Supplement of Atmos. Chem. Phys., 26, 9257–9275, 2026
https://doi.org/10.5194/acp-26-9257-2026-supplement

Atmospheric Chemistry and Physics
Open Access

Supplement of

Investigating information transfer in CO_2 flux inversions: an analysis of ensemble Kalman filter based on Monte Carlo simulations

Shidong Fan and Ying Li

Correspondence to: Ying Li (liy66@sustech.edu.cn)

S1. The derivation of Kalman Filter

In this work, we adopt a traditional statistical estimation methods to derive the KF (Rao, 1994; Sengupta and Jammalamadaka, 2020). Besides, we follow Anderson (2001) to work on the joint state space. Set the state variables as $\boldsymbol{Z} = (\boldsymbol{y}, \boldsymbol{E_1}, \dots, \boldsymbol{E_k})^T$, where $\boldsymbol{y} = \mathcal{H}\big(\mathcal{M}(\boldsymbol{y_0}, \boldsymbol{E_1}, \dots, \boldsymbol{E_k})\big)$ is the vector of the observed quantities at specific locations and times and has a dimension n. Letters in bold face denote vectors or matrices. $\mathcal{M}$ and $\mathcal{H}$ are the dynamical model (e.g., atmospheric transport model) and the observation operator, respectively, both of which need not to be linear (they will be denoted as $\boldsymbol{M}$ and $\boldsymbol{H}$, if they are linear).

From the perspective of constrained linear estimation, we have a general linear statistical model

$$\boldsymbol{Z}^b \coloneqq \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{E}_1 \\ \vdots \\ \boldsymbol{E}_k \end{pmatrix}^b \coloneqq \begin{pmatrix} \boldsymbol{y}^b \\ \boldsymbol{E}_1^b \\ \vdots \\ \boldsymbol{E}_k^b \end{pmatrix} = \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{E}_1 \\ \vdots \\ \boldsymbol{E}_k \end{pmatrix} + \boldsymbol{\varepsilon}^b = \boldsymbol{Z} + \boldsymbol{\varepsilon}^b \tag{S1}$$

where superscript b stands for background and $\boldsymbol{\varepsilon^b}$ is the background error with covariance matrix $\boldsymbol{B}^b$. This linear statistical model can be interpreted as that background values $\boldsymbol{Z}^b$ are equal to true values $\boldsymbol{Z}$ plus errors $\boldsymbol{\varepsilon}^b$, where true values are not random variables but errors are. In this study we do not symbolically distinguish random variables ($\boldsymbol{Z}^b$, $\boldsymbol{\varepsilon}^b$, and so on) and their realizations. In addition to this linear statistical model, we have a constraint

$$\boldsymbol{y}^o = \boldsymbol{P}\boldsymbol{Z} + \boldsymbol{\varepsilon}^o \tag{S2}$$

where superscript o stands for observation. $\boldsymbol{P}$ is just a projection matrix that mapping $\boldsymbol{Z}$ to $\boldsymbol{y}$ and here equals to $\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ with proper dimensions (depending on the dimensions of $\boldsymbol{E}_1$ to $\boldsymbol{E}_k$). $\boldsymbol{\varepsilon}^o$ is the observation error, a random variable with covariance matrix $\boldsymbol{R}$. According to the linear estimation theory (Sengupta and Jammalamadaka, 2020), when observations and backgrounds are uncorrelated, the constrained best linear unbiased estimator (BLUE, linear estimator with minimum variance) of this general linear statistical model with constraint is

$$\widehat{\boldsymbol{Z}} = \boldsymbol{Z}^b + Cov(\boldsymbol{Z}^b, \boldsymbol{P}\boldsymbol{Z}^b - \boldsymbol{y}^o)[Var(\boldsymbol{P}\boldsymbol{Z}^b - \boldsymbol{y}^o)]^{-1}(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b). \tag{S3}$$

This state update equation can be understood in a general sense as follows. The constrained estimator is a modification to the unconstrained estimator that incorporates information from the constraint (i.e., the observation). The unconstrained estimator of $\boldsymbol{Z}$ is $\boldsymbol{Z}^b$, which reflects the estimates of true values based solely on background information, without any observations. When we apply the constraint, we must properly adjust the background information provided by $\boldsymbol{Z^b}$ via subtracting it from the information provided by the constraint. This adjustment yields the new information given by $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$, the observation innovation. To develop a linear estimator for $\boldsymbol{Z}$, we need to find a linear combination of these two sources of information: $\boldsymbol{Z}^b$ and $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$. To ensure that we derive an estimator with minimal variance, we first introduce an ordering relation that allows us to define what "minimal" means. We utilize the Löwner order of matrices ($\succcurlyeq$), which is a partial order that possesses desirable properties, such as the nondecreasing of diagonal elements and consequently the nondecreasing of trace for non-negative definite square matrices (as covariance matrices are non-negative definite), see Seber (2007) for more

details. Then, to achieve minimal (in terms of the order) variance, we require the constrained estimator to be uncorrelated with $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$. If this condition is not met, the variance of the estimator $\widehat{\boldsymbol{Z}}$ would not be minimal. From a geometric perspective, since $E(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b) = \boldsymbol{0}$ (due to unbiasedness), the condition $Cov\left(\widehat{\boldsymbol{Z}}, \boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b\right) = \boldsymbol{0}$ implies that $E\left[\widehat{\boldsymbol{Z}}(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b)^T\right] = \boldsymbol{0}$, which means that $\widehat{\boldsymbol{Z}}$ and $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$ are orthogonal ("perpendicular" to each other). Thus, this uncorrelation ensures that the information of $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$ is maximally utilized and the estimator should be the optimal one. From an algebraic perspective, assume that there is another different linear unbiased estimator of $\boldsymbol{Z}$ that is correlated to $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$:

$$\widehat{\boldsymbol{Z}}' = \boldsymbol{Z}^b + \boldsymbol{L}(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b) = \widehat{\boldsymbol{Z}} + \boldsymbol{L}'(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b)$$

where $\boldsymbol{L}$ and $\boldsymbol{L}'$ are two linear operators and $\boldsymbol{L}' \neq \boldsymbol{0}$. Then we can show its variance is not less than the variance of $\widehat{\boldsymbol{Z}}$:

$$Var\left(\widehat{\boldsymbol{Z}}'\right) = Var\left(\widehat{\boldsymbol{Z}} + \boldsymbol{L}'(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b)\right) = Var\left(\widehat{\boldsymbol{Z}}\right) + Var\left(\boldsymbol{L}'\left(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b\right)\right) \succcurlyeq Var\left(\widehat{\boldsymbol{Z}}\right).$$

Therefore, for a linear estimator to be optimal, it must be uncorrelated with $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$. The condition is also sufficient since $\widehat{\boldsymbol{Z}}'$is arbitrary. By employing the covariance adjustment technique, we can identify the unique linear combination of $\boldsymbol{Z}^b$ and $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$ that is uncorrelated with $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$, which is equation (S3). For more technical details, please refer to Chapter Three and Seven of Sengupta and Jammalamadaka (2020). Note that even if the "observation" and "background" are correlated, which is relevant to sequential KF when the observations are mutually correlated, the validity of this equation persists, albeit with some necessary modifications. This is because the new information introduced by the constraint can be represented by a more complicated term than $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$. However, we will not derive the modified equation here.

