cardinali et al. (2004) suggested a method for calculating the influence matrix within the general data assimilation framework and applied the method to a forecast model of the European Centre for Medium Weather Forecasts (ECMWF). The diagonal elements of the influence matrix are the analysis sensitivities (i.e., self-sensitivity), which are proportional to the spread of the analysis and are inversely proportional to the predetermined observation error. The trace of the diagonal elements of the influence matrix reflects the information content, which is the amount of information extracted from observations. The influence matrix provides objective diagnostics regarding the impact of observations on the analysis and hence the performance of the data assimilation system because inaccurate observations can be identified by analyzing the observation impact (Cardinali et al., 2004). Liu et al. (2009) suggested a method for calculating self-sensitivity and cross-sensitivity (i.e., off-diagonal elements of the influence matrix) within the EnKF framework and diagnosed the relative importance of individual observations within an observation system using the idealized Lorenz-40 model and the simplified hydrostatic model.

Although Cardinali et al. (2004) and Liu et al. (2009) suggested methods for calculating the impact of individual observations on an analysis, their studies focused on NWP. Therefore, the impact of

individual observations on surface CO_2 flux analysis has not been diagnosed in a study on atmospheric CO_2 inversion using the state-of-the-art data assimilation method. Because the analysis is more important than the forecast in atmospheric CO_2 inversion, the methods suggested by Cardinali et al. (2004) and Liu et al. (2009) can be applied to diagnose the impact of observations on the CO_2 flux analysis.

CarbonTracker is a system developed by the National Oceanic and Atmospheric Administration (NOAA), which optimizes the surface CO_2 flux by assimilating mole fraction observations (i.e., concentration) of surface CO_2 (Peters et al., 2005). CarbonTracker has been applied in studies on atmospheric CO_2 inversion in North America (Peters et al., 2010), Europe (Peters et al., 2010), and Asia (Kim et al. 2014). To develop CarbonTracker for use in Asia, Kim et al. (2012) performed an experiment employing CarbonTracker in this region and demonstrated that CarbonTracker produces optimized surface CO_2 fluxes for Asia. Kim et al. (2014) showed that the estimates of the surface CO_2 flux are more consistent with observed CO_2 concentrations in Asia when using the nesting domain of the transport model on Asia in CarbonTracker. Zhang et al. (~~2013~~2014) conducted a study on the assimilation of aircraft CO_2 observations from the Comprehensive Observation Network for TRace gases by AIrLiner (CONTRAIL) (Machida et al., 2008) in Asia using CarbonTracker.

In this study, an influence matrix is calculated in CarbonTracker to evaluate the impact of mole fraction observations of CO_2 on the analyzed surface CO_2 fluxes. The relative importance of each observation site and each observation site category is evaluated by analyzing the self-sensitivity and information content, and the characteristics of the self-sensitivity and information content are subsequently investigated. Section 2 presents the experimental framework, which includes CarbonTracker, EnKF, observations, the methodology for calculating the influence matrix, and the experimental framework. Section 3 presents the results, and Section 4 provides a summary and conclusion.

2 Methodology

2.1 CarbonTracker

CarbonTracker is an atmospheric CO_2 inversion system that estimates the surface CO_2 flux consistent with CO_2 observations. In CarbonTracker, the optimized flux with a 1°×1° horizontal

resolution is calculated by

$$F(x,y,t) = \lambda_r \cdot F_{bio}(x,y,t) + \lambda_r \cdot F_{ocn}(x,y,t) + F_{ff}(x,y,t) + F_{fire}(x,y,t), \quad (1)$$

where $F_{bio}(x,y,t)$ is the prescribed prior biosphere flux from the Carnegie Ames Stanford Approach Global Fire Emissions Database (CASA GFED) version 3.1 (van der Werf et al., 2010); $F_{ocn}(x,y,t)$ is the prescribed prior ocean flux based on Jacobson et al. (2007); $F_{ff}(x,y,t)$ is the prescribed prior fossil fuel flux determined using the Carbon Dioxide Information and Analysis Center (CDIAC) and the Emission Database for Global Atmospheric Research (EDGAR) inventories; $F_{fire}(x,y,t)$ is the prescribed prior fire flux derived from CASA GFED version 3.1 ~~2~~ (van der Werf et al., 2010~~6~~)~~3.1 다시살펴보니 version 3.1이 사용되어서 수정해야할 것 같습니다. 이로 인해 해당 reference도 제외돼어야 합니다.~~; and λ_r is the scaling factor to be optimized in the data assimilation process, corresponding to 156 ecoregions around the globe. CarbonTracker adopts a smoother window to reflect the transport speed of CO_2, which is based on the temporal relationship between the surface CO_2 flux and atmospheric CO_2 observations, as found in Bruhwiler et al. (2005) (Peters et al., 2005). For this reason, the scaling factor is optimized for five weeks of lag, which implies that the observations made in the most recent week affect the optimized surface CO_2 flux in the preceding four weeks. The optimization of the scaling factor during the data assimilation process is presented in Fig. 1. In each assimilation cycle, five weeks of analysis scaling factors are estimated by observations from the most recent week. After the fifth cycle, the scaling factor estimated by these five weeks of observations is saved as the optimized scaling factor and used to calculate the optimized surface CO_2 flux in Eq. (1). During this process, a new mean background scaling factor for the next week is calculated by the estimated mean scaling factors of the previous two weeks using a simple dynamic model, as follows:

$$\lambda_t^{\mathrm{b}} = \frac{(\lambda_{t-2}^{\mathrm{a}} + \lambda_{t-1}^{\mathrm{a}} + \lambda^{\mathrm{p}})}{3}, \quad (2)$$

where λ_t^{b} is a prior mean scaling factor for the new analysis week; $\lambda_{t-2}^{\mathrm{a}}$ and $\lambda_{t-1}^{\mathrm{a}}$ are posterior mean scaling factors estimated two weeks and one week previous, respectively; and λ^{p} is a prior value fixed as 1. Thus, the information from the previous observations is included in λ_t^{b}.

The TM5 model (Krol et al., 2005) is used as a transport model that calculates model CO_2 concentrations corresponding to the observed CO_2 concentrations. The TM5 model uses the

서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
변경된 필드 코드

surface CO_2 fluxes calculated from Eq. (1) and the ECMWF meteorological field to calculate model CO_2 concentrations and is used as the observation operator, which will be explained in Section 2.2.

2.2 Ensemble Kalman Filter

The EnKF data assimilation method used in CarbonTracker is the ensemble square root filter (EnSRF) suggested by Whitaker and Hamill (2002). The analysis equation for data assimilation is expressed as

$$\mathrm{x}^{\mathrm{a}} = \mathbf{K}\mathrm{y}^{\mathrm{o}} + (\mathrm{I}_n - \mathbf{K}\mathbf{H})\mathrm{x}_{\mathrm{b}}, \tag{3}$$

where x^{a} is the n-dimensional analysis (posterior) state vector; y^{o} is the p-dimensional observation vector; $\mathbf{K}$ is the n × p dimensional Kalman gain; $\mathbf{I}_n$ is the identical matrix; $\mathbf{H}$ is the linearized observation operator, which transforms the information in the model space to the information in the observation space; and x^{b} is the background state vector. In EnSRF, the ensemble mean and perturbed state vectors are updated independently using the following equations:

$$\mathrm{x}^{\mathrm{a}} = \mathrm{x}^{\mathrm{b}} + \mathbf{K}\left(\mathrm{y}^{\mathrm{o}} - \mathbf{H}\mathrm{x}^{\mathrm{b}}\right), \tag{4}$$

$$\mathrm{x}_i'^{\mathrm{a}} = \mathrm{x}_i'^{\mathrm{b}} - \tilde{\mathbf{k}}\mathbf{H}\mathrm{x}_i'^{\mathrm{b}}, \tag{5}$$

where x^{a} and x^{b} are mean state vectors of the analysis and background, respectively, and $\mathrm{x}_i'^{\mathrm{a}}$ and $\mathrm{x}_t'^{\mathrm{b}}$ are perturbation state vectors of the analysis and background, respectively. Many inflation techniques (e.g., Wang and Bishop, 2003; Bowler et al., 2008; Whitaker et al., 2008; Li et al., 2009; Anderson, 2009; Miyoshi, 2011; Kang et al., 2012~~의중 reference에 없던 것을 추가하였습니다~~) have been used to maintain proper ensemble spread and to improve the performance of EnKF data assimilation. Although the EnSRF in CarbonTracker does not use the inflation method, Kim et al. (2012) demonstrated that the ensemble spread measured by rank histograms is maintained properly. In CarbonTracker, the state vector corresponds to the scaling factor, as described in Section 2.1. $\mathbf{K}$ and the reduced Kalman gain, $\tilde{\mathbf{k}}$, are defined as

서식 있음: 강조

$$\mathbf{K} = \left(\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}\right)\left(\mathbf{H}\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1}, \tag{6}$$

$$\tilde{\mathbf{k}} = \mathbf{K} \cdot \alpha, \tag{7}$$

where $\mathbf{P}^{\mathrm{b}}$ is the background error covariance; $\mathbf{R}$ is the observation error covariance, which is predefined at each observation site; and α is a scalar value that varies whenever each observation is used in the analysis process and is calculated as

$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}, \tag{8}$$

$\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ and $\mathbf{H}\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ in Eqs. (6) and (8) can be calculated as

$$\mathbf{P}\mathbf{H}^{\mathrm{T}} \approx \frac{1}{m-1}\left(\mathrm{x}_1', \mathrm{x}_2', \ldots, \mathrm{x}_m'\right) \cdot \left(\mathbf{H}\mathrm{x}_1', \mathbf{H}\mathrm{x}_2', \ldots, \mathbf{H}\mathrm{x}_m'\right)^{\mathrm{T}}, \tag{9}$$

$$\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} \approx \frac{1}{m-1}\left(\mathbf{H}\mathrm{x}_1', \mathbf{H}\mathrm{x}_2', \ldots, \mathbf{H}\mathrm{x}_m'\right) \cdot \left(\mathbf{H}\mathrm{x}_1', \mathbf{H}\mathrm{x}_2', \ldots, \mathbf{H}\mathrm{x}_m'\right)^{\mathrm{T}}, \tag{10}$$

where m is the number of ensembles.

To reduce the sampling error and filter divergence due to the underestimation of background error covariance in EnSRF, the covariance localization method is used (Houtekamer and Mitchell, 2001). Because the physical distance between the scaling factors cannot be defined in CarbonTracker, correlations between the ensemble of the scaling factor and the ensemble of the model CO_2 concentration are calculated, and a statistical significance test is performed on the correlations. Then, the Kalman gain which has an insignificant statistical value is set to zero. This type of localization is applied to all observation sites except for Marine Boundary Layer (MBL) sites, because the observations at MBL sites are considered to include information on large footprints of flux signals (Peters et al., 2007).

2.3 Influence matrix

The influence matrix for EnKF is calculated as in Liu et al. (2009). The projection of Eq. (1→3) onto the observation space becomes

$$\mathbf{H}\mathrm{x}^{\mathrm{a}} = \mathrm{y}^{\mathrm{a}} = \mathbf{H}\mathbf{K}\mathrm{y}^{\mathrm{o}} + (\mathbf{I}_p - \mathbf{H}\mathbf{K})\mathrm{y}^{\mathrm{b}}, \tag{11}$$

where y^{a} is the analysis value in the observation space and the projection of the state vector x^{a}

on the observation space. The influence matrix $\mathbf{S}^{o}$, representing the sensitivity of the analysis state vector y^{a} to the observation vector y^{o} (i.e., analysis sensitivity to observation) in the observation space, is calculated as follows:

$$\mathbf{S}^{o}=\frac{\partial \mathbf{y}^{a}}{\partial \mathbf{y}^{o}}=\mathbf{K}^{T}\mathbf{H}^{T}=\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^{a}\mathbf{H}^{T}, \quad (12)$$

where $\mathbf{S}^{o}$ is proportional to the analysis error covariance and is inversely proportional to the observation error covariance. By contrast, the analysis sensitivity to background is

$$\mathbf{S}^{b}=\frac{\partial \mathbf{y}^{a}}{\partial \mathbf{y}^{b}}=\frac{\partial \mathbf{y}^{a}}{\partial (\mathbf{H}\mathrm{x}^{b})}=\mathbf{I}_{p}-\mathbf{K}^{T}\mathbf{H}^{T}=\mathbf{I}_{p}-\mathbf{S}^{o}, \quad (13)$$

where y^{b} is the projection of the background on the observation space, and $\mathbf{I}_{p}$ is an identity matrix with the size of the number of observations. Consequently, the sum of the analysis sensitivity to observation in Eq. (12) and the analysis sensitivity to background in Eq. (13) is one.

