Diagnosing spatial error structures in CO_2 mole fractions and XCO_2 column mole fractions from atmospheric transport

Thomas Lauvaux[1,a]**, Liza I. Díaz-Isaac**[1,b]**, Marc Bocquet**[2]**, and Nicolas Bousserez**[3,c]

[1]Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA
[2]CEREA, joint laboratory École des Ponts ParisTech and EDF R& D, Université Paris-Est, Champs-sur-Marne, France
[3]University of Colorado Boulder, Boulder, CO
[a]Now at Laboratoire des Sciences du Climat et de l'Environnement, CEA, CNRS, UVSQ/IPSL, Université Paris-Saclay, Orme des Merisiers, 91191 Gif-sur-Yvette cedex
[b]Now at Scripps Institution of Oceanography, University of California, San Diego, 92093, USA
[c]Now at European Centre for Medium-Range Weather Forecasts, Reading, UK

Correspondence to: Thomas Lauvaux (thomas.lauvaux@lsce.ipsl.fr)

Abstract. Atmospheric inversions inform about the magnitude and variations of greenhouse gas (GHG) sources and sinks from global to local scales. Deployment of observing systems such as spaceborne sensors and ground-based instruments distributed around the globe has started to offer an unprecedented amount of information to estimate surface exchanges of GHG at finer spatial and temporal scales. However, all inversion methods still rely on imperfect atmospheric transport models whose error structures directly affect the inverse estimates of GHG fluxes. The impact of spatial error structures on the retrieved fluxes increase concurrently with the density of the available measurements. In this study, we diagnose the spatial structures due to transport model errors affecting modeled in situ carbon dioxide (CO_2) mole fractions and total column dry air mole fractions of CO_2 (XCO_2). We implement a cost-effective filtering technique recently developed in the meteorological data assimilation community to describe spatial error structures using a small-size ensemble. This technique can enable ensemble-based error analysis for multi-year inversions of sources and sinks. The removal of noisy structures due to sampling errors in our small-size ensembles is evaluated by comparison to larger-size ensembles. A second filtering approach for error covariances is proposed (Wiener filter), producing similar results over the 1-month simulation period compared to a Schur filter. Based on a comparison to a reference 25-member calibrated ensemble, we demonstrate that error variances and spatial error correlation structures are recoverable from small-size ensembles of about eight (8) to ten (10) members, improving the representation of transport errors in mesoscale inversions of CO_2 fluxes. Moreover, error variances of in situ near-surface and free-tropospheric CO_2 mole fractions differ significantly from total column XCO_2 error variances. We conclude that error variances for remote sensing observations need to be quantified independently of in situ CO_2 mole fractions due to the complexity of spatial error structures at different altitudes. However, we show the potential use of meteorological error structures such as the mean horizontal wind speed, directly available from Ensemble Prediction Systems, to approximate spatial error correlations of in situ CO_2 mole fractions, with similarities in seasonal variations and characteristic error length scales.

1 Introduction

Atmospheric carbon dioxide (CO_2) mole fraction has been increasing steadily since the first industrial revolution, primarily due to fossil fuel emissions and land use change (Ciais et al., 2015). Recent estimates of sources and sinks at the global scale suggest a coincidental reinforcement of natural sinks balancing the continuously increasing anthropogenic emissions (Le Quéré et al., 2016; Keenan et al., 2016). Therefore, the fraction of fossil fuel CO_2 remaining in the atmosphere was kept constant at 2ppm per year[1], excluding short-time anomalies such as El Niño events (Feely et al., 1999; Kim et al., 2016). In the objective of characterizing the natural sink mechanisms, atmospheric inversion methods have provided some evidences of a fertilization effect possibly increasing the effective absorption by plants of the exceeding CO_2 in the atmosphere (Schimel et al., 2015). But large uncertainties still affect atmospheric inversions of CO_2 fluxes and limit the interpretation of continental-scale CO_2 budgets (Peylin et al., 2013). Therefore, more robustness in these findings first requires better characterization of the error affecting inverse estimates (Baker et al., 2007; Stephens et al., 2007; Díaz-Isaac et al., 2014).