Equation (S3) can be simplified since the observation $\boldsymbol{y}^o$ and the background $\boldsymbol{Z}^b$ are uncorrelated (i.e., $Cov(\boldsymbol{y}^o, \boldsymbol{Z}^b) = \boldsymbol{0}$):

$$\widehat{\boldsymbol{Z}} = \boldsymbol{Z}^b + Cov(\boldsymbol{Z}^b, \boldsymbol{P}\boldsymbol{Z}^b)[Var(\boldsymbol{P}\boldsymbol{Z}^b) + Var(\boldsymbol{y}^o)]^{-1}\left(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b\right) \tag{S4}$$

Furthermore, insert the covariance matrices $\boldsymbol{B}^b = Var(\boldsymbol{Z}^b) = Cov(\boldsymbol{Z}^b, \boldsymbol{Z}^b)$ and $\boldsymbol{R} = Var(\boldsymbol{y}^o)$, we have

$$\widehat{\boldsymbol{Z}} = \boldsymbol{Z}^b + \boldsymbol{B}^b\boldsymbol{P}^T(\boldsymbol{P}\boldsymbol{B}^b\boldsymbol{P}^T + \boldsymbol{R})^{-1}(\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b) \tag{S5}$$

which is the widely used form of KF (where the projection $\boldsymbol{P}$ is usually replaced with a more general linear operator $\boldsymbol{H}$ or, in a more general case, $\boldsymbol{H}\boldsymbol{M}$) and in this form the equation can be treated as the analytical solution to the cost function of 4D-Var when the model and observation operator is linear (Evensen et al., 2022). $\boldsymbol{B}^b\boldsymbol{P}^T(\boldsymbol{P}\boldsymbol{B}^b\boldsymbol{P}^T + \boldsymbol{R})^{-1}$ is usually referred to the Kalman gain matrix and denoted by $\boldsymbol{K}$, while $\boldsymbol{y}^o - \boldsymbol{P}\boldsymbol{Z}^b$ is known as the innovation vector (Asch et al., 2016). When the matrix $\boldsymbol{P}\boldsymbol{B}^b\boldsymbol{P}^T + \boldsymbol{R}$ is singular, its inverse needs to be replaced with its pseudo-inverse such as the Moore–Penrose inverse. In sequential KF, $\boldsymbol{P}\boldsymbol{B}^b\boldsymbol{P}^T + \boldsymbol{R}$ could be a scalar and always invertible (Houtekamer and Mitchell, 2001).

Also, we can write $\widehat{\boldsymbol{Z}}$ in another form

$$\widehat{\boldsymbol{Z}} = \boldsymbol{Z}^b + Cov(\boldsymbol{Z}^b, \boldsymbol{y}^b)[Var(\boldsymbol{y}^b) + \boldsymbol{R}]^{-1}(\boldsymbol{y}^o - \boldsymbol{y}^b), \tag{S6}$$

which can be easily evaluated in EnKF.

Given the updated state variable, we can also update the covariance matrix

$$\widehat{\boldsymbol{B}} = Var(\widehat{\boldsymbol{Z}}) = \boldsymbol{B}^b - \boldsymbol{B}^b\boldsymbol{P}^T(\boldsymbol{P}\boldsymbol{B}^b\boldsymbol{P}^T + \boldsymbol{R})^{-1}\boldsymbol{P}\boldsymbol{B}^b$$

or

$$\widehat{\boldsymbol{B}} = \boldsymbol{B}^b - Cov(\boldsymbol{Z}^b, \boldsymbol{y}^b)[Var(\boldsymbol{y}^b) + \boldsymbol{R}]^{-1} Cov(\boldsymbol{y}^b, \boldsymbol{Z}^b)$$

In data assimilation literature, $\widehat{\boldsymbol{Z}}$ and $\widehat{\boldsymbol{B}}$ are generally written as $\boldsymbol{Z}^a$ and $\boldsymbol{B}^a$, where a stands for analysis.

This derivation does not require a normality assumption since it relies solely on the covariance matrix and no information about the PDF behind is needed. However, the normality will influence the interpretation of "optimal" estimation. When the PDFs of $\boldsymbol{\varepsilon}^b$ and $\boldsymbol{\varepsilon}^o$ are normal, it can be shown that the BLUE is also the estimator with minimum variance among all unbiased estimators (MVUE), not necessarily among linear unbiased estimators (Chen, 2009; Rasch and Schott, 2018). Conversely, this indicates that when the PDFs are not normal, the BLUE may not necessarily be "optimal" and there may be a nonlinear estimator that achieves smaller variances. Nevertheless, there is no simple way to construct a nonlinear estimator, that is, to find a nonlinear function f such that $\boldsymbol{Z}^a = f(\boldsymbol{Z}^b, \boldsymbol{y}^o)$, even in such a simple linear statistical model (not to mention that a MVUE may not exist at all).

Linear estimation theory primarily addresses the updates of state variables and the covariance matrix, which represent only part of the KF. In addition to the update step, the KF also includes a prediction step. The prediction of state variables and covariance matrix is accomplished through the dynamical model. State variables can be forwarded by the dynamical model straightforwardly (or provided independently if they are parameters of the dynamical model), while the forward of background covariance matrix can present more significant challenges. When the model is linear (denoted as $\boldsymbol{M}$), predicting the background covariance matrix for the next time step is straightforward because covariance is also a linear operator. We have

$$\boldsymbol{B}_{k+1}^b = Var\left(\boldsymbol{Z}_{k+1}^b\right) = Var(\boldsymbol{M}\boldsymbol{Z}_k^a + \boldsymbol{\varepsilon}_k^m) = \boldsymbol{M}\boldsymbol{B}_k^a\boldsymbol{M}^T + \boldsymbol{Q}_k$$

where $\boldsymbol{\varepsilon}^m$ is the model error and $\boldsymbol{Q}$ its covariance matrix which, with other sources of errors, are usually absorbed into $\boldsymbol{R}$ and are empirically estimated. Here we also assume that the model error and the linear estimation error are uncorrelated. Subscript k indicates discrete time steps.

When the model is not linear, however, predicting $\boldsymbol{B}$ becomes more challenging. Over the past few decades, two popular methods have been developed and widely applied to address this issue. The first method involves linearizing the model and use the linearized version to forward the covariance matrix. This approach leads to the development of the extended KF (EKF). The second method focuses on propagating the probability density function (PDF) of the state rather than directly forwarding the covariance matrix (Evensen, 1994). In this approach, the PDF of a random state is represented by an ensemble of states each of which is evolved using the full dynamical model. The resulting ensemble of predicted states is then used to compute the predicted covariance matrix. This is method is known as EnKF, which is the focus of this study.

S2. Illustrative inversion experiments for consistency with diagnostic results

To complement the correlation-based diagnostic analysis presented in the main text, we conducted a set of simplified inversion experiments within an observing system simulation experiment (OSSE) framework. These experiments are not intended as a comprehensive evaluation of inversion system performance, but rather as a qualitative consistency check of the mechanisms identified in this study.