Substituting Eq. (10) into Eq. (12) becomes

$$\mathbf{S}^{o}=\mathbf{R}^{-1}\mathbf{H}\mathbf{P}^{a}\mathbf{H}^{T}=\frac{1}{m-1}\mathbf{R}^{-1}(\mathbf{H}\mathbf{X}^{a})(\mathbf{H}\mathbf{X}^{a})^{T}, \quad (14)$$

where $\mathbf{HX}^{a}$ is the analysis ensemble perturbation matrix in the observation space, and the i^{th} column of $\mathbf{HX}^{a}$ is calculated as

$$\mathbf{HX}_{i}^{a}\cong h\left(\mathrm{x}_{i}^{a}\right)-\frac{1}{m}\sum_{i=1}^{m}h\left(\mathrm{x}_{i}^{a}\right), \quad (15)$$

where x_{i}^{a} is the i^{th} analysis ensemble member; m is the number of ensembles (i.e., 150); and $h(\cdot)$ is the linear or nonlinear observation operator. More specifically, if the observation errors are not correlated, the diagonal elements of the influence matrix (i.e., self-sensitivity) are calculated as

$$\mathbf{S}_{jj}^{o}=\frac{\partial \mathbf{y}_{j}^{a}}{\partial \mathbf{y}_{j}^{o}}=\left(\frac{1}{m-1}\right)\frac{1}{\sigma_{j}^{2}}\sum_{i=1}^{m}(\mathbf{HX}_{i}^{a})_{j}\times(\mathbf{HX}_{i}^{a})_{j},$$

(16)

where σ_j^2 is the error variance of the j^{th} observation. The ~~and the~~ cross-sensitivity~~, which~~ is the off-diagonal elements of the influence matrix. ~~, is calculated as~~

~~$$\mathbf{S}^{o}_{jl} = \frac{\partial \mathbf{y}^{a}_{l}}{\partial \mathbf{y}^{o}_{j}} = \left(\frac{1}{m-1}\right)\frac{1}{\sigma_j^2}\sum_{i=1}^{m}(\mathbf{HX}^{a}_{i})_j \times (\mathbf{HX}^{a}_{i})_l \,,$$ (17)~~

~~where σ_j^2 is the error variance of the j^{th} observation.~~ The influence matrix is calculated for the most recent week of each cycle because the background at the most recent week of each cycle is updated once by observations.

The cumulative impact of the influence matrix for the five weeks of lag can be calculated because the background in the lagged window already includes the effect from previous observations. For example, Fig. 2 shows that $\mathbf{S}^{b}(5)$ is affected by $\mathbf{S}^{o}(1)$, $\mathbf{S}^{o}(2)$, $\mathbf{S}^{o}(3)$, and $\mathbf{S}^{o}(4)$, where the number inside of parenthesis represent the week of the five-week assimilation lag. If $\mathbf{S}^{o}(\bullet)$ has a value between 0 and 1, $\mathbf{S}^{b}(1)$, the analysis sensitivity to background at the first week, represents an information from previous analysis cycle and is calculated as

$$\mathbf{S}^{b}(1) = (1-\mathbf{S}^{o}(1))(1-\mathbf{S}^{o}(2))(1-\mathbf{S}^{o}(3))(1-\mathbf{S}^{o}(4))(1-\mathbf{S}^{o}(5))$$

(178)

Using Eq. (13), the cumulative impact of the influence matrix is

$$\mathbf{S}^{o}_{cum} = 1-\mathbf{S}^{b}(1) = 1-(1-\mathbf{S}^{o}(1))(1-\mathbf{S}^{o}(2))(1-\mathbf{S}^{o}(3))(1-\mathbf{S}^{o}(4))(1-\mathbf{S}^{o}(5)),$$ (189)

where $\mathbf{S}^{o}_{cum}$ is the cumulative impact of observations during the lagged window. The cumulative impact was defined within the five-week assimilation lag and calculated when $\mathbf{S}^{o}(5)$ exists.~~CarbonTracker에서는 optimized surface CO_2 flux를 계산하기 위해 five week assimilation lag을 사용하기 때문에 여러 weeks의 관측 자료에 대한 cumulative impact를 계산할 수 있다. Fig. 2는 cumulative impact를 정의하는 schematic 이다. 이 때 cumulative impact는 five week assimilation lag 안에서 정의된다. 이 값을 계산하기 위해 다음과 같은 방법을 이용하여~~

서식 있음: 글꼴: Times New Roman, 12 pt, 글꼴 색: 검정

서식 있음: 글꼴: Times New Roman, 12 pt, 글꼴 색: 검정

서식 있음: 글꼴: Times New Roman, 12 pt, 글꼴 색: 검정

서식 있음: 들여쓰기: 첫 줄: 0 글자, 간격 앞: 6 pt, 단락 뒤: 0 pt, 줄 간격: 1.5줄

서식 있음: 글꼴 색: 검정

서식 있음: 글꼴: Times New Roman, 12 pt

서식 있음: 글꼴: Times New Roman, 12 pt

~~cumulative impact를 계산하였다. 먼저 each assimilation cycle에서 계산된 self-sensitivity는 계산의 편의성을 위해 평균 값을 사용하였다. 만약 $A(t)$가 0에서 1사이의 값을 갖는다면 five-week assimilation lag 이전의 previous analysis cycles로부터 전달된 정보를 나타내는 $B(1)$은 다음과 같이 계산된다.~~

~~$$B(1) = (1 - A(1))(1 - A(2))(1 - A(3))(1 - A(4))(1 - A(5)), \quad (18)$$~~

~~여기서 $A(t)$ 는 five-week assimilation lag안의 each analysis cycle에서 계산된 analysis sensitivity to observation을 나타낸다. 예를 들어, $A(5)$는 the most recent week의 average self-sensitivity를 나타낸다. 이때 eq. (13)을 이용하면 cumulative impact는 다음과 같이 계산할 수 있다.~~

~~$$A_c = 1 - (1 - A(1))(1 - A(2))(1 - A(3))(1 - A(4))(1 - A(5)), \quad (19)$$~~

~~여기서 A_c는 cumulative impact를 나타내며, 이 값은 A(5)가 존재할 때만 (i.e most recent week에서 관측 자료가 존재할 때만) 계산된다. 이를 통해 각 관측 지점에 대해 cumulative impact를 계산할 수 있다.~~

The information content (i.e., degrees of freedom for signal), which is a measure of the information extracted from the observations, is calculated by the trace of the influence matrix. As suggested by Cardinali et al. (2004), the globally averaged influence of the observations can be calculated by averaging the global self-sensitivities as

$$GAI = \frac{\mathrm{tr}(\mathbf{S}^{o})}{p}, \quad (19\text{~~18~~})$$

where p is the total number of observations used in each assimilation cycle. The partial influence of a subset of observations is calculated as

$$PAI = \frac{\sum_{i \in I} S_{ii}^{o}}{p_I}, \quad (20\text{~~19~~})$$

where p_I represents the number of observations in subset I, which can either be set as specific

observation types or as specific vertical and horizontal domains.

2.4 Observations

The observations used in this study are surface CO_2 mole fraction data observed at sites distributed around the globe (Table 1 and Fig. 32). As in Peters et al. (2007), the surface CO_2 mole fraction data used in this study includes

sSurface air samples collected around the globe and from tall towers.

서식 있음: 강조 없음

These data were observed by NOAA, the Commonwealth Scientific and Industrial Research Organization (CSIRO), Environment Canada (EC), the National Center for Atmospheric Research (NCAR), and Lawrence Berkeley National Laboratory (LBNL) (Masarie et al., 2011). Observations from three additional sites made by the Japan Meteorological Agency (JMA) are also used in this study. The site categories and model-data mismatch values (i.e., observation error) are shown in Table 2. The model-data mismatch is determined as the innovation χ^2 in Eq. (2120) becomes one at each observation site (Peters et al., 2007).

$$\chi^2 = \frac{(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b)^2}{\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}}, \quad (210)$$

The innovation χ^2 statistics for each observation site in Asia during the experimental period are presented in Table 3. The model-data mismatch for the TAP site (Tae-ahn peninsula, South Korea; 36.73ºN, 126.13ºE, 20 m) was changed from the value of 7.5 ppm used in previous studies to 5 ppm because the innovation χ^2 value obtained using 5 ppm was closer to one. However, TAP was still included in the Difficult category in the statistical analysis in Section 3. The model-data mismatches of the three JMA sites were set to 3 ppm, as in Zhang et al. (2014).

2.5 Experimental framework

The surface carbon flux analysis system used in this study is based on the CarbonTracker 2010 release (CT2010). However, the system employed in this study is different from CT2010 in two aspects: first, the nesting domain of the TM5 model, with 1°×1° horizontal resolution, is centered in Asia rather than in North America, which enables a more detailed analysis of the surface CO_2 fluxes over Asia, as shown in Kim et al. (2014); second, as mentioned in Section 2.4, three new

JMA observation sites are added in this system, which also enhances the analysis of surface CO_2 fluxes over Asia. The global horizontal resolution is 3°×2°, as in CT2010. The experimental period is from 1 January 2000 to 31 Dec 2009. The number of ensembles is 150, and the scaling factor includes five weeks of lag, as in Peters et al. (2007, 2010) and Kim et al. (2012, 2014).

3 Results

~~3.1 Validation~~

~~Cardinali et al. (2004) showed that the influence matrix is calculated approximately in the four-dimensional variational data assimilation method (4DVAR) because the analysis error covariance in 4DVAR is numerically calculated by the inverse of the Hessian matrix of the cost function. If the analysis error covariance is not calculated appropriately, the self-sensitivity can show a value greater than one. In contrast, the self-sensitivity in EnKF theoretically has a value lesser than one. Nevertheless, the self-sensitivity in this study can have a value greater than one because the observation operator used has nonlinearity in calculating the transport of CO_2 concentrations. In this study, only 13 observations from the total of 76,692 observations used for data assimilation present a value greater than one. This is only 0.02% of the total number of observations, which implies that the calculated self-sensitivity is generally valid.~~

~~3.2~~3.1 Self-sensitivity

~~3.2.1~~3.1.1 Average self-sensitivity

Cardinali et al. (2004) demonstrated that the self-sensitivity is theoretically between 0 and 1 if observations are not correlated. In 4D-VAR, Cardinali et al. (2004) denoted that analysis error covariance based on the Hessian representation with truncated eigenvector expansion can introduce the self-sensitivities greater than 1 for only a small percentage of the cases. In contrast, the self-sensitivity in EnKF theoretically has a value lesser than 1~~one~~ (Liu et al. 2009). Nevertheless, the self-sensitivity in this study shows a value greater than one because the sparse observations cause insufficient reduction of the background and observation operator used has nonlinearity in calculating the transport of CO_2 concentrations. In this study, ~~only~~ 13 observations from the total of 76,801➔~~다시 살펴보니 76,801개가 맞는 것으로 확인되어 수정해야 할 듯 합니다. Fig. 5와 10 에서는 2000년 관측수가 제외되었기 때문에 본문과는 다른 숫자를 나타내게 됩니다.~~ observations used for data

서식 있음: 표준
서식 있음: 글꼴: (영어) Times New Roman, 12 pt, 글꼴 색: 검정
서식 있음: 글꼴: (영어) Times New Roman, 12 pt, 글꼴 색: 검정
서식 있음: 글꼴: (영어) Times New Roman, 12 pt, 글꼴 색: 검정
서식 있음: 글꼴: (영어) Times New Roman, 12 pt, 글꼴 색: 검정
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음

assimilation present a value greater than one. This is only 0.02% of the total number of observations, which implies that the calculated self-sensitivity is generally valid.

Because the spatial coverage and number of observations varies during the experimental period, the average self-sensitivity throughout the experimental period was analyzed to evaluate the overall characteristics of the self-sensitivity at each observation site. As in previous studies (e.g., Peters et al., 2007, 2010; Kim et al., 2014), the results for the year 2000 were excluded from the data analysis because 2000 is considered as the spin-up period.