Atmospheric inversions of Greenhouse Gases (GHG) are now widely used to infer surface fluxes from natural (e.g. Enting, 2002; Gurney et al., 2002; Lauvaux et al., 2012; Peylin et al., 2013) and anthropogenic (e.g. McKain et al., 2015; Lauvaux et al., 2016) sources at global, regional, and local scales. However, key information in carbon cycle science lies in multi-year time scales, therefore confining the development of inverse methodologies to cost-effective approaches (e.g. Bruhwiler et al., 2005). Based on similar methodologies than those of meteorology or geophysics, atmospheric inversions have used primarily fast approaches to produce multi-decadal fluxes such as variational approaches (Baker et al., 2006; Chevallier et al., 2010), avoiding large ensemble of simulations based on Monte Carlo formulation (Evensen, 1994). In parallel, assumptions made in prior flux errors and transport errors impact the inverse solution in similar ways (Engelen et al., 2002). Concerning the prior flux errors, few studies have proposed to constrain the spatial and temporal structures more rigorously (Wu et al., 2013; Ganesan et al., 2014), some of them based on terrestrial biogeochemical models and eddy-covariance flux measurements to estimate the spatial structures in the prior flux errors of CO_2 (e.g. Chevallier et al., 2006; Hilton et al., 2013). For transport errors, correlations remained small at the global scale, primarily due to sparse atmospheric GHG observation networks. However, denser tower networks (Andrews et al., 2014) and recent satellite missions have significantly increased the sampling density (e.g. the Greenhouse gases Observing SATellite (GOSAT; Yokota et al. (2009); Houweling et al. (2015)) and the Orbital Carbon Observatory (OCO-

2) missions (Crisp et al., 2004)) requiring the characterization of their correlated errors in inversion systems.

The increased density in existing tower networks and the availability of fine-scale satellite retrievals raised concerns about spatial and temporal structures in transport model errors (Rayner and O'Brien, 2001; Lauvaux et al., 2009; Miller et al., 2015). The proximity of the measurements (e.g. couple kilometers between OCO-2 retrievals) means that spatial correlations in model errors are significant and can no longer be ignored (Chevallier, 2007). This issue becomes critical to greenhouse gas inversion problems when applied to urban scales (Lauvaux et al., 2016) but remains poorly studied to date. Recent deployment of path-integrated instruments also increased the complexity of the problem from the ground when trying to invert for emissions from single facilities such as a large dairy (Viatte et al., 2017).

Ensemble approaches are useful to describe flow-dependent errors (e.g. Anderson, 2001; Evensen, 2003) but remain computationally expensive due to the number of model simulations required to correctly represent model error statistics (Houtekamer and Mitchell, 1998). In general, a small number of members leads to incomplete descriptions of error structures which require the use of localization to avoid spurious correlations due to sampling errors (Houtekamer and Mitchell, 2001; Raynaud and Pannekoucke, 2013). But small-size ensembles are efficient computationally and able to provide information on flow-dependent error structures compared to prescribed static error structures (Brousseau et al., 2012). With the development of new perturbation methods, the number of members may decrease significantly thanks to optimal perturbations combining physics, parameter sensitivity and energy-based perturbations (Jankov et al., 2017). In any case, small-size ensembles remain affected by sampling noise which has to be removed before extracting spatial structures, either by modeling (Pannekoucke et al., 2008; Lauvaux et al., 2009) or by filtering unphysical structures (Hamill et al., 2001; Houtekamer and Mitchell, 2001). Here, we apply a newly developed approach based on local filtering and a localization technique (Ménétrier et al., 2015a; Ménétrier et al., 2015b). There are only a few approaches for the optimal localization of covariance matrices in the field of data assimilation for the geosciences (Lei and Anderson, 2014; Flowerdew, 2015; De La Chevrotière and Harlim, 2017). To our knowledge, the method is the only one so far which is both (i) mathematically consistent and (ii) a priori, i.e. not based on learning on past or present datasets. Besides, in spite of its sophistication, the filtering approach is rather straightforward to implement.