Experimental setup

The inversion system is built on the Community Multiscale Air Quality model (CMAQ) with an offline version of the Vegetation Photosynthesis and Respiration Model (VPRM), and implemented within the Data Assimilation Research Testbed (DART) framework (Anderson et al., 2009; Data Assimilation Research Section, 2022). To reduce computational cost, the inversions cover a two-month period (November–December 2015), with analysis focused on the second month. An ensemble size of 100 is used. We adopted the traditional fixed-lag smoother. Following previous studies (Liu et al., 2019; Kong et al., 2022), we used a seven-day lag. Each assimilation cycle, also the atmospheric transport model cycle, advanced by one day. At each assimilation cycle, the one-day observations are assimilated and the fluxes within the seven-day smoother lag. In addition, the fluxes and initial concentrations for the next cycle were also updated. A localization function with a horizontal length scale of 1000 km is applied (Gaspari and Cohn, 1999). Hourly fluxes are inverted but only the monthly mean values are analyzed.

The OSSE setup assumes identical meteorological fields and initial and boundary conditions between the nature run (the "true" atmosphere) and the inversions. All the fluxes (F_{BIO}, F_{ANT}, and F_{OCE}) in the nature run differed from those used as priors in the inversion system. Synthetic observations include surface observations and satellite data, all of which are constructed using real observational metadata (e.g., time, location, and uncertainty) drawn from valid entries in datasets such as ObsPack, GOSAT, and OCO-2.

Four experiments are performed: (1) OFF100surf using 100 km correlation length and surface observations only; (2) OFF600surf using 600 km correlation length and surface observations only; (3) OFF100 using 100 km correlation length and surface and satellite observations; (4) OFF600 using 600 km correlation length and surface and satellite observations.

Results

Figure S9 illustrates the OSSE results. The objective of the inversion is to adjust the prior flux toward the true flux. Accordingly, in the OSSE, the posterior correction (posterior minus prior, Figs. S9h-k) should resemble the "true innovation" (true minus prior, Fig. S9l). When only surface observations are used, the inversion behavior differs markedly between the two correlation length settings. In OFF100surf, the posterior flux remains very close to the prior (Figs. S9a,b), and the posterior correction is both weak and spatially localized near observation sites (Fig. S9h). This indicates that observational constraints are unable to propagate effectively beyond immediate site neighborhoods. The spatial distribution of corrections closely follows the observation network, confirming that the inversion influence is highly localized. In contrast, in OFF600surf, the posterior flux departs more substantially from the prior and moves closer to the true flux (Figs. S9a,c,f). The posterior corrections are larger in magnitude and spatially more coherent, and show improved similarity to the true innovation, particularly near observation sites (Figs. S9i,l). This comparison demonstrates that, under sparse observational coverage, short correlation lengths strongly limit the spatial propagation of observational constraints, leading to weak and localized flux adjustments.

The inclusion of satellite observations increases the overall constraint on fluxes, but does not fundamentally change the role of correlation length. In OFF100, satellite data provide additional constraints, particularly in South and Southeast Asia where sampling density is relatively high (Fig. S9j). In these regions, the inversion is able to recover some aspects of the true flux even with a short correlation length. However, in regions with strong true innovations (e.g., China and Russia), the inversion

remains poorly constrained under the 100 km setting. The posterior corrections are still weak and fail to reproduce the true innovation pattern. In OFF600, the addition of satellite data further improves the inversion (Fig. S9k). The posterior corrections become both stronger and more spatially consistent with the true innovation, especially across southern and southwestern China in addition to South and Southeast Asia. These results indicate that while denser satellite observations can partially compensate for limited information propagation, they cannot fully overcome the constraints imposed by an overly short prior correlation length.

To further assess inversion performance, we compare posterior fluxes directly with the true flux using mean bias (MB) and root mean square error (RMSE). In OFF100surf, both MB and RMSE remain nearly unchanged relative to the prior, indicating minimal improvement. In OFF600surf, MB is reduced by 44% and RMSE by 9%. With satellite data, OFF100 shows moderate improvement (MB reduced by 53%, RMSE by 17%). In OFF600, the best performance is achieved, with MB (absolute values) reduced by 88% and RMSE by 32%.

In summary, these two-month inversion experiments support the main conclusions of the diagnostic analysis that short spatial autocorrelation lengths severely limit the ability of observations to constrain fluxes under sparse networks. When the correlation length increases, both the magnitude and spatial extent of posterior corrections are enhanced. Satellite observations improve constraint but cannot fully compensate for insufficient correlation structure. Overall, the results provide independent, inversion-based evidence that prior autocorrelation structure critically controls the effectiveness of observational constraint in CO_2 flux inversions.

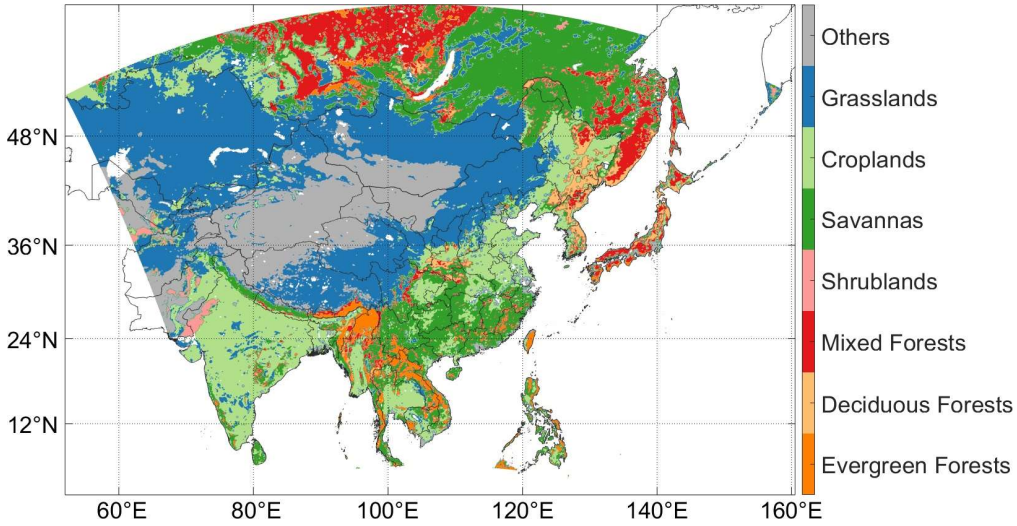

Fig. S1. The seven dominate land covers used by WRF-VPRM in the domain of this study.

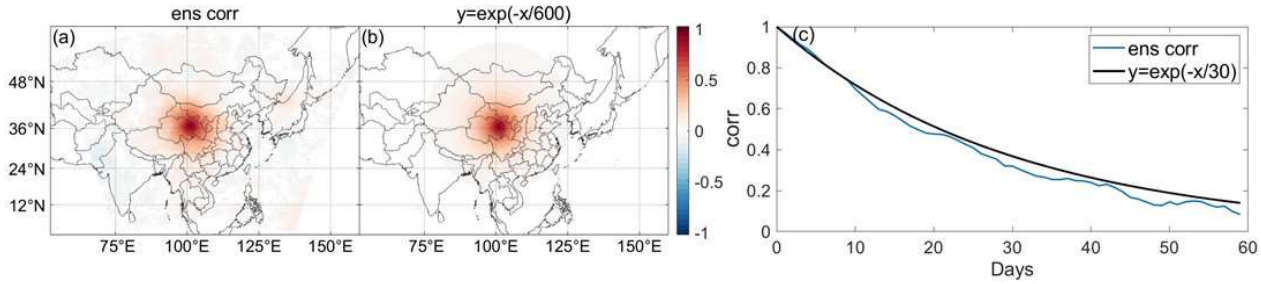

Fig. S2. Correlation structures of flux at the WLG site. (a) Spatial correlations from ensembles at the first day, (b) ideal spatial correlation of an exponential function, and (c) temporal correlations from ensembles and from an exponential function.