Figure 4~~3~~ shows the average self-sensitivities at each observation site during the experimental period. Different sizes of circles are used in some locations to distinguish sites at the same location or at geographically close locations. In the globe, negative correlations between the spatial density of the observation sites and the self-sensitivities are not as apparent as those reported by Cardinali et al. (2004) and Liu et al. (2009). Negative correlations between the spatial density of the observation sites and the self-sensitivities are apparent in the Northern Hemisphere (NH). In particular, some observation sites in Asia show high sensitivities and a low spatial density of observation sites. The observation sites located in deserts, remote oceans, and high altitude regions generally exhibit low sensitivities.

The average self-sensitivities of each observation site category over the globe, in the NH, Tropics, and Southern Hemisphere (SH) are shown in Fig. 5~~4~~. The average global self-sensitivity is 4.8% (Fig. 5~~4~~a), which implies that the analysis extracts 4.8% of its information from the observations and 95.2% from the background each assimilation cycle. Although the average self-sensitivity seems low, the background contains the observation information included in the previous analysis cycle, as reported in Cardinali et al. (2004). Moreover, the surface CO_2 fluxes in CarbonTracker are optimized by five weeks of observations during the assimilation process. Therefore, the cumulative impact over five weeks is 19.1% much ~~would be~~ greater than 4.8%, which only represents the most recent week of each~~the~~ cycle. Although the cumulative impact shows a higher value, the non-cumulative impact measured in the most recent week of each cycle is used to discuss the impact of observations because the non-cumulative impact has been generally used as the observation impact.

In the globe, the Mixed site category shows the highest average self-sensitivity, and the Difficult site category shows the lowest average self-sensitivity (Fig. 5~~4~~a), which is related to the model-

data mismatch values shown in Table 1. The model-data mismatch for the Mixed site category is relatively low, while that of the Difficult site category is high. Although the MBL site category has the lowest model data mismatch, the MBL site category does not show the highest average self-sensitivity due to the small spread of the analysis CO_2 concentrations at MBL sites. As shown in Eq. (167~~7~~), the model-data mismatch and the spread of the analysis CO_2 concentrations are two factors determining the self-sensitivity. Because MBL sites are located far from strong source and sink regions of CO_2, the spread of the analysis CO_2 concentrations at these sites is small. The average self-sensitivity in the NH is 5.3%, which is the highest of all global regions (Fig. 5~~4~~b). Similar to the global results, the average self-sensitivity is highest for the Mixed site category, while that for the Difficult site category is lowest. The average self-sensitivity in the Tropics is 3.6% (Fig. 5~~4~~c); the Mixed site category shows the highest values, but they are not significantly higher than those of other categories. In the Tropics, there is no Continuous site category. In the SH, the average self-sensitivity is 3.0%, which is the lowest among the global regions (Fig. 5~~4~~d); the MBL site category shows the highest values, and there is no Continuous site category.

~~3.2.2~~3.1.2 Time series of self-sensitivity

Figure 6~~5~~ shows the time series of the average self-sensitivity and number of observations around the globe and in each region. Globally, two apparent characteristics can be identified in the time series (Fig. 6~~5~~a): first, the average self-sensitivity decreases as the number of observations increases, showing an inversely proportional relationship; second, there is seasonal variability in the average self-sensitivity, showing high values in summer and low values in winter. In the NH, the above two features are more apparent than in other regions (Fig. 6~~5~~b). Because most of the observation sites are located in the NH, characteristics of the average global self-sensitivity are mostly determined by those in the NH. As the number of observations in the Tropics increases in the late 2000s, a slight inversely proportional relationship between the average self-sensitivity and the number of observations appears in the Tropics (Fig. 6~~5~~c). However, the average self-sensitivity in the Tropics does not show distinct seasonal variability. In the SH, an inverse relationship between the average self-sensitivity and the number of observations is not clearly shown (Fig. 6~~5~~d), which is due to the insufficient increase of the number of observations assimilated in the SH compared with the other regions. However, the seasonal variability of the average self-sensitivity appears clearly in the SH. Therefore the inverse relationship is distinctly shown when the increase of the number of observations is enough to cause the decrease of the average self-sensitivity.

Figure 7~~6~~ shows the average self-sensitivity for each observation site category. Although the MBL site category has the second largest number of observations, the average self-sensitivity shows little variation with respect to time (Fig. 7~~6~~a). Similarly, the average self-sensitivity for the Continental site category does not vary with respect to time (Fig. 7~~6~~b). The average self-sensitivity of the Mixed site category shows distinct seasonal variation (Fig. 7~~6~~c). The Continuous site category displays distinct seasonal variability in the average self-sensitivity and an inversely proportional relationship between the average self-sensitivity and the number of observations (Fig. 7~~6~~d). Because Continuous sites are mostly located in North America with relatively large numbers (Fig. 3~~2~~), the impact of a single observation decreases as the number of observations increases. Therefore, the inversely proportional relationships between the average self-sensitivity and the number of observations around the globe (Fig. 6~~5~~a) and in the NH (Fig. 6~~5~~b) are mainly attributed to the Continuous site category. The Difficult site category shows a slight inverse relationship between the average self-sensitivity and the number of observations (Fig. 7~~6~~e).

~~3.2.3~~3.1.3 Effect of the ensemble spread of the model surface CO_2 flux on the average self-sensitivity

Despite the inversely proportional relationship between the self-sensitivity and the number of observations in the NH time series (Fig. 6~~5~~a), the average self-sensitivity in the NH is higher than in the other regions (Fig. 5~~4~~). In addition, the average self-sensitivities in the NH and SH are greater in summer than in winter (Fig. 6~~5~~). The above two characteristics imply that another factor in addition to the number of observations affects the self-sensitivity. As briefly mentioned in Section 3.2.1, another factor that affects the self-sensitivity is the spread of the analysis CO_2 concentrations. Therefore, the average standard deviations of surface CO_2 fluxes are evaluated in Fig. 8~~7~~ to investigate the influence of the surface CO_2 flux uncertainties on the seasonal and regional characteristics of the self-sensitivities. The ensemble spread of the background surface CO_2 fluxes reflects the uncertainties, which are projected onto the ensemble spread of the background and analysis CO_2 concentrations (i.e., $\mathbf{HX}^a$ in Eq. (16)) by the transport model. The uncertainties of the background surface CO_2 fluxes over the terrestrial portion of the NH are high in summer months (i.e., June, July, and August: JJA) (Fig. 8~~7~~a) compared with those in winter months (i.e., December, January, and February: DJF) (Fig. 8~~7~~b). Due to the high surface CO_2 flux uncertainties in North America (Fig. 8~~7~~a), the self-sensitivities in North America are not lower than those in the other regions (Fig. 4~~3~~), regardless of the large number of observations in this region.

By contrast, despite the high uncertainties of the surface CO_2 fluxes in the Eurasian Boreal region, the self-sensitivities in this region cannot be evaluated owing to the absence of observations. Instead, the self-sensitivities of the observation sites near the Eurasian Boreal region show high values (Fig. 43).

The uncertainties of the optimized biosphere and ocean fluxes by one week of observations, shown in Figs. 87c and d, are reduced compared with those of the background fluxes, shown in Figs. 87a and b. The magnitude of the reduction of the surface CO_2 flux uncertainties in North America is relatively greater than in other regions, which is consistent with the greater self-sensitivities found in North America. By contrast, when using five weeks of observations, the magnitude of the reduction of the surface CO_2 flux uncertainties is greater in Asia than in North America (Figs. 87e and f).

Therefore, the surface CO_2 flux uncertainty is one of the components to determine the magnitude and seasonal variation of the self-sensitivities~~, and the seasonal variability of the surface CO_2 flux uncertainties leads to the seasonal variation of the self-sensitivities~~.

~~3.3~~3.2 Information content

~~3.3.1~~3.2.1 Average information content

Figure 98 shows the average information content at each observation site during the experimental period. This value was calculated by averaging the ratio of information contents for each cycle at each site during the experimental period. Note that this average value is not the amount of information content extracted from observations but rather the relative ratio of each site's information content, normalized by the total influence of all observations. Because the magnitude of the information content at one observation site is proportional to the self-sensitivity and the number of observations, the observation sites with a high average self-sensitivity or a large number of observations show high information content. The number of observations at one station depends on the temporal resolution, missing rate, and total period of observations. Therefore, the observation sites located in North America and Asia generally show high average information content.

To investigate the distribution of the information content in each region, histograms of the average information content around the globe and in the NH, Tropics, and SH were generated

(Fig. 109). The average information content was 80.2% in the NH, 13.3% in the Tropics, and 6.5% in the SH, which implies that the observations in the NH are the most informative. This is due to the large number of observations and high self-sensitivities in the NH. Around the globe, the most informative observation site category is the Continuous category (Fig. 109a). The MBL, Continental, and Mixed site categories show a similar magnitude of information content, but the Difficult site category shows the lowest information content. As in the ~~NH and for the~~ globe ~~as a whole~~, the Continuous site category is the most informative in the NH (Fig. 109b). In the current CarbonTracker system, the observation sites of the Continuous site category are mainly located in North America, except for the three JMA sites, which are located in Asia. Therefore, most of the information extracted from the Continuous site category is used to constrain the surface CO_2 fluxes of North America. In the Tropics, the MBL and Mixed site categories provide the most information (Fig. 109c). In the SH, the MBL site category provides the most information, but information extracted from the Continental, Mixed, and Difficult site categories is rare (Fig. 109d). In addition, the information from the Continuous site category is zero because there is no Continuous data in the SH.

~~3.3.2~~3.2.2 Time series of information content

Figure 110 shows the time series of the weekly averaged information content for each site category in each region. In the globe, the proportion of the information content of the Continuous site category increases steadily over time (Fig. 110a), which is associated with the steady increase in the number of observations of the Continuous site category over time. In the NH, the increase of the proportion of the information content and the number of observations of the Continuous site category is more readily apparent (Fig. 110b). ~~Because of the high self-sensitivity in summer in the NH, the proportion of the information content of the Continuous category in the NH is greater in summer than in winter.~~ In the Tropics, the MBL and Mixed site categories provide the most information, while the Difficult site category provides limited information from 2004 onward (Fig. 110c) because, after this date, observations from only one Difficult observation site (Bukit Kotobang (BKT), Indonesia: 0.2 ºS, 100.32 ºE, 864 m) are used in the data assimilation. In the SH, most information is extracted from observations made in the MBL site category (Fig. 110d). Because the number of observations in the SH is much lower than in the other regions, the information content extracted from the observations made in this region is also lower. The information content in summer is greater than in winter in the SH owing to the

seasonal variability in self-sensitivity.

To investigate the regional distribution of the information content in the NH, the time series of the information contents in Asia, North America, and Europe are shown in Fig. 12~~1~~. The information content in North America is greater than that in the other regions because the self-sensitivities are high and the number of observations increases with time in North America. However, the rate of increase in the information content is lower than that of the number of observations because self-sensitivity decreases as the number of observations increases in North America.

~~3.3.3~~3.2.3 Relationship between the information content and the optimized flux

Because CarbonTracker is a system that optimizes the surface CO_2 flux using measurements of surface CO_2 concentrations, the effect of the observations on the optimized surface CO_2 fluxes is important. To investigate the relationship between the information content and the optimized surface CO_2 fluxes, the root mean square differences (RMSDs) between the optimized surface CO_2 fluxes and the background fluxes were calculated (Fig. 13~~2~~). The surface CO_2 fluxes predicted by the dynamic model in Eq. (2) (i.e., background) show a high RMSD in the NH, with the highest values in North America, followed by Asia (Fig. 13~~2~~a). In terms of seasonal variation, the impact of the observations in JJA is greater than in DJF (Figs. 13~~2~~a and b). The large difference between the prior fluxes and the surface CO_2 fluxes predicted by the dynamic model implies that the assimilation of previous observations substantially affects the results. The RMSD of the analyzed surface CO_2 fluxes constrained by one week of observations from the background fluxes in JJA is greater in the NH compared with the other regions. The JJA RMSD value for North America (especially in the mid-continental region of the US) is the highest in the NH (Fig. 13~~2~~c). Although the RMSD of North America in DJF is lower than that in JJA, the RMSD of North America is still greater than that of other regions in DJF (Fig. 13~~2~~d). The regions with a high average information content are consistent with the regions with a high RMSD (compare Fig. 9~~8~~ and Fig. 13~~2~~), which implies that the observations from North America provide more information in the first cycle than those from other regions because the observations in North America are characterized by high self-sensitivities and abundant observations. By contrast, the RMSD values obtained in the first cycle in other regions are not as high as those in North America. The RMSD in Asia and other regions increases after five weeks of optimization (Figs. 13~~2~~e and f). In particular, the magnitude

of the RMSD in the Eurasian Boreal region increases after five weeks of optimization (Fig. 132e), which implies that, by the transport of the CO_2 concentrations, the observation information from remote regions affects the optimization of the surface CO_2 fluxes in the Eurasian Boreal region. This remote influence is due to the absence of observations in this region. In addition, the five-week assimilation lag is effective in optimizing the surface CO_2 flux in this region. Therefore, a longer, smoother window is necessary to optimize the surface CO_2 flux in Asia, where there are sparse observations; this may implyies that in the current version of CarbonTracker, the uncertainty of the surface CO_2 flux in Asia maycan be reduced when using a longer, smoother window than that used for North America. A study on the effect of various assimilation window and ensemble size on the estimation of the surface CO_2 flux in Asia is under way to investigate which determine more appropriate lag window and ensemble size are appropriate for Asia in CarbonTracker.