In this study, we apply the filter of variances and the covariance localization developed in Ménétrier et al. (2015a), and propose an additional solution to using the optimality condition, both for Gaussian and non-Gaussian error statistics cases. The filter is applied to several calibrated ensembles of different sizes to evaluate the impact of our filter on small (5 members) to larger (25 members) to the full ensem-

---

[1]www.esrl.noaa.gov/gmd/ccgg/trends/

ble (45 members) based on multi-physics simulations (Díaz-Isaac et al., 2018a). Results are presented for in situ CO_2 mole fractions, XCO_2 dry air mole fractions, mean horizontal winds, and

5 . We discuss the results in Section 4.

2 Methods

2.1 Calibration of WRF-CO_2 ensembles

We generate an ensemble using the Weather Research and Forecasting (WRF) model version 3.5.1 (Skamarock et al.,
10 2008), including the chemistry module modified in this study for CO_2 (Lauvaux et al., 2012). The ensemble consists of 45 members that were generated by varying the different physics parameterization and meteorological data. The land surface models, surface layers, planetary boundary layer schemes,
15 cumulus schemes, microphysics schemes, and meteorological data (i.e., initial and boundary conditions) are alternated in the ensemble (Díaz-Isaac et al., 2018b). All the simulations use the same radiation schemes, both long and shortwave. All simulations were run using the one-way nesting
20 method, with two nested domains. The coarse domain uses a horizontal grid spacing of 30 km and covers most of the United States and part of Canada. The inner domain uses a 10 km grid spacing, is centered in Iowa and covers the Midwest region of the United States. The vertical resolution of
25 the model is described with 59 vertical levels, with 40 of them within the first 2 km of the atmosphere. This work focuses on the WRF simulation with the higher resolution, therefore only the 10-km domain will be analyzed. Simulations were performed from 27 June, 2008 to 21 July 2008,
30 with a 10-day spin-up for initial conditions. The CO_2 fluxes for summer 2008 were obtained from NOAA Global Monitoring Division's CarbonTracker version 2009 (CT2009) data assimilation system (Peters et al. (2007), with updates documented at http://carbontracker.noaa.gov). The different
35 fluxes that CT2009 propagates into the models are fossil fuel burning, terrestrial biosphere exchange, and the exchange with oceans. The CO_2 lateral boundary conditions were obtained from CT2009 mole fractions. Only the meteorological transport fields vary between each model configuration or en-
40 semble member.

The ensemble was calibrated over the Midwest U.S. using the available meteorological observations and the 10-km model simulation as described in Díaz-Isaac et al. (2018a). The measurements used included
45 balloon soundings collected over the Midwest region (http://weather.uwyo.edu/upperair/sounding.html) for 14 rawinsonde stations. The ensemble was calibrated for three different meteorological variables: wind speed, wind direction and Planetary Boundary Layer Height (PBLH) in
50 the late afternoon data (i.e., 0000 UTC) from the different rawinsondes. Daytime data was used to represent well mixed conditions, at the selected time when CO_2 mole fraction

are assimilated in atmospheric inversions to avoid stable conditions near the surface. The calibration algorithm is described in Garaud and Mallet (2011), selecting optimal
55 ensembles of different sizes using Simulated Annealing and Genetic Algorithm techniques. The metric used in Díaz-Isaac et al. (2018a) is the flatness of the Rank Histogram which is a measure of the ensemble dispersion. By eliminating members with redundant information, smaller
60 ensembles were able to better match the variability in the observations. We refer to Díaz-Isaac et al. (2018a) for a full description of the calibration process and the final selection of optimal ensembles.