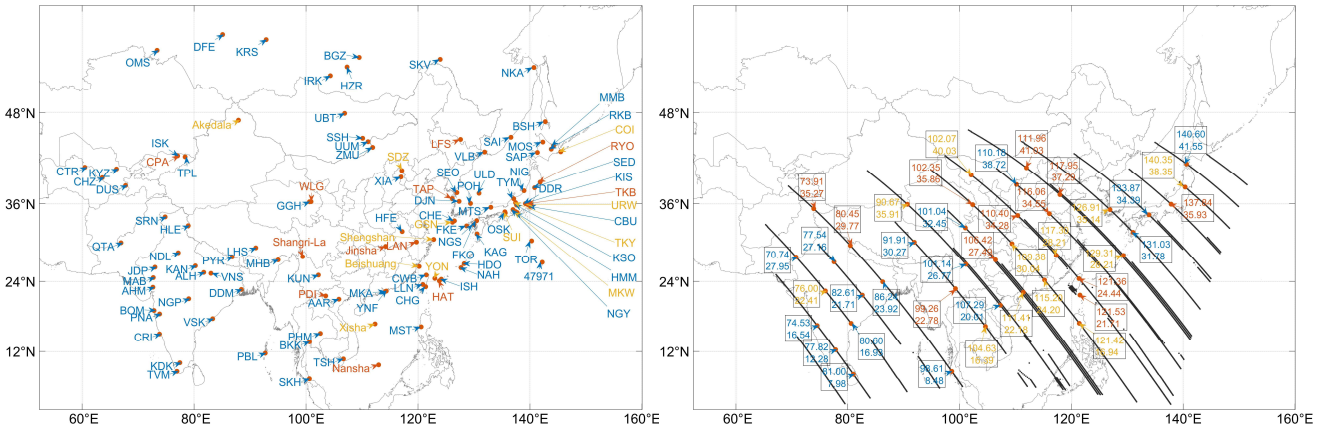

Fig. S3. Spatial distributions of CO_2 observation sites (a) and XCO_2 locations (b) used in this study. Orange and yellow colors indicate the 24 sites or locations used to present the site-specific results.

155

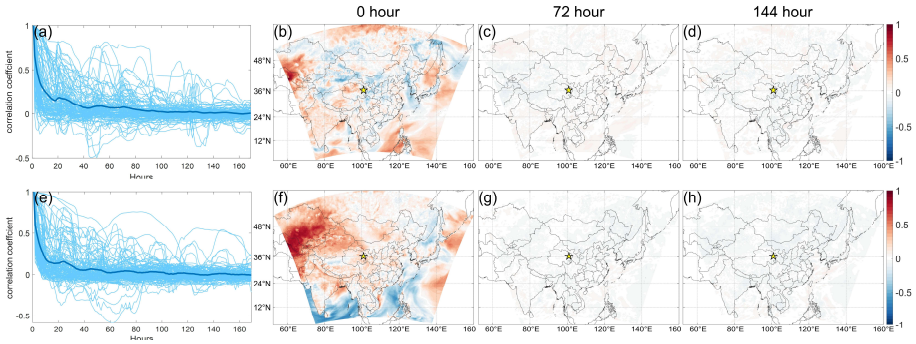

Fig. S4. (a) Temporal variations of correlations between modelled surface CO_2 and initial concentrations (IC) at the 111 sites in Fig. S2. The bold line is their mean. (b)–(d) Spatial distributions of correlations between surface CO_2 at the continental background site WLG (yellow star) and IC in the whole domain. The modelled surface CO_2 is at the initial time (b), the 72th hour (c), and the 144th hour (d) starting from
160 06:00 UTC 1 January 2016, respectively. (e)–(h) are the same as (b)–(d) but for 7 January 2016.