4 Summary and Conclusion

In this study, the effect of observations of CO_2 concentrations on the optimized surface CO_2 fluxes in CarbonTracker was evaluated by calculating the influence matrix for a 10-year period from 2000 to 2009. CarbonTracker is a system used to optimize the surface CO_2 flux using EnKF as a data assimilation algorithm. Most of the calculated influence values were in the range of the theoretical limit, from 0 to 1, which makes it possible to objectively diagnose the performance of the data assimilation system used in this study.

The average global self-sensitivity is 4.8%, which implies that the impact of the background on the optimized flux is 95.2%. The value of 4.8% obtained in CarbonTracker is lower than the 15% value obtained from NWP models, as reported by Cardinali et al. (2004) and Liu et al. (2009). However, as indicated by Cardinali et al. (2004), the background fluxes predicted by the dynamic model already include information extracted from earlier observations used in previous cycles. Because the state vector used in CarbonTracker includes five weeks of lag, the cumulative impact of the observations on the analysis is greater than the impact calculated for a single assimilation cycle. The cumulative impact over five weeks is 19.1%, much greater than 4.8%, and tThe large cumulative impact is confirmed by the RMSD of the surface CO_2 fluxes associated with each assimilation process.

The self-sensitivity and spatial coverage of the observation sites are inversely correlated in the NH,

whereasile these factors are not apparently related in the Tropics and SH. The lower correlation between the self-sensitivity and the spatial coverage of the observation sites in the Tropics and SH is attributed to either the sparseness of the observation sites or the locations of the observation sites which are not appropriate for detecting the variability of CO_2 concentrations with a high temporal resolution but are appropriate for detecting the global trend of the background CO_2 concentrations. By contrast, the observation sites near the Eurasian Boreal region show high self-sensitivity because there are no available observations in the Eurasian Boreal region.

The self-sensitivity time series is characterized by seasonal variations. In both hemispheres, the self-sensitivity in summer is greater than in winter, which is clearly evident in the Mixed and Continuous site categories and is associated with the background surface CO_2 flux uncertainties. The number of observations used in data assimilation increases over time, which causes the average self-sensitivities to decrease. The decreasing trend of the self-sensitivity over time for the Continuous site observations in North America may indicate the limited impact of additional observations in this region. Schuh et al. (2013) reported that additional tower measurements (i.e., observations in the Continuous site category) in the Corn Belt region of the US did not significantly alter the surface CO_2 flux estimates for 2008, which is consistent with the low self-sensitivity detected over North America in the same period. Therefore, under the current CarbonTracker framework, to obtain the beneficial effect of additional observations on the surface CO_2 flux analysis, new observations should be added in regions with a low spatial density of observation sites (e.g., Asia).

The observation sites with a high average self-sensitivity and a small number of observations show low average information content, whereas the observation sites with a low average self-sensitivity and a large number of observations show high average information content because the range of average self-sensitivity is bounded from o to 1, but the range of the number of observations is large. Therefore, the Continuous site category shows high average information content. In general, the information extracted from observations is concentrated in the NH, especially in North America. A strong correlation exists between the information content and the optimized surface CO_2 fluxes. The high information content found in regions with a large number of observations implies that much of the information is extracted from observations, and as a result, the fluxes are optimized quickly in a relatively short period. However, the surface CO_2 fluxes in regions with no local observation sites (e.g., Siberia) are optimized by remote observations

during relatively long assimilation windows with a lag.

The effect of various observations on the analyzed surface CO_2 fluxes can be calculated using the method suggested in this study. More CO_2 observations become available in data assimilation for estimating the surface CO_2 fluxes. These additional sources include CONTRAIL data, which are aircraft observations (Machida et al., 2008); column-averaged CO_2 concentrations retrieved from the Japanese Greenhouse gases Observing SATellite (GOSAT) (Yokoda et al., 2009); and data from the Total Carbon Column Observing Network (TCCON), which are observed by ground-based Fourier Transform Spectrometers (Wunch et al., 2011). As a next step, the impact of various observations on the optimization of surface CO_2 fluxes can be evaluated using the method suggested in this study.

Acknowledgements

The authors thank the two anonymous reviewers for their valuable comments. The authors thank Dr. Andrew R. Jacobson for providing the resources necessary for this study. This study was funded by the Korea Meteorological Administration Research and Development Program under the Grant CATER 2012-3032.

References

Anderson, J. I.: Spatially and temporally varying adaptive covariance inflation for ensemble filters, Tellus, 61A, 72-83, doi:10.1111/j.1600-0870.2008.00361.x, 2009.

Baker, D. F., Doney, S. C., and Schimel, D. S.: Variational data assimilation for atmospheric CO_2, Tellus, 58B, 359-365, 2006.

Baker, D. F., BÖsch, H., Doney, S. C., O'Brien, D., and Schimel, D. S.: Carbon source/sink information provided by column CO_2 measurements from the Orbiting Carbon Observatory, Atmos. Chem. Phys., 10, 4145-4165, doi:10.5194/acp-10-4145-2010, 2010.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO_2 fluxes estimated from GOSAT retrievals of total column CO_2, Atmos. Chem. Phys., 12, 86958717, doi:10.5194/acp-13-8695-2013, 2013.

Boweler, N. E, Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E.: The MOGREPS short-range ensemble prediction system, Q. J. R. Meteorol. Soc., 134, 703-722, 2008.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmos. Chem. Phys., 5, 2691-2702, 2005.

Cardinali, C., Pezzulli S., and Andersson E.: Influence-matrix diagnostic of a data assimilation system, Q. J. R. Meteorol. Soc., 130, 2767-2786, 2004.

Chatterjee, A., Michalak, A. M., Anderson, J. L., Mueller, K. L., Yadav, V.: Toward reliable ensemble Kalman filter estimates of CO_2 fluxes, J. Geophys. Res., 117, D22306, doi:10.1029/2012JD018176, 2012.

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

Chevallier, F., Fisher, M., Peylin, P., Seerrar, S., Bousque, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24309, doi:10.1029/2005JD006390, 2005.

Chevallier, F., Maksyutov, S., Bousquet, P., Bréon, F.-M., Saito, R., Yoshida, Y., and Yokota, T.: On the accuracy of the CO_2 surface fluxes to be estimated from the GOSAT observations, Geophys. Res. Lett., 36, L19807, doi:10.1029/2009GL040108, 2009b.

Ciais P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives, Climate Change, 103, 69-92, 2010.

Engelen R. J., and Stephens, G. L.: Information content of infrared satellite sounding measurements with respect to CO_2, J. Appl. Meteorol., 43, 373–378, 2004.

Feng, L., Palmer, P. I., Bosch, H., and Dance, S.: Estimating surface CO_2 fluxes from space-born CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmos. Chem. Phys., 9, 2619-2633, 2009.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.

Houtekamer, P. L., and H. L., Mitchell, A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Wea. Rev., 129, 123-137, 2001.

Jacobson, A. R., Fletcher, S. E. M., Gruber, N., Sarmiento, J. L., and Gloor, M.: A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 2. Regional results, Glob. Biogeochem. Cycles, 21, GB1019, doi:10.1029/2006GB002703, 2007.

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

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

Kim, J., Kim, H. M., and Cho, C.-H.: Application of Carbon Tracking System based on ensemble Kalman Filter on the diagnosis of Carbon Cycle in Asia, Atmosphere, 22(4), 415-447, 2012. (in

Korean with English abstract)

~~Kim, J., Kim, H. M., and Cho, C. H.: The effect of optimization and the nesting domain on carbon flux analyses in Asia using a carbon tracking system based on the ensemble Kalman filter, Asia-Pacific J. Atmos. Sci., accepted, 2014.~~

Kim, J., Kim, H. M., and Cho, C.-H.: The effect of optimization and the nesting domain on carbon flux analyses in Asia using a carbon tracking system based on the ensemble Kalman filter, Asia-Pacific J. Atmos. Sci., 50, 327–344, doi:10.1007/s13143-014-0020-y, 2014.

Krol, M., Houweling, S., Bregman, B., Broek, M., van der Segers, A., Velthoven, P. V., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications, Atmos. Chem. Phys., 5, 417-432, 2005.

Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, Q. J. R. Meteorol. Soc., 135, 523-533, 2009.

Liu, J, Kalnay, E., Miyoshi, T., and Cardinali, C.: Analysis sensitivity calculation in an ensemble Kalman filter, Q. J. R. Meteorol. Soc., 135, 1842-1851, 2009.

Lokupitiya, R., S., Zupanski, D., Denning, A. S., Kawa, S. R., Gurney, R., and Zupanski, M.: Estimation of global CO_2 fluxes at regional scale using the maximum likelihood ensemble filter, J. Geophys. Res., 113, D20110, doi:10.1029/2007JD009679, 2008.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., Goto, K., Nakazawa, T., Ishikawa, K., and Ogawa, T.: Worldwide measurements of atmospheric CO_2 and other trace gas species using commercial airlines, J. Atmos. Oceaninc. Technol., 25, 1744-1754, doi:10.1175/2008JTECHA1082.1, 2008.

Masarie, K. A., Pétron, G., Andrews, A., Bruhwiler, L., Conway, T. J., Jacobson, A. R., Miller, J. B., Tans, P. P., Worthy, D. E., and Peters, W.: Impact of CO_2 measurement bias on CarbonTracker surface flux estimates, J. Geophys. Res., 116, D17305, doi:10.1029/2011JD016270, 2011.

Meirink, J. F., Bergamaschi, P., Frankenberg, C., d'Amelio, M. T. S., Dlugokencky, E. J., Gatti, L. V., Houweling, S., Miller, J. B., RÖckmann, T., Villani, M. G., and Krol, M. C.: Four-dimensional variational

서식 있음: 첫 줄: 0 글자

서식 있음: 독일어(독일)

data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations, J. Geophys. Res., 113, D17301, doi:10.1029/2007JD009740, 2008.

Miyazaki, K., Maki, T., Patra, P., and Nakazawa, T.: Assessing the impact of satellite, aircraft, and surface observations on CO_2 flux estimation using an ensemble-based 4-D data assimilation system, J. Geophys. Res., 116, D16306, doi:10.1029/2010JD015366, 2011.

Miyoshi, T.: 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, 2011.

Talagrand O.: Assimilation of observations, an introduction. J. Meteorol Soc. Jpn., 75, 191-209, 1997.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304, doi:10.1029/2005JD006157, 2005.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proc. Nat. Acd. Sci. U.S.A., 104, 18925-18930, 2007.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R. Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Strӧm, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, A., Rodӧ, X., Morguí, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P. Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Global Change Biol., 16, 1317-1337, doi:10.1111/j.1365-2486.2009.02078.x, 2010.

Peylin P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699-6720,

doi:10.5194/bg-10-6699-2013, 2013.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, World Scientific, London, 2000.

Schuh, A. E., Lauvaux, T., West, T. O., Denning, A. S., Dvais, K. J., Miles, N., Richardson, S., Uliasz, M., Lokupitiya, E., Cooley, D., Andrews, A., and Ogle S.: Evaluating atmospheric CO_2 inversions at multiple scales over a highly inventoried agricultural landscape, Global Change. Biol., 19, 1424-1439, doi:10.1111/gcb.12141, 2013.

Wang, X., and Bishop, C. H.: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, J. Atmos. Sci., 60, 1140-1158, 2003.

~~van der Werf, G. R., Randersom, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423-3441, 2006. 본문에서 빠지기 때문에 없어져야 합니다.~~

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.

Whitaker, J. S., and Hamill, T. M.: Ensemble Data Assimilation without Perturbed Observations, Mon. Wea. Rev., 130, 1913-1924, 2002.