Here, we will compare the different ensembles generated
65 in Díaz-Isaac et al. (2018a) from 5 to 8 to 10 members. An additional ensemble was created for our study with a larger number of members in order to address the potential lack of representativeness of model errors with small-size ensembles. Therefore, we generated a 25-member ensemble and
70 applied the same calibration process. This ensemble has not been documented in Díaz-Isaac et al. (2018a) but is described here in the Appendix A. We compare the results of the filtering for small sizes to the 25-member calibrated ensemble instead of the original 45-member ensemble that was not cal-
75 ibrated.

2.2 Variance Filtering and Covariance Localization

Ménétrier et al. (2015a); Ménétrier et al. (2015b) have proposed a new theory for the optimal filtering of sample variances and covariances. These are defined by the following
80 empirical second-order moment statistics. Assume we have an ensemble of states $\mathbf{x}_k \in \mathbb{R}^n$ for $k = 1, \ldots, N$ of mean $\overline{\mathbf{x}}$, from which to infer the statistics. Define the associated anomalies $\delta\mathbf{x}^k = \mathbf{x}^k - \overline{\mathbf{x}}$, also called perturbations. Then, the sample covariance matrix is:
85

$$\widetilde{\mathbf{B}} = \frac{1}{N-1} \sum_{k=1}^{N} \delta\mathbf{x}^k \left(\delta\mathbf{x}^k\right)^{\mathrm{T}}, \tag{1}$$

which is an unbiased estimator of the true covariance matrix $\mathbf{B}^\star$, i.e., $\mathbb{E}\left[\widetilde{\mathbf{B}}\right] = \mathbf{B}^\star$ where $\mathbb{E}$ is the expectation operator over the reference distribution from which the $\mathbf{x}_k$ are sam-
90 pled. In the following, we denote by $\widetilde{B}_{ij}$ the entries of this covariance matrix.

Filtering of variances and covariances is made necessary because of the finite-size of the sample ensembles, which can generate significant sampling errors. The sampling errors can
95 be filtered out by applying a linear filter on the variances and covariances. The most general linear filter is of the form

$$\widehat{B}_{ij} = \sum_{kl} F_{ijkl} \widetilde{B}_{kl}, \tag{2}$$

where $\widehat{\mathbf{B}}$ is the filtered error covariance matrix. Typical examples are the application of a convolution to the vector of

variances to smooth them out, or the application of a Schur product with a non-degenerate, short-range correlation matrix to the sample covariance matrix. The linear filter often requires parameters, a correlation length typically, that must be tuned for the filter to be optimal.

The theory proposed in Ménétrier et al. (2015a) to achieve optimality of the filter is based on three key ingredients:

1. The first one consists in requiring that the residual sampling error be minimal. Assume that we have an estimator $\widetilde{\mathbf{x}}$ of some statistics of a reference distribution with true statistics $\mathbf{x}^\star$, obtained from sampling from this distribution. We regularize $\widetilde{\mathbf{x}}$ with a linear filter $\mathbf{F}$ (a matrix here) in order to minimize the sampling error: $\widehat{\mathbf{x}} = \mathbf{F}\widetilde{\mathbf{x}}$. A typical criterion to minimize is

$$\mathcal{L}(\mathbf{F}) = \mathbb{E}\left[(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})^{\mathrm{T}}(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})\right]. \quad (3)$$

The variation of this criterion with respect to a variation $\delta\mathbf{F}$ of $\mathbf{F}$ is $\delta\mathcal{L}(\mathbf{F}) = -2\mathrm{Tr}\left\{\delta\mathbf{F}^{\mathrm{T}}\mathbb{E}\left[(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})\widetilde{\mathbf{x}}^{\mathrm{T}}\right]\right\}$, which implies, that, at the minimum, we have an optimality condition in the form of an orthogonality of random vectors:

$$\mathbb{E}\left[(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})\widetilde{\mathbf{x}}^{\mathrm{T}}\right] = \mathbb{E}\left[(\mathbf{x}^\star - \widehat{\mathbf{x}})\widetilde{\mathbf{x}}^{\mathrm{T}}\right] = \mathbf{0}. \quad (4)$$