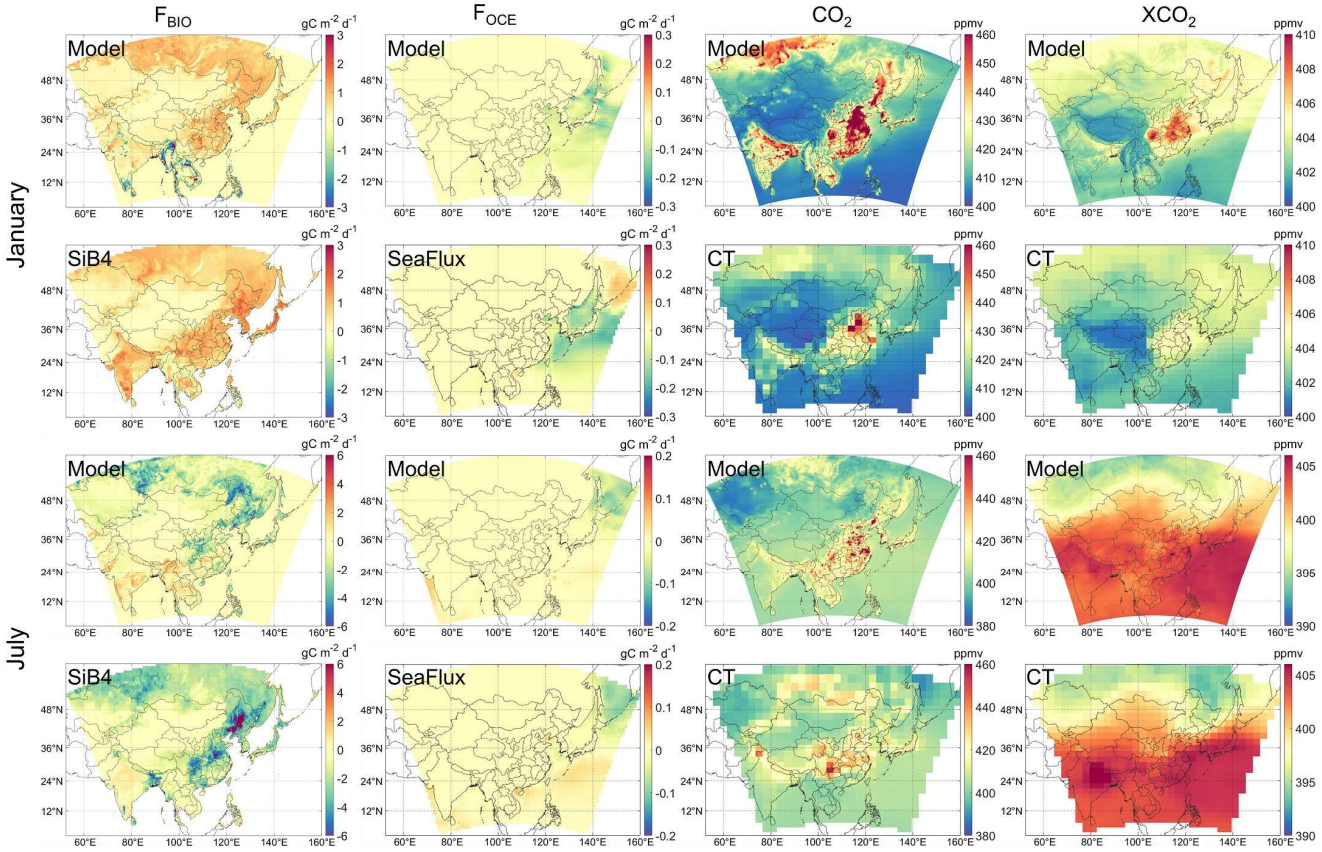

Fig. S5. Comparison between our modelled values with other widely used datasets: F_{BIO} from SiB4.2 (Haynes et al., 2021), F_{OCE} from SeaFlux (Fay et al., 2021), surface CO_2 and XCO_2 from CT (Jacobson et al., 2023). The first two rows are for January, and the last two rows for July. All data are weekly average, except for SeaFlux that are monthly.

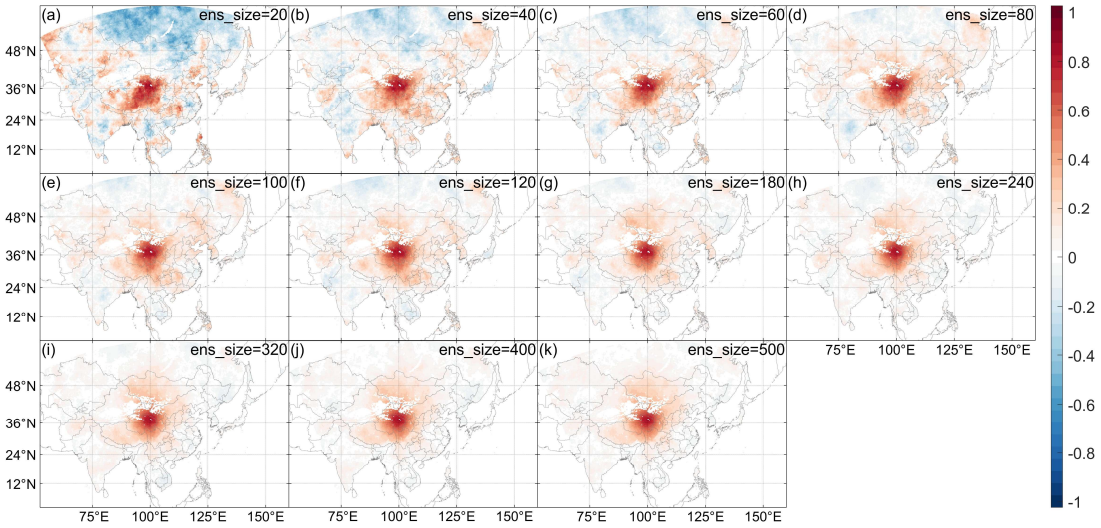

Fig. S6. Correlation patterns at the WLG site. Different panels have different ensemble sizes labeled at the right upper corner of the panels.

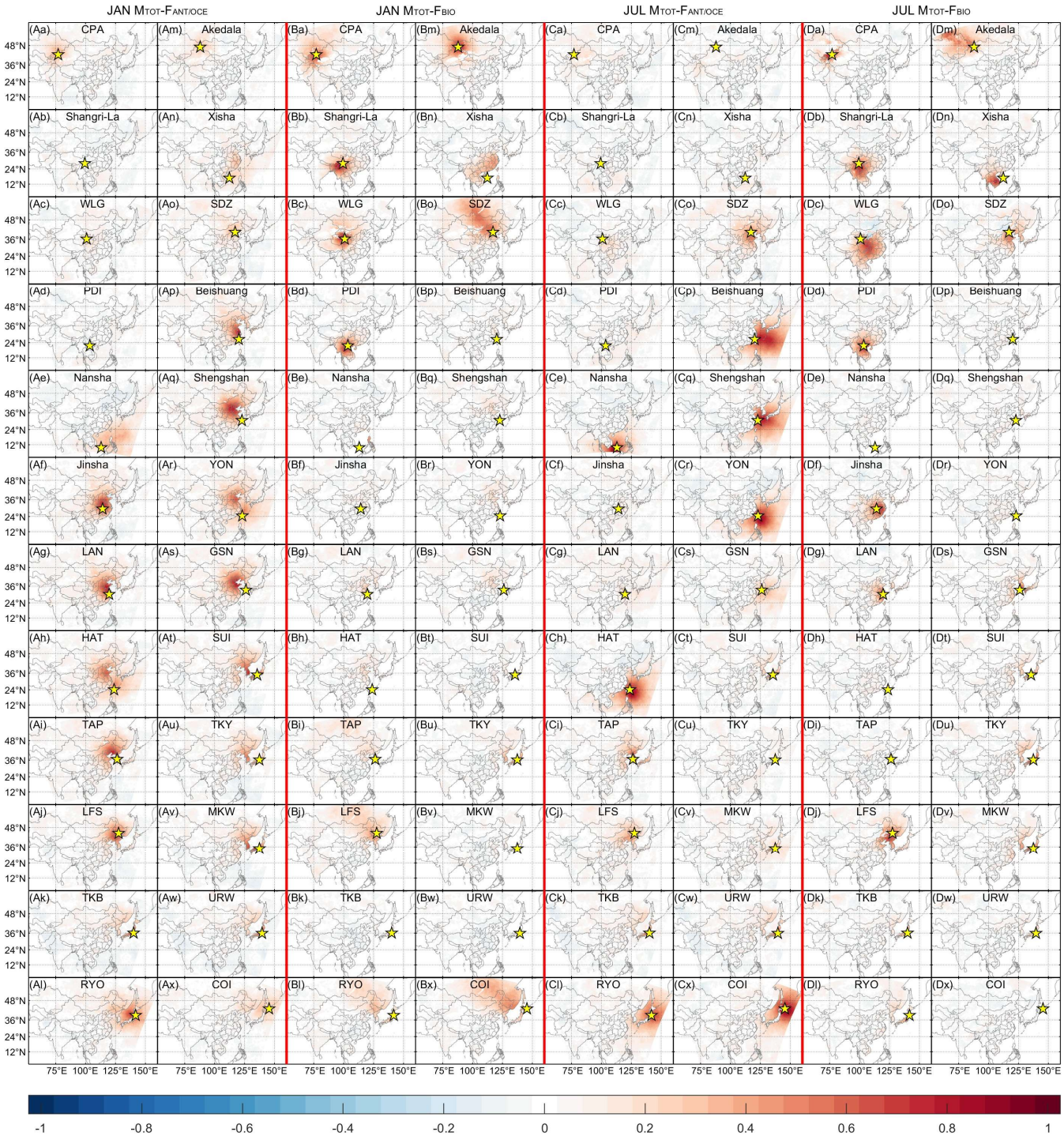

Fig S7. Spatial patterns of 500-member correlations between the 168th-hour M_{TOT} at sites and the zeroth-hour fluxes (F_{ANT}, F_{OCE}, and F_{BIO}) across domain. (A) includes the 24 sites shown in Fig. S3 and is for F_{ANT} and F_{OCE} in January. (B) is for F_{BIO} in January. (C) and (D) are for July.