Whitaker, J. S., Hamill, T. M., Wei, X., Song, Y., and Toth, Z.: Ensemble data assimilation with the NCEP global forecast system, Mon. Wea. Rev., 136, 463-482, doi:10.1175/2007MWR2018.1, 2008.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jim´enez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., Maksyutov, S.: Global

서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음
서식 있음: 강조 없음

Concentrations of CO_2 and CH_4 Retrieved from GOSAT: First Preliminary Results, Sci. Online. Lett. Atmos., 5, 160-163, doi:10.2151/sola.2009-041, 2010.

~~Zhang, H. F, Chen, B. Z., van der Laan Luijkx, Machida, T., Matsueda, H., Sawa, Y, Fukuyama, Y., Labuschange, C., Langenfelds, R., van der Schoot, M., Xu, G., Yan, J. W., Zhou, L. X., Tans, P. P., and Peters, W.: Estimating Asian terrestrial carbon fluxes from CONTRAIL aircraft and surface CO_2 observations for the period 2006 to 2010, Atmos. Chem. Phys. Discuss., 13, 27597-27639, 2013.~~

Zhang, H. F., Chen, B. Z. , van der Laan-Luijkx, I. T., Machida, T., Matsueda, H., Sawa, Y., Fukuyama, Y., Langenfelds, R., van der Schoot, M., Xu, G., Yan, J. W., Cheng, M. L., Zhou, X., Tans, P. P., and Peters, W.: Estimating Asian terrestrial carbon fluxes from CONTRAIL aircraft and surface CO_2 observations for the period 2006-2010, Atmos. Chem. Phys., 14, 5807-5824, doi:10.5194/acp-14-5807-2014, 2014.

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

서식 있음: 첫 줄: 0 글자

서식 있음: 아래 첨자

Table 1. Information on the observation sites used in this study. MDM represents the model-data mismatch, which is the observation error.

Site code	Location	Latitude	Longitude	Height	Laboratory	MDM [ppm]
ALT_01D0	Alert, Nunavut, Canada	82.45°N	62.51°W	200 m	ESRL	1.5
ALD_06C0	Alert, Nunavut, Canada	82.45°N	62.51°W	200 m	ESRL	2.5
AMT_01C3	Argyle, Maine, United States	45.03°N	68.68°W	50 m	ESRL	3
AMT_01P0	Argyle, Maine, United States	45.03°N	68.68°W	50 m	ESRL	3
ASC_01D0	Ascension Island, United Kingdom	7.92°S	14.42°W	54 m	ESRL	0.75
ASK_01D0	Assekrem, Algeria	23.18°N	5.42°E	2728 m	ESRL	1.5
AZR_01D0	Terceira Island, Azores, Portugal	38.77°N	27.38°W	40 m	ESRL	1.5
BAL_01D0	Baltic Sea, Poland	55.35°N	17.22°E	3 m	ESRL	7.5
BAO_01C3	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1584 m	ESRL	3
BAO_01P0	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1584 m	ESRL	3
BKT_01D0	Bukit Kotobang, Indonesia	0.20°S	100.32°E	864 m	ESRL	7.5

BME_01D0	St. Davids Head, Bermuda, United Kingdom	32.27°N	64.65°E	30 m	ESRL	1.5
BMW_01D0	Tudor Hill, Bermuda, United Kingdom	32.27°N	64.88°E	30 m	ESRL	1.5
BRW_01D0	Barrow, Alaska, United States	71.32°N	156.61°W	11 m	ESRL	1.5
BRW_01C0	Barrow, Alaska, United States	71.32°N	156.61°W	11 m	ESRL	2.5
BSC_01D0	Black Sea, Constanta, Romania	44.17°N	28.68°E	3 m	ESRL	7.5
CBA_01D0	Cold Bay, Alaska, United States	55.21°N	162.72°W	21 m	ESRL	1.5
CDL_06C0	Candle Lake, Saskatchewan, Canada	53.99°N	105.12°W	600 m	ESRL	3
CFA_02D0	Cape Ferguson, Queensland, Australia	19.28°S	147.06°E	184 m	ESRL	2.5
CGO_01D0	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94 m	ESRL	0.75
CGO_02D0	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94 m	CSIRO	0.75
CHR_01D0	Christmas Island, Republic of Kiribati	1.70°N	157.17°W	3 m	ESRL	0.75
CRZ_01D0	Crozet Island, France	46.45°S	51.85°E	120 m	ESRL	0.75
cya_02D0	Casey, Antarctica, Australia	66.28°S	110.5°E	51 m	CSIRO	0.75

EGB_06C0	Egbert, Ontario, Canada	44.23ºN	79.78ºW	251 m	EC	3
EIC_01D0	Easter Island, Chile	27.15ºS	109.45ºW	50 m	ESRL	7.5
ESP_06C0	Estevan Point, British Columbia, Canada	49.38ºN	126.54ºW	7 m	EC	3
ETL_06C0	East Trout Lake, Saskatchewan, Canada	54.35ºN	104.98ºW	492 m	EC	3
FEF_03C0	Fraser, Colorado, United Stated	39.91ºN	105.88ºW	2745 m	NCAR	3
FSD_06C0	Fraserdale, Canada	49.88ºN	81.57ºW	210 m	EC	3
GMI_01D0	Mariana Islands, Guam	13.43ºN	144.78ºE	2 m	ESRL	1.5
HBA_01D0	Halley Station, Antarctica, United Kingdom	75.58ºS	26.50ºW	30 m	ESRL	0.75
HDP_03C0	Hidden Peak (Snowbird), Utah, United States	40.56ºN	111.65ºW	3351 m	NCAR	3
HUN_01D0	Hegyhatsal, Hungary	46.95ºN	16.65ºE	248 m	ESRL	7.5
ICE_01D0	Storhofdi, Vestmannaeyjar, Iceland	63.40ºN	20.29ºW	118 m	ESRL	1.5
KEY_01D0	Key Biscayne, Florida, United States	25.67ºN	80.16ºW	3 m	ESRL	2.5
KUM_01D0	Cape Kumukahi, Hawaii, United States	19.52ºN	154.82ºW	3 m	ESRL	1.5

KZD_01D0	SaryTaukum, Kazakhstan	44.06ºN	76.82ºE	601 m	ESRL	2.5
KZM_01D0	Plateau Assy, Kazakhstan	43.25ºN	77.88ºE	2519 m	ESRL	2.5
LEF_01C3	Park Falls, Wisconsin, United States	45.95ºN	90.27ºW	472 m	ESRL	3
LEF_01P0	Park Falls, Wisconsin, United States	45.95ºN	90.27ºW	472 m	ESRL	3
LLB_06C0	Lac La Biche, Alberta, Canada	54.95ºN	112.45ºW	540 m	EC	3
MAA_02D0	Mawson Station, Antarctica, Australia	67.62ºS	62.87ºE	32 m	CSIRO	0.75
MHD_01D0	Mace Head, County Galway, Ireland	53.33ºN	9.90ºW	5 m	ESRL	2.5
MID_01D0	Sand Island, Midway, United States	28.21ºN	177.38ºW	4 m	ESRL	1.5
MKN_01D0	MT. Kenya, Kenya	0.05ºS	37.30ºE	3897 m	ESRL	2.5
MLO_01C0	Mauna Loa, Hawaii, United States	19.54ºN	155.58ºW	3397 m	ESRL	0.75
MLO_01D0	Mauna Loa, Hawaii, United States	19.54ºN	155.58ºW	3397 m	ESRL	1.5
MNM_19C0	Minamitorishima, Japan	24.29ºN	153.98ºE	8 m	JMA	3
MQA_02D0	Macquarie Island, Australia	54.48ºS	158.97ºE	12 m	CSIRO	0.75

NMB_01D0	Gobabeb, Namibia	23.58ºS	15.03ºE	456 m	ESRL	2.5
NWR_01D0	Niwot Ridge, Colorado, United States	40.05ºN	105.58ºW	3523 m	ESRL	1.5
NWR_03C0	Niwot Ridge, Colorado, United States	40.05ºN	105.58ºW	3523 m	NCAR	3
OBN_01D0	Obninsk, Russia	55.11ºN	36.60ºE	183 m	ESRL	7.5
OXK_01D0	Ochsenkopf, Germany	50.03ºN	11.80ºE	1022 m	ESRL	2.5
PAL_01D0	Pallas-Sammaltunturi, GAW Station, Germany	67.97ºN	24.12ºE	560 m	ESRL	2.5
POC_01D1	Pacific Ocean, N/A	0.39ºS	132.43ºW	10 m	ESRL	0.75
PSA_01D0	Palmer Station, Antarctica, United States	64.92ºS	64.00ºW	10 m	ESRL	0.75
PTA_01D0	Point Arena, California, United States	38.95ºN	123.74ºW	17 m	ESRL	7.5
RPB_01D0	Ragged Point, Barbados	13.17ºN	59.43ºW	45 m	ESRL	1.5
RYO_19C0	Ryori, Japan	39.03ºN	141.82ºE	260 m	JMA	3
SCT_01C3	Beech Island, South Carolina, United States	33.41ºN	81.83ºW	115 m	ESRL	3
SEY_01D0	Mahe Island, Seychelles	4.67ºS	55.17ºE	3 m	ESRL	0.75

SGP_01D0	Southern Great Plains, Oklahoma, United States	36.80ºN	97.50ºW	314 m	ESRL	2.5
SGP_64C3	Southern Great Plains, Oklahoma, United States	36.80ºN	97.50ºW	314 m	ESRL	3
SHM_01D0	Shemya Island, Alaska, United States	52.72ºN	174.10ºE	40 m	ESRL	2.5
SMO_01C0	Tutuila, American Samoa	14.25ºS	170.56ºW	42 m	ESRL	0.75
SMO_01D0	Tutuila, American Samoa	14.25ºS	170.56ºW	42 m	ESRL	1.5
SNP_01C3	Shenandoah National Park, United States	38.62ºN	78.35ºW	1008 m	ESRL	3
SPL_01C3	Storm Peak Laboratory (Desert Research Institute), United States	40.45ºN	106.73ºW	3210 m	NCAR	3
SPO_01C0	South Pole, Antarctica, United Stated	89.98ºS	24.80ºW	2810 m	ESRL	0.75
SPO_01D0	South Pole, Antarctica, United Stated	89.98ºS	24.80ºW	2810 m	ESRL	1.5
STM_01D0	Ocean Station M, Norway	66.00ºN	2.00ºE	0 m	ESRL	1.5
STR_01P0	Sutro Tower, San Francisco, California, United States	37.76ºN	122.45ºW	254 m	ESRL	3
SUM_01D0	Summit, Greenland	72.57ºN	38.48ºW	3238 m	ESRL	1.5
SYO_01D0	Syowa Station, Antarctica, Japan	69.00ºS	39.58ºE	11 m	ESRL	0.75

TAP_01D0	Tae-ahn Peninsula, Republic of Korea	36.73ºN	126.13ºE	20 m	ESRL	5
TDF_01D0	Tierra Del Fuego, Ushuaia, Argentina	54.87ºS	68.48ºW	20 m	ESRL	0.75
THD_01D0	Trinidad head, California, United States	41.73ºN	91.35ºW	107 m	ESRL	2.5
UTA_01D0	Wendover, Utah, United States	39.90ºN	113.72ºW	1320 m	ESRL	2.5
UUM_01D0	Ulaan Uul, Mongolia	44.45ºN	111.10ºE	914 m	ESRL	2.5
WBI_01C3	West Branch, Iowa, United States	41.73ºN	91.35ºW	242 m	ESRL	3
WBI_01P0	West Branch, Iowa, United States	41.73ºN	91.35ºW	242 m	ESRL	3
WGC_01C3	Walnut Grove, California, United States	38.27ºN	121.49ºW	0 m	ESRL	3
WGC_01P0	Walnut Grove, California, United States	38.27ºN	121.49ºW	0 m	ESRL	3
WIS_01D0	WIS Station, Negev Desert, Israel	31.13ºN	34.88ºE	400 m	ESRL	2.5
WKT_01C3	Moody, Texas, United States	31.32ºN	97.33ºW	251 m	ESRL	3
WKT_01C3	Moody, Texas, United States	31.32ºN	97.33ºW	251 m	ESRL	3
WLG_01D0	Mt. Waliguan, Peoples Republic of China	36.29ºN	100.90ºE	3810 m	ESRL	1.5

WSA_06C0	Sable Island, Nova Scotia, Canada	49.93ºN	60.02ºE	5 m	EC	3
YON_19C0	Yonagunijima, Japan	24.47ºN	123.02ºE	30 m	JMA	3
ZEP_01D0	Ny-Alesund, Svalbard, Norway and Sweden	78.90ºN	11.88ºE	475 m	ESRL	1.5

Table 2. Observation site categories and corresponding MDM~~model-data-mismatch~~ values ~~[ppm]~~.