This is a linear equation in $\mathbf{F}$ whose solution is

$$\mathbf{F}^\star = \left\{\mathbb{E}\left[\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^{\mathrm{T}}\right]\right\}^{-1}\mathbb{E}\left[\mathbf{x}^\star\widetilde{\mathbf{x}}^{\mathrm{T}}\right]. \quad (5)$$

If $\mathbf{F}$ is a Schur filter, i.e., $\widehat{\mathbf{x}} = \mathbf{f} \circ \widetilde{\mathbf{x}}$, given by the Schur or Hadamart product (which is a subcase of the above problem) – hence $\mathbf{F}$ is now a vector $\mathbf{f}$ – then the solution has the form

$$\mathbf{f}^\star = \frac{\mathbb{E}\left[\mathbf{x}^\star \circ \widetilde{\mathbf{x}}\right]}{\mathbb{E}\left[\widetilde{\mathbf{x}} \circ \widetilde{\mathbf{x}}\right]}, \quad (6)$$

where the division of vectors is component-wise. Eqs. (5,6) can be applied to the filtering of $\widetilde{\mathbf{B}}$, storing the entries B_{ij} in $\mathbf{x}$. Hence, they provide optimality conditions for linear filtering of $\widetilde{\mathbf{B}}$. They are known in the signal scientific community as Wiener filters. We note that $\mathbf{F}^\star$, or $\mathbf{f}^\star$, still depends on the unknown true statistics $\mathbf{x}^\star$

2. The second one is to exploit the structure relationships that bind the moments of sample estimators of the reference distribution. For any reference distribution (referred to as the non-Gaussian case in the following), the second-order moments of the sample covariances B_{ij} are functions of the second-order and fourth-order moments of the reference distribution. If, in addition, one assumes this reference distribution to be a Gaussian, then the covariances of the sample covariances $\widetilde{B}_{ij}$ are only functions of the second-order moments of the reference distribution. This will be naturally referred to

later as the Gaussian case. For instance, in the Gaussian case, the relation has the well-known form:

$$\mathbb{E}\left[\left(B_{ij}^\star - \widetilde{B}_{ij}\right)^2\right] = \frac{1}{N-1}\left(B_{ij}^\star B_{ij}^\star + B_{ii}^\star B_{jj}^\star\right). \quad (7)$$

3. In spite of the above key ideas, some local spatial averaging will additionally be needed to obtain robust estimators for the filters and their correlation lengths. Such averaging can be justified by ergodic assumptions on the statistics of the errors.

In the following, we make the difference between the cases where the true distribution is assumed Gaussian or not, since we saw it has an impact on the structure function such as Eq. (7), and could yield distinct optimal filtering results.

2.2.1 Gaussian case

It turns out that it is more convenient to filter the variance and the correlation independently, in particular using a general linear filter for the variances and a Schur filter for the correlation (Ménétrier et al., 2015a).

We denote $\mathbf{v}$ the vector of variances, i.e., $v_i \equiv B_{ii}$. Combining the optimality criterion (4) with the structure relationship (7), without reference to any explicit filter at this stage, the filtered and the sampled variances are related by (see Eq. (50) of Ménétrier et al., 2015a):

$$\mathfrak{C}_i^{\mathrm{G}} \equiv \mathbb{E}[\widetilde{v}_i^2] - \frac{N+1}{N-1}\mathbb{E}[\widetilde{v}_i\widehat{v}_i] = 0, \quad (8)$$

with $\mathfrak{C}_i^{\mathrm{G}} = 0$ the optimality criterion in Gaussian conditions. If we filter the covariances with a Schur filter, i.e. $\widehat{\mathbf{B}} = \mathbf{F} \circ \widetilde{\mathbf{B}}$, then one obtains (see Eq. (64) of Ménétrier et al., 2015a):

$$F_{ij}^{\mathrm{G}} = \frac{N-1}{(N+1)(N-2)}\left\{(N-1) - \frac{\mathbb{E}[\widetilde{v}_i\widetilde{v}_j]}{\mathbb{E}[\widetilde{B}_{ij}^2]}\right\}. \quad (9)$$