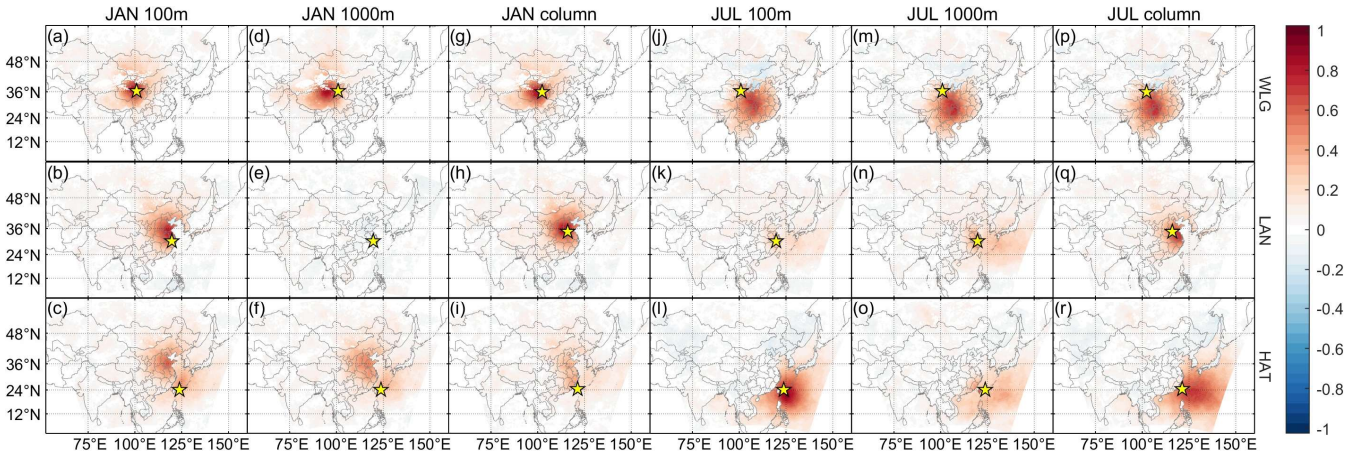

Fig. S8. Spatial patterns of correlations for the M_{TOT} at different heights. (a)–(c) ~100 m in January. (d)–(f) ~1000 m in January. (g)–(i) XCO_2 in January. (j)–(r) July. The first row is for F_{BIO} and the others for F_{ANT} and F_{OCE}. The XCO_2 site locations are not identical with surface site locations.

175

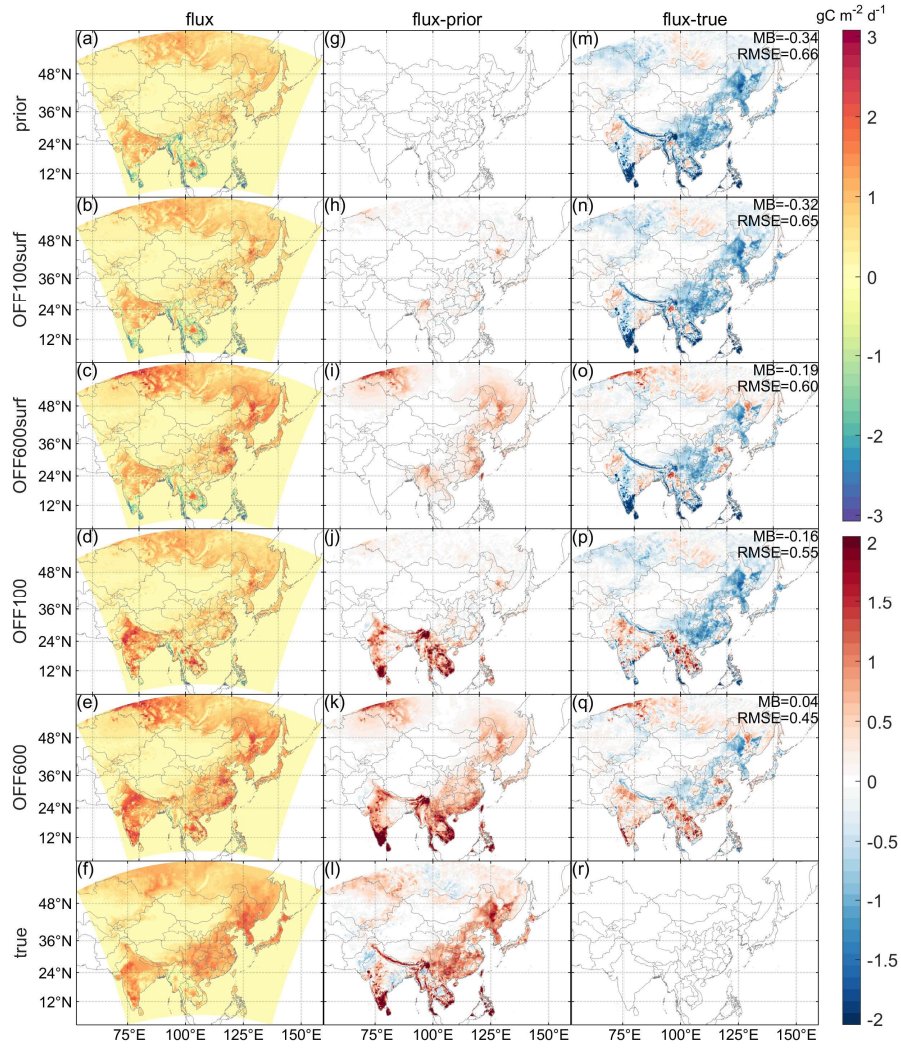

Fig. S9. Monthly biospheric fluxes in the inversions. (a) Prior flux. (b)–(e) Posterior fluxes. (f) True flux. The second column presents differences from the prior (a); the third column presents differences from the true flux (f). Values in the third column panels are the mean bias (MB) and root mean square error (RMSE) across all land grids.

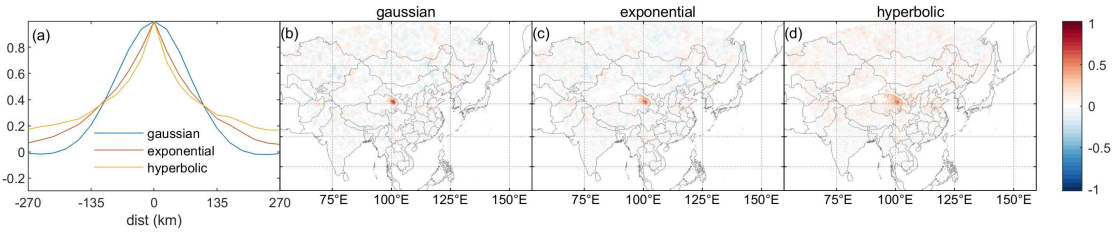

Fig. S10. Influence of correlation function shape. (a) Spatial autocorrelations along latitude centered at the WLG site. Lines are averaged from 21 lines with centers near the WLG site for smoothness. Colors denote correlation functions: blue—Gaussian, orange—exponential, yellow—hyperbolic (1/(1+d)). (b)–(d) Spatial patterns of M_{TOT}–F_{BIO} correlations using (b) Gaussian, (c) exponential, and (d) hyperbolic format functions. The experiments are conducted using CMAQ instead of WRF-VPRM, and the temporal coverage is from November 1 to November 8, 2015, without perturbing the initial conditions.