서식 있는 표

Observation category	Description	Observation frequency	MDM~~odel-data-mismatch~~ [ppm]
Marine Boundary Layer (MBL)	Observation site close to Marine boundary layer	Once a week	0.75
Mixed land/ocean and mountain (Mixed)	Observation site located in mixed land, ocean, and mountain	Once a week	1.5
Continental	Observation site located in the continent	Once a week	2.5
Continuous	Observation site with continuous observations	Once a day	3
Difficult	Difficult	Once a week	7.5 (5.0)

Table 3. Information on the observation sites located in Asia, including the number of observations, number of rejected observations, MDM values, innovation χ^2 statistics, and the average bias of the model CO_2 concentrations calculated by optimized fluxes. For the TAP_01D0 site, the numbers in parentheses are values used in previous studies, and the numbers without parentheses are the modified values based on the innovation χ^2 statistics in this study.

Site name	Number of observations	Number of rejected observations	MDM	Innovation χ2	Bias of model CO_2 concentration
BKT_01D0	207	0	7.5	0.57	-4.01
KZD_01D0	430	11	2.5	1.25	-0.4
KZM_01D0	384	9	2.5	1.22	-0.67
MNM_19C0	3304	0	3	0.16	-0.45
RYO_19C0	3149	108	3	0.53	-0.9
TAP_01D0	339	10	5	0.59	0.01
	(269)	(3)	(7.5)	(0.37)	(-0.26)
UUM_01D0	454	10	2.5	1.03	0.26
WIS_01D0	489	3	2.5	0.72	-0.15
WLG_01D0	347	10	1.5	1.14	0.04
YON_19C0	2947	8	3	0.53	-0.9

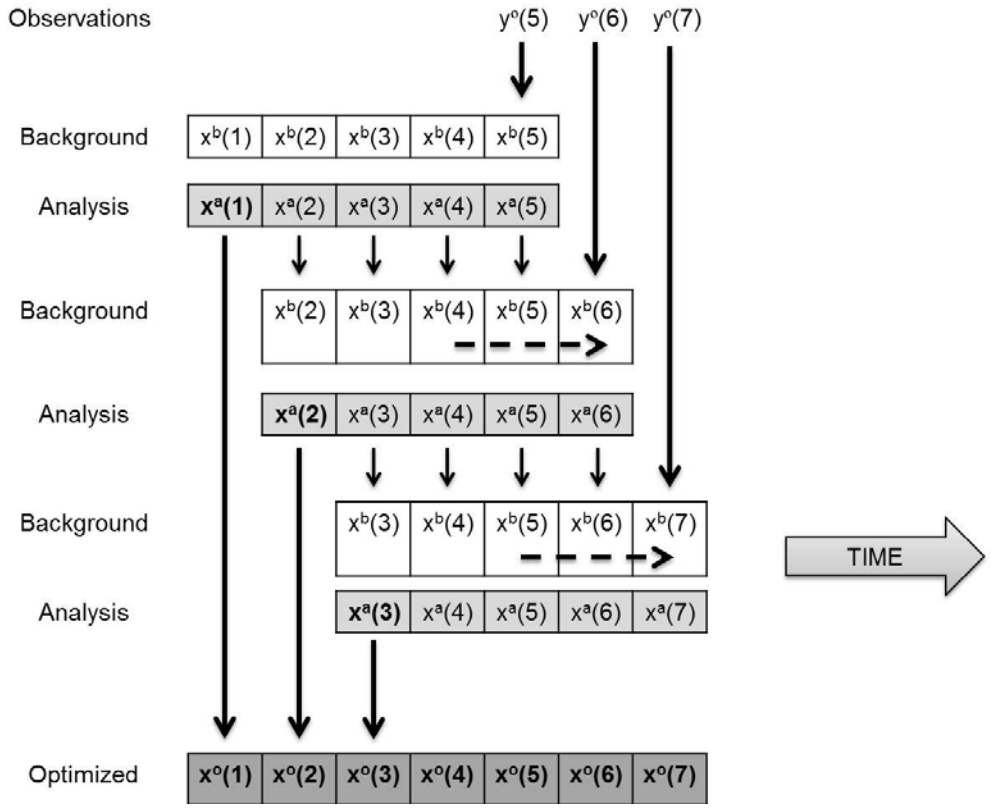
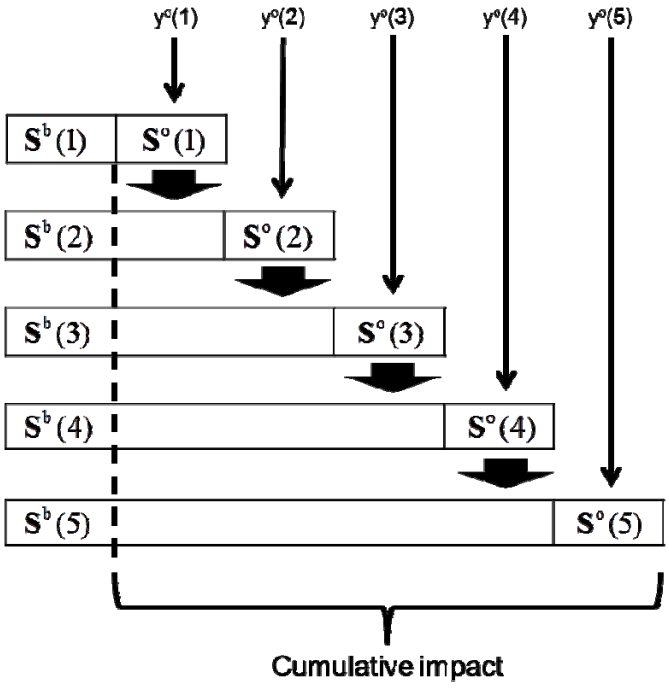


Figure 1. Schematic diagram of the assimilation process employed in CarbonTracker. In each analysis cycle, observations made within one week are used to update the state vectors with a five-week lag. The dashed line indicates how the simple dynamic model uses analysis state vectors from the previous one and two weeks to produce a new background state vector for the current analysis time. The TM5 model is used as the observation operator to calculate the model CO_2 concentration for each corresponding observation location and time.


y°(1)
y°(2)
y°(3)
y°(4)
y°(5)
S^b(1)
S°(1)
S^b(2)
S°(2)
S^b(3)
S°(3)
S^b(4)
S°(4)
S^b(5)
S°(5)
Cumulative impact

서식 있음: 가운데

1

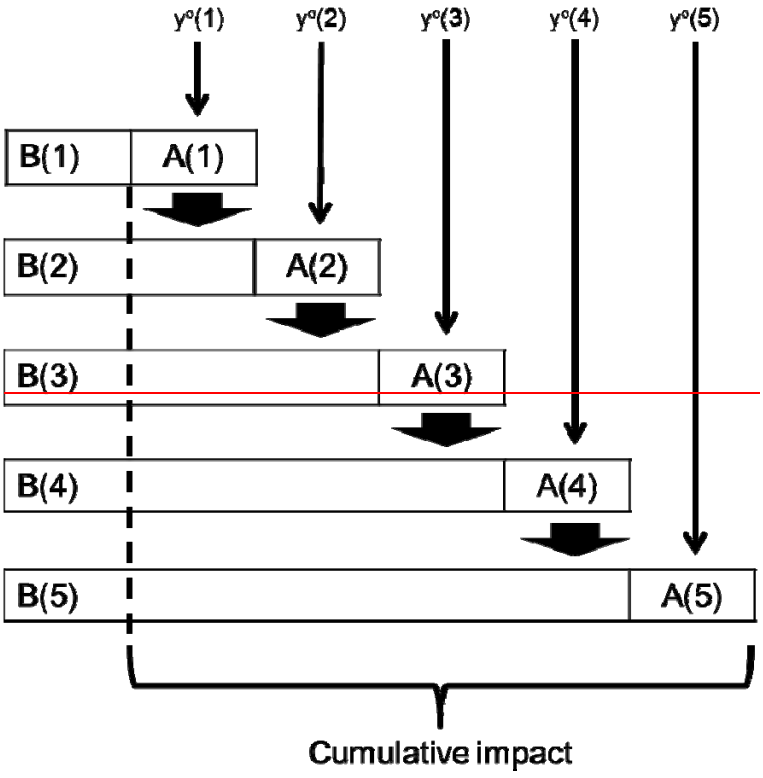


Figure 2. Schematic diagram of calculating ~~definition of~~ cumulative impact in ~~the~~ CarbonTraker ~~framework~~. $\mathbf{S}^{b}(\bullet)$ ~~B~~ indicates the analysis sensitivity to background at each analysis cycle within five weeks of lag, where $\bullet$ denotes each week from 1 to 5. $\mathbf{S}^{o}(\bullet)$ ~~A~~ indicates the analysis sensitivity to observation at each analysis cycle.

서식 있음: 글자 위치 내림: 2 pt

서식 있음: 가운데

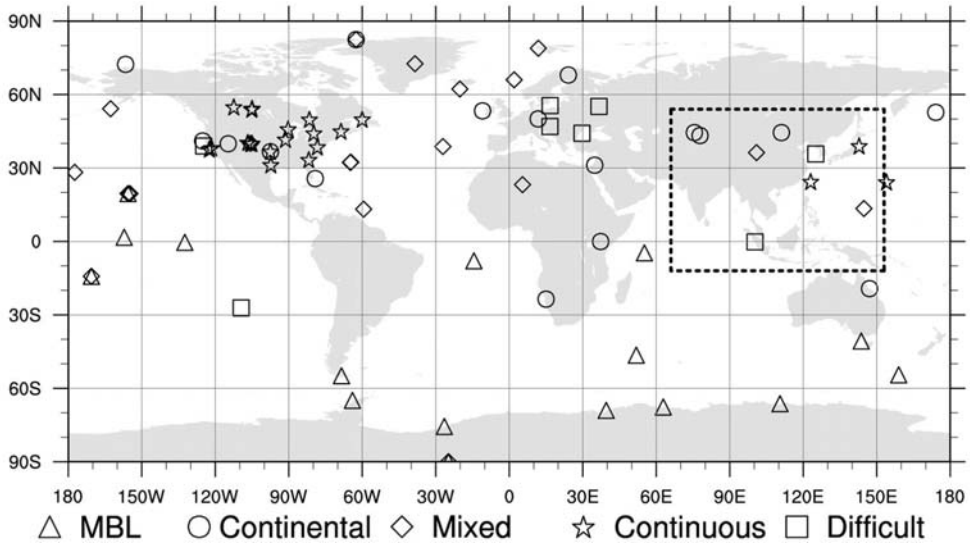


Figure 3~~2~~. Observation network of CO_2 concentrations around the globe and the nested domain of the TM5 transport model over Asia (dashed box). Each observation site is assigned to different categories (△: MBL; ○: Continental; ◇: Mixed land/ocean and mountain; ☆: Continuous; □: Difficult).

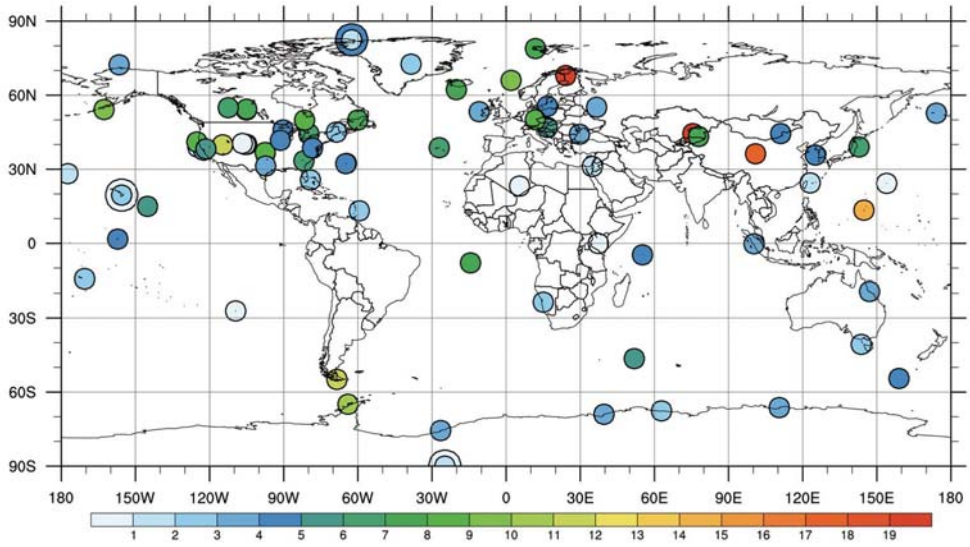


Figure 43. Average self-sensitivity at each observation site from 2000 to 2009. The overlapping observation sites at the same locations or at close locations are distinguished by different sizes of circles.

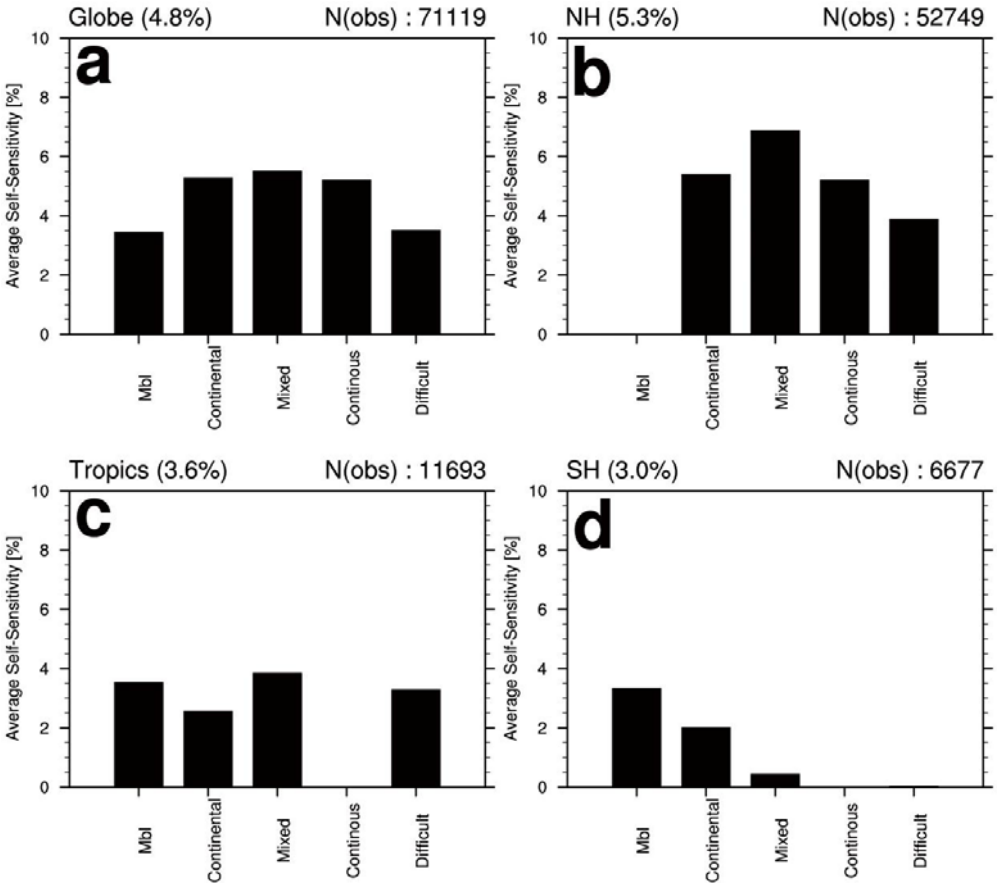


Figure 54. Histograms of the average self-sensitivity for each observation site category from 2000 to 2009 (a) around the globe and in the (b) Northern Hemisphere, (c) Tropics, and (d) Southern Hemisphere. N(obs) in the upper right corner represents the number of observations used in data assimilation.

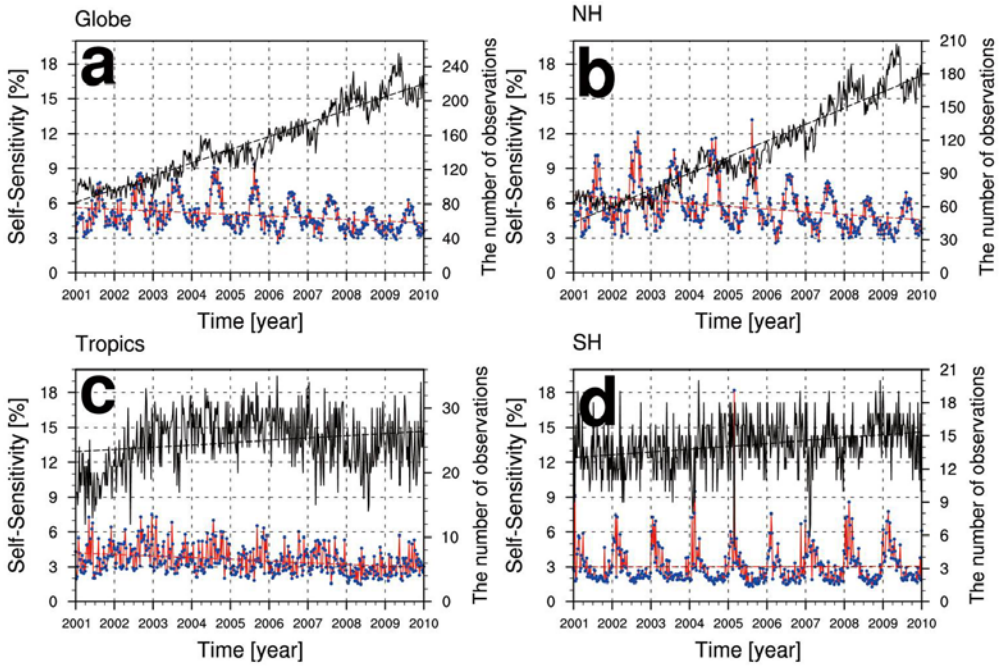


Figure 65. Time series of the average self-sensitivity (red solid line with blue dots) and the number of observations (black solid line) with a weekly temporal resolution (a) around the globe and in the (b) Northern Hemisphere, (c) Tropics, and (d) Southern Hemisphere from 2000 to 2009. The dashed lines represent the regression lines for the average self-sensitivity (red dashed line) and the number of observations (black dashed line).

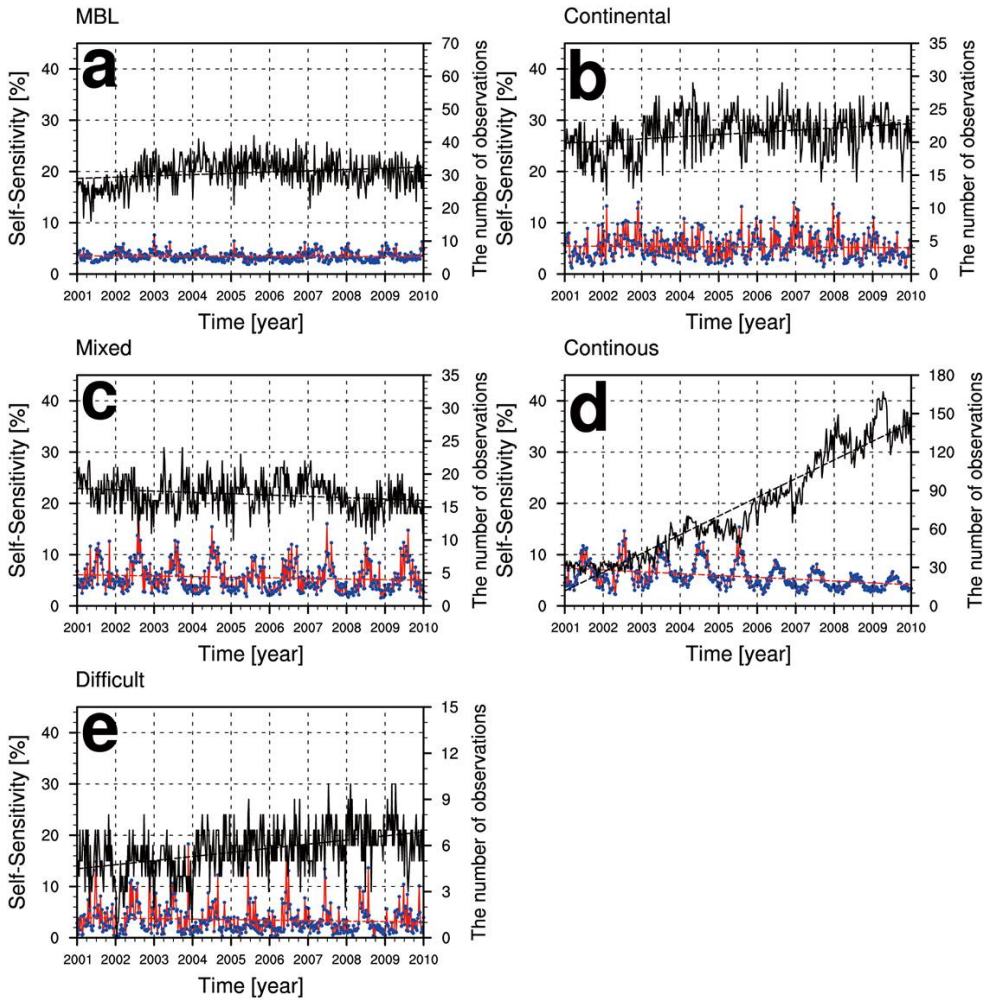


Figure 7~~6~~. Time series of the average self-sensitivity (red solid line with blue dots) and the number of observations (black solid line) with a weekly temporal resolution for the (a) MBL, (b) Continental, (c) Mixed, (d) Continuous, and (e) Difficult observation site categories from 2000 to 2009. The dashed lines represent the regression lines for the average self-sensitivity (red dashed line) and the number of observations (black dashed line).

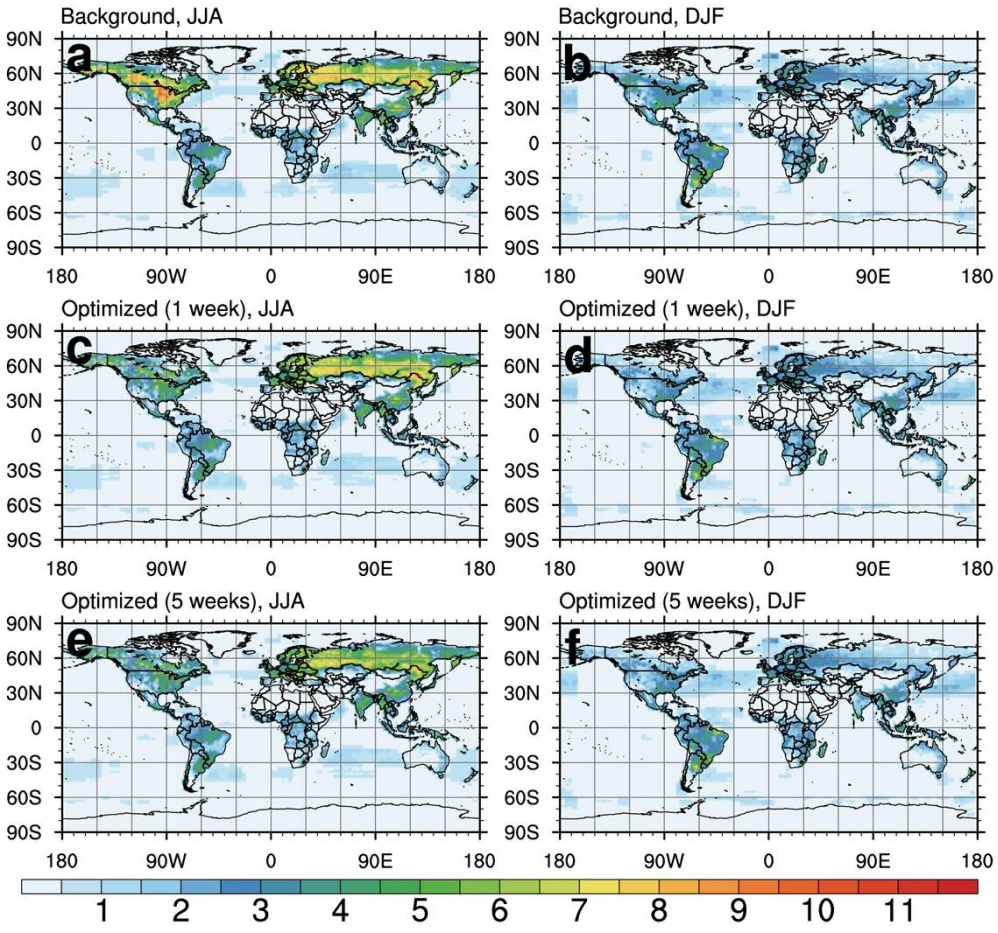


Figure 8~~7~~. Average standard deviation of background biosphere and ocean fluxes in (a) JJA and (b) DJF; the posterior biosphere and ocean fluxes optimized by one-week observations in (c) JJA and (d) DJF; and the posterior biosphere and ocean fluxes optimized by five weeks of observations in (e) JJA and (f) DJF. The units are g C m^{-2} $week^{-1}$.