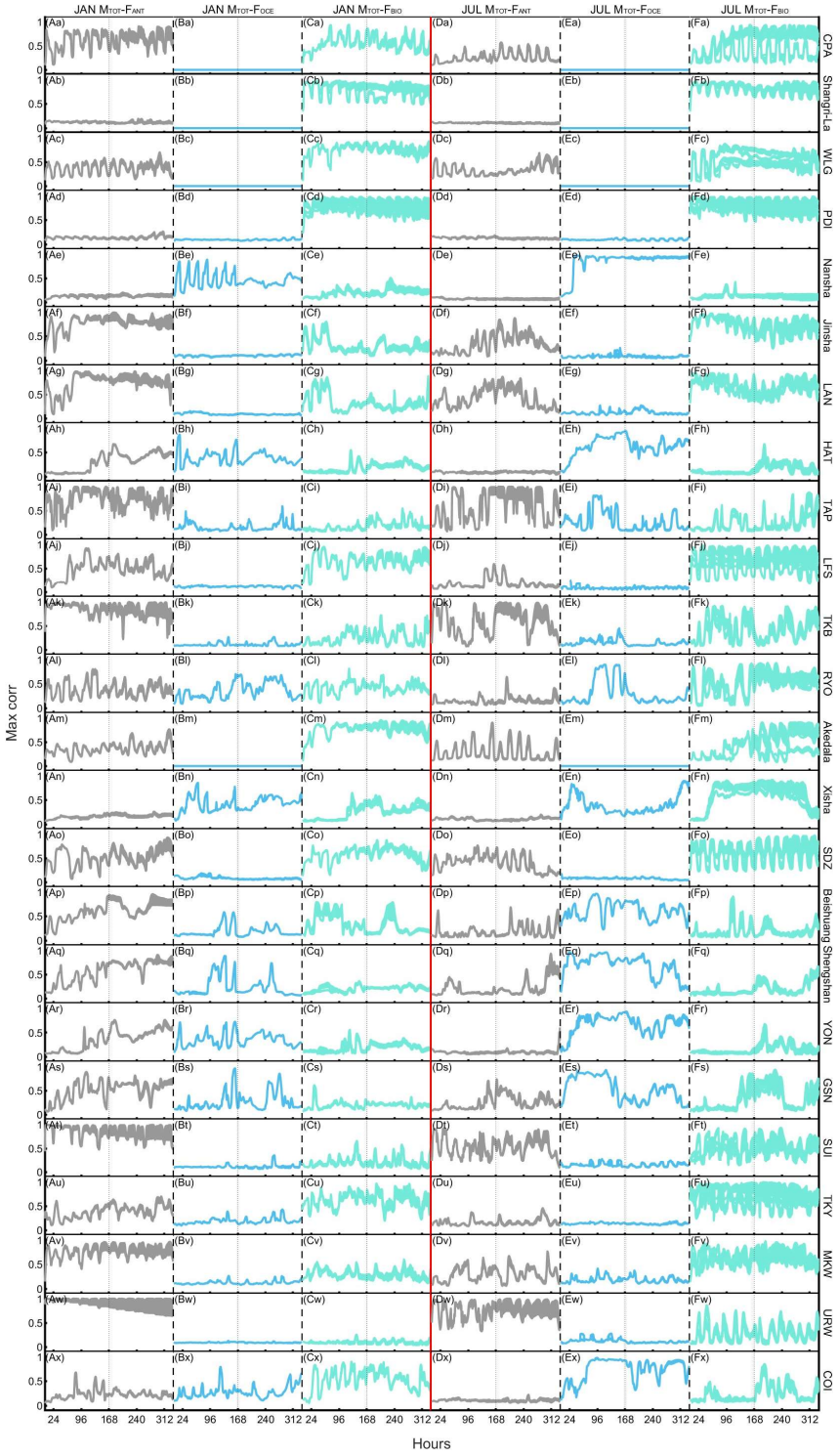

Fig. S11. Timeseries of maximum correlations the 168th-hour M_{TOT} at sites and the zeroth-hour fluxes (F_{ANT}, F_{OCE}, and F_{BIO}) near the sites. Each row represents a site, labeled at right. (A) M_{TOT}-F_{ANT} in January; (B) M_{TOT}-F_{OCE} in January; (C) M_{TOT}-F_{BIO} in January; (D)–(F) July.