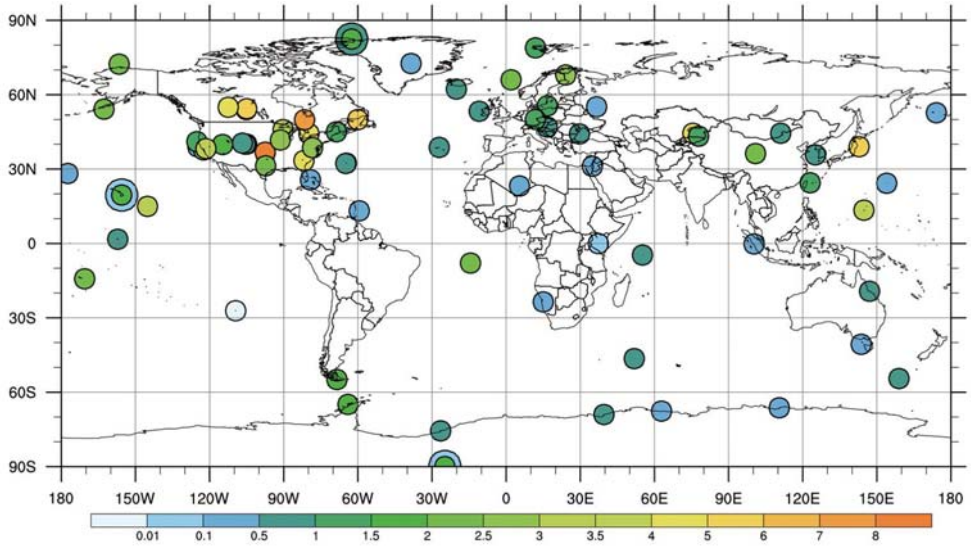


Figure 98. Average normalized information content for each observation site from 2000 to 2009. The overlapping observation sites at the same locations or at close locations are distinguished using different sizes of circles.

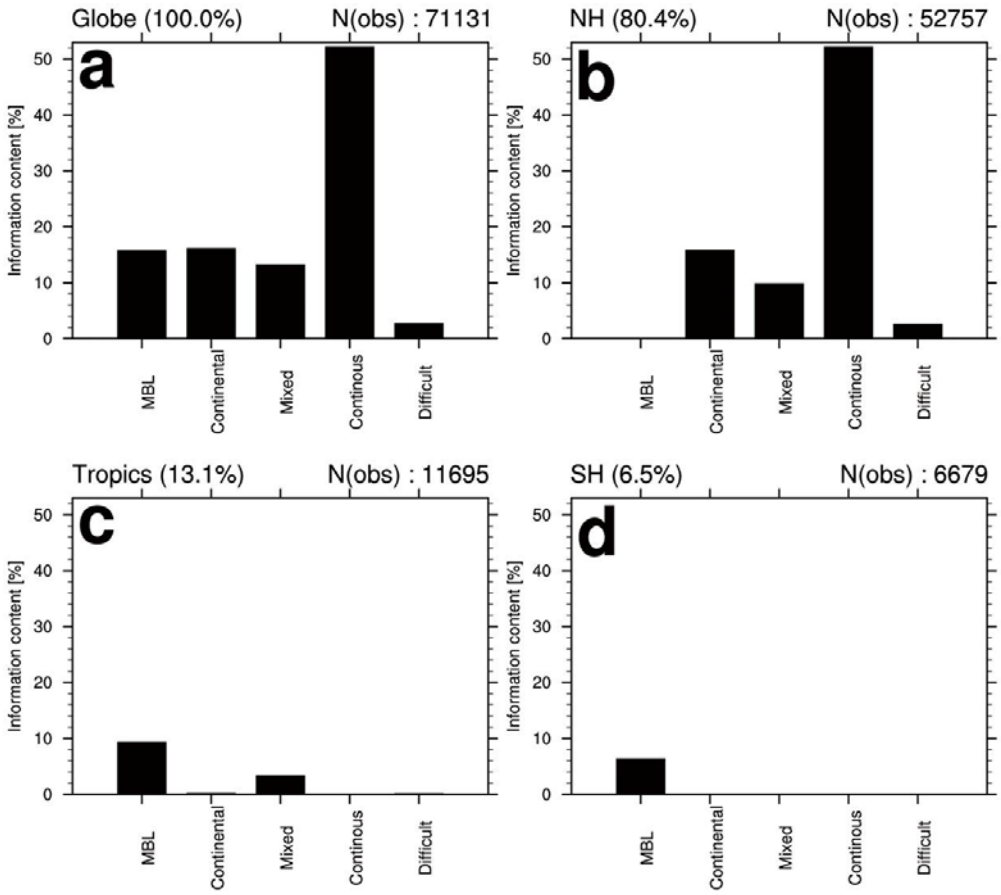


Figure 109. Histograms of the average information content for each observation site category (a) around the globe and in the (b) Northern Hemisphere, (c) Tropics, and (d) Southern Hemisphere from 2000 to 2009. N(obs) in the upper right corner represents the number of observations used in data assimilation.

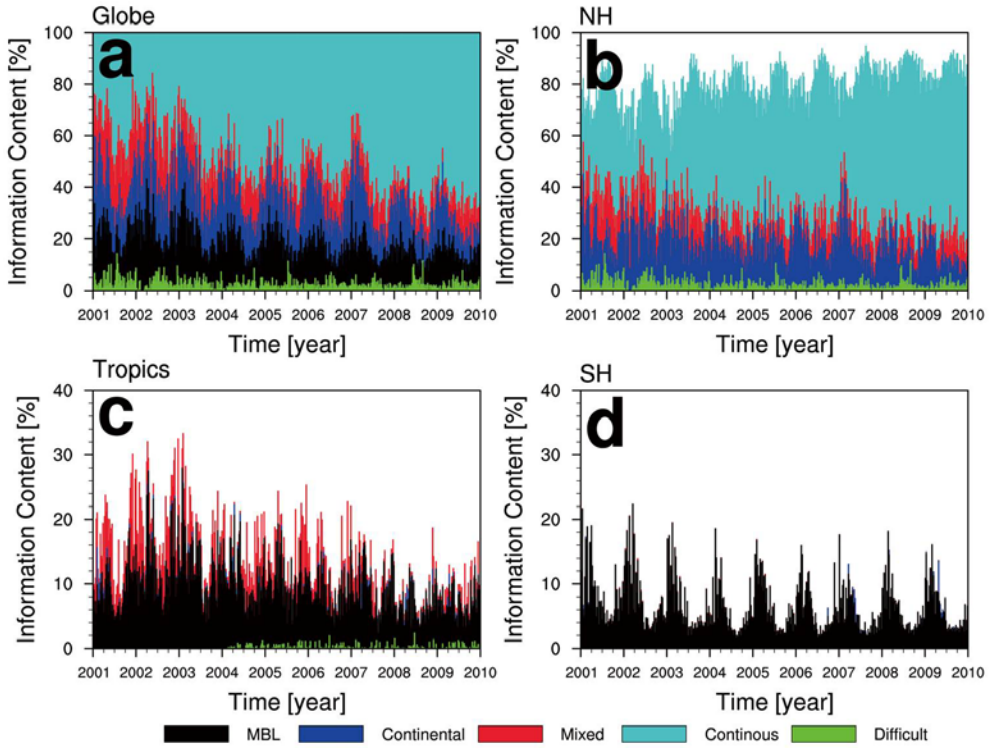


Figure 110. Time series of the average information content for each observation site category (a) around the globe and in the (b) Northern Hemisphere, (c) Tropics, and (d) Southern Hemisphere from 2000 to 2009.

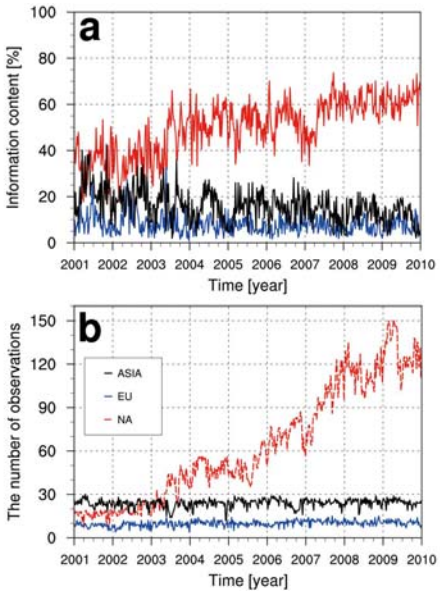


Figure 1~~2~~1. Times series of the (a) weekly averaged information content and (b) number of observations in Asia (black line), Europe (blue line), and North America (red line) from 2000 to 2009.

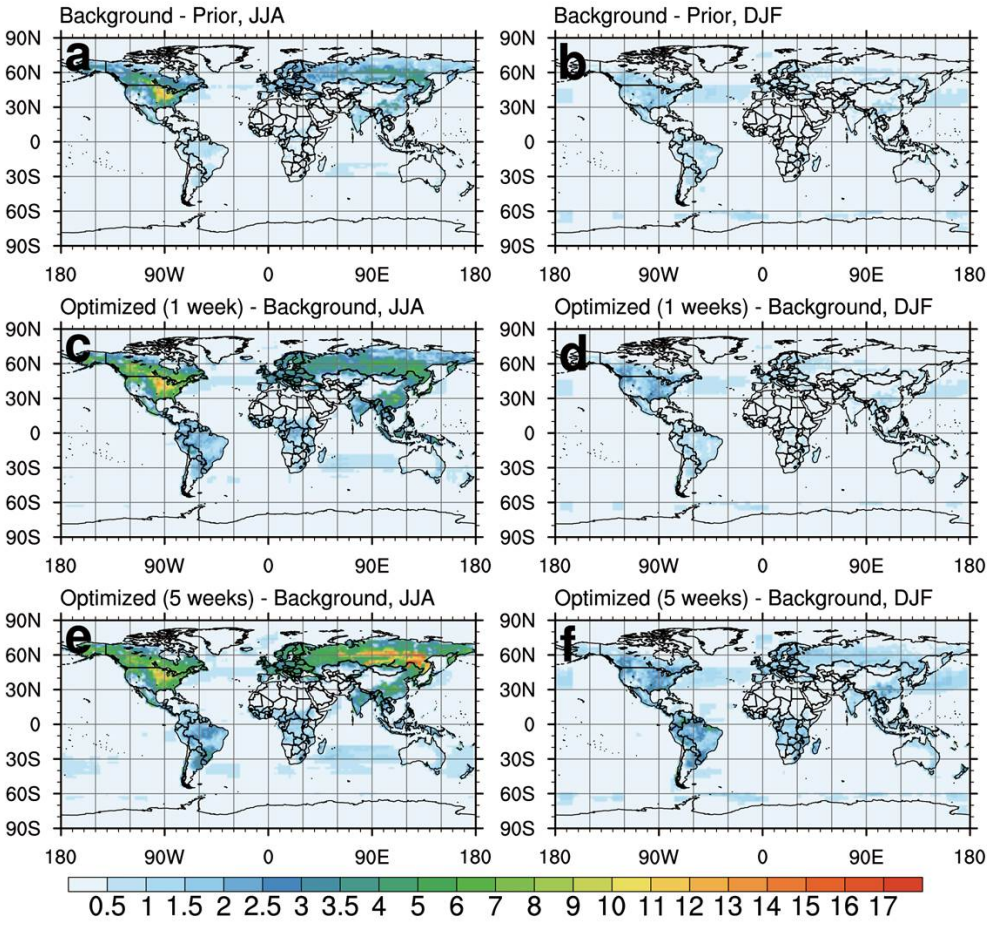


Figure 132. Root mean square difference (RMSD) between the background flux and prior flux in (a) JJA and (b) DJF; RMSD between the background flux and posterior flux optimized by one-week observations in (c) JJA and (d) DJF; and RMSD between the background flux and posterior flux optimized by five weeks of observations in (e) JJA and (f) DJF. The units are g C m^{-2} $week^{-1}$.