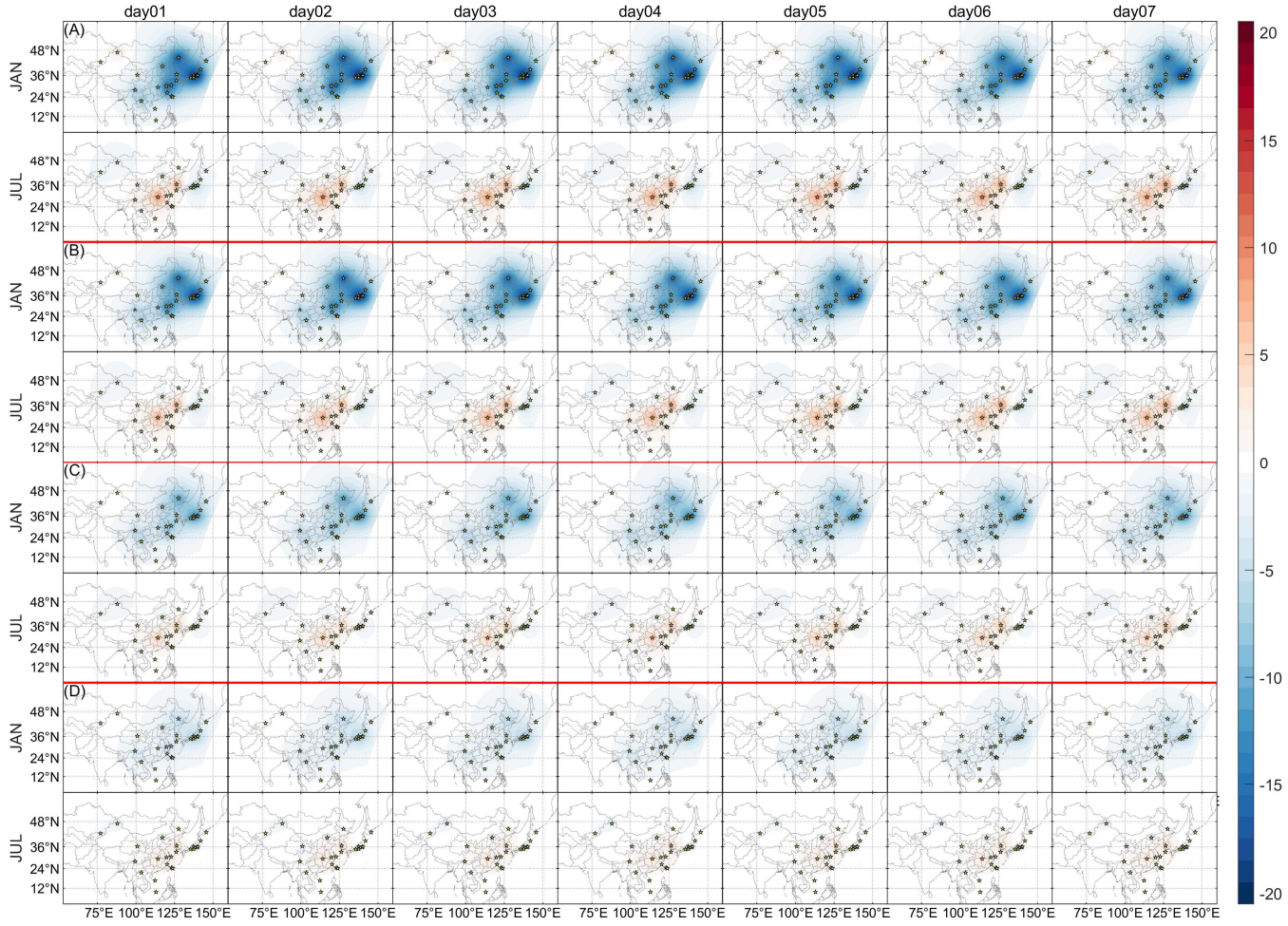

Fig. S12. The same as Fig. 6B (the spatial patterns of $\boldsymbol{B}\boldsymbol{g}_0$) but the gradient $\boldsymbol{g}_0$ keeps only values near the observation sites. A keeps a 15×15 grids box, B 9×9, C 3×3, and D only one grid. Each grid is of size 27 km × 27 km.

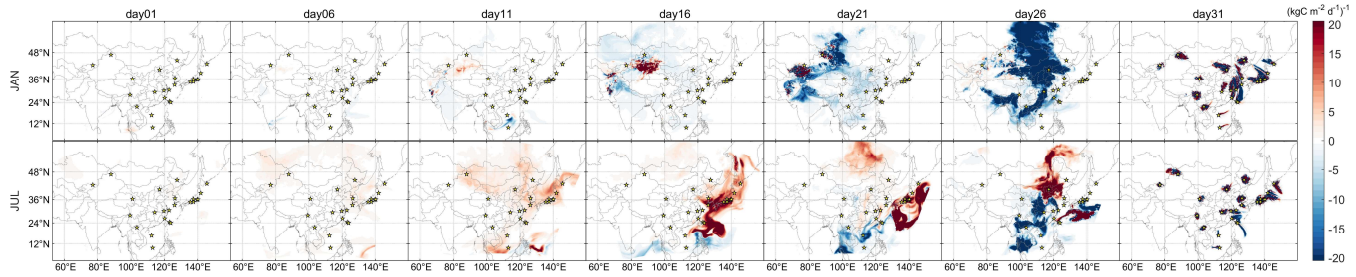

Fig. S13. The same as Fig. 6C but the cost function uses only concentrations at the last hour. Note that the unit is smaller than that of Fig. 6C by three orders of magnitude.

References

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A.: The Data Assimilation Research Testbed: A Community Facility, Bulletin of the American Meteorological Society, 90, 1283–1296, https://doi.org/10.1175/2009bams2618.1, 2009.

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Monthly Weather Review, 129, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129%3C2884:Aeakff%3E2.0.Co;2, 2001.

Asch, M., Bocquet, M., and Nodet, M.: Data Assimilation: Methods, Algorithms, and Applications, Society for Industrial and Applied Mathematics, 2016.

Chen, X.: Advanced Mathematical Statistics (in Chinese), University of Science and Technology of China Press, 2009.

Data Assimilation Research Section: DART Release 10.5.1, 2022.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10143–10162, https://doi.org/10.1029/94JC00572, 1994.

Evensen, G., Vossepoel, F. C., and Van Leeuwen, P. J.: Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-96709-3, 2022.

Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., Iida, Y., Laruelle, G. G., Rödenbeck, C., Roobaert, A., and Zeng, J.: SeaFlux: harmonization of air–sea CO_2 fluxes from surface pCO_2 data products using a standardized approach, Earth System Science Data, 13, 4693–4710, https://doi.org/10.5194/essd-13-4693-2021, 2021.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, https://doi.org/10.1002/qj.49712555417, 1999.

Haynes, K. D., Baker, I. T., and Denning, A. S.: SiB4 Modeled Global 0.5-Degree Hourly Carbon Fluxes and Productivity, 2000-2018, , https://doi.org/10.3334/ORNLDAAC/1847, 2021.

Houtekamer, P. L. and Mitchell, H. L.: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 129, 123–137, https://doi.org/10.1175/1520-0493(2001)129%3C0123:ASEKFF%3E2.0.CO;2, 2001.

Jacobson, A. R., Kenneth N. Schuldt, Pieter Tans, Arlyn Andrews, John B. Miller, Tomohiro Oda, Sourish Basu, John Mund, Brad Weir, Lesley Ott, Tuula Aalto, James Brice Abshire, Ken Aikin, Shuji Aoki, Francesco Apadula, Sabrina Arnold, Bianca Baier, Jakub Bartyzel, Andreas Beyersdorf, Tobias Biermann, Sebastien C. Biraud, Harald Boenisch, Gordon Brailsford, Willi A. Brand, Gao Chen, Huilin Chen, Lukasz Chmura, Shane Clark, Aurelie Colomb, Roisin Commane, Sébastien Conil, Cédric Couret, Adam Cox, Paolo Cristofanelli, Emilio Cuevas, Roger Curcoll, Bruce Daube, Kenneth J. Davis, Stephan De Wekker, Julian Della Coletta, and et al: CarbonTracker CT2022, 2023.

Kong, Y., Zheng, B., Zhang, Q., and He, K.: Global and regional carbon budget for 2015-2020 inferred from OCO-2 based on an ensemble Kalman filter coupled with GEOS-Chem, Atmospheric Chemistry and Physics, 22, 10769–10788, https://doi.org/10.5194/acp-22-10769-2022, 2022.

Liu, Y., Kalnay, E., Zeng, N., Asrar, G., Chen, Z., and Jia, B.: Estimating surface carbon fluxes based on a local ensemble transform Kalman filter with a short assimilation window and a long observation window: an observing system simulation

experiment test in GEOS-Chem 10.1, Geoscientific Model Development, 12, 2899–2914, https://doi.org/10.5194/gmd-12-2899-2019, 2019.

Rao, C. R.: Some Statistical Problems in Multitarget Tracking, in: Statistical Decision Theory and Related Topics V, 513–521, https://doi.org/10.1007/978-1-4612-2618-5_40, 1994.

Rasch, D. and Schott, D.: Mathematical statistics, John Wiley & Sons, 2018.

Seber, G. A.: A matrix handbook for statisticians, John Wiley & Sons, 2007.

Sengupta, D. and Jammalamadaka, S. R.: Linear models and regression with R: an integrated approach, second., World Scientific, 2020.