2.2.2 Non-Gaussian case

In the non-Gaussian case, the structure relationship incorporates a term that depends on the fourth-order moments Ξ_{ijkl} of the true error statistics. Using these relationships and the optimality criterion (4), without reference at this stage to any particular filter, one obtains (see Eq. (48) of Ménétrier et al., 2015a):

$$\mathfrak{C}_i^{\mathrm{NG}} \equiv \mathbb{E}[\widetilde{v}_i^2] - \frac{N(N-2)(N-3)}{(N-1)(N^2-3N+3)}\mathbb{E}[\widetilde{v}_i\widehat{v}_i]$$
$$- \frac{N^2}{(N-1)(N^2-3N+3)}\mathbb{E}[\widetilde{\Xi}_{iiii}] = 0. \quad (10)$$

Again, but in the non-Gaussian case, if we regularize the covariances with a Schur filter, i.e., $\widehat{\mathbf{B}} = \mathbf{F} \circ \widetilde{\mathbf{B}}$, then ones

obtains the optimal filter (see Eq. (62) of Ménétrier et al., 2015a):

$$F_{ij}^{NG} = \frac{(N-1)^2}{N(N-3)} - \frac{N}{(N-2)(N-3)}\frac{\mathbb{E}[\widetilde{\Xi}_{ijij}]}{\mathbb{E}[\widetilde{B}_{ij}^2]}$$
$$+ \frac{N-1}{N(N-2)(N-3)}\frac{\mathbb{E}[\widetilde{v}_i\widetilde{v}_j]}{\mathbb{E}[\widetilde{B}_{ij}^2]}. \tag{11}$$

For the localization of the covariances and hence the correlations, Eqs. (9,11) provide the optimal Schur localization. For the filtering of the variances, one uses Eqs. (8,10) but still need to specify a filter, such as a convolution with a short-range kernel of correlation length l. Then Eqs. (8,10) are implicit equations for l, which can be solved iteratively using, for instance, a fixed point method.

We note that all these formulae still depend on some statistical expectation, such as $\mathbb{E}[\widetilde{B}_{ij}^2]$. To make those formulae practical, we identify these expectations as local, if not global, spatial averages.

2.2.3 Wiener filter

There is an alternative to using the optimality condition (4) in conjunction with the structure relationships of the moments of $\mathbf{B}$. We propose to solely use the optimality condition (4) and upon choosing the generic form of the filter use the optimal filters given by Eqs. (5) or (6). We will call them Wiener filters in the following.

For instance, assuming Schur regularization, we obtain the Wiener filter:

$$L_{ij} = \frac{\mathbb{E}[\widetilde{B}_{ij}B_{ij}^{\star}]}{\mathbb{E}[\widetilde{B}_{ij}^2]}. \tag{12}$$

Using the sample estimator $B_{ij}^{\star} = \mathbb{E}[\widetilde{B}_{ij}]$, we obtain:

$$L_{ij} = \frac{\mathbb{E}^2[\widetilde{B}_{ij}]}{\mathbb{E}[\widetilde{B}_{ij}^2]}. \tag{13}$$

Both Wiener and Schur filters will be applied to sub-domains defined around instrumented tower locations measuring continuously CO_2 mole fractions in the US Upper Midwest (Miles et al., 2012). The sub-domains cover an area of 400x400 km^2 around each site (here seven sites across the domain) which also correspond to the spatial extent of the local spatial averaging (3rd item in Section 2.2). Due to computational limitations, we performed additional experiments with larger sub-domains for our 25-member ensemble, as show in the Section 3.5.

2.3 Meteorology and CO_2 error structures

We want to explore the relationships between the different variables especially in situ mole fractions of CO_2, total column XCO_2, and PBLH. We will compare both the error variances and covariances to identify possible links between error structures in PBLH and CO_2/XCO_2 mole fractions. We will explore the spatial correlation lengths for CO_2 mole fractions, mean horizontal wind (zonal and latitudinal components), and PBLH to quantify and possible utilize error structures in meteorological fields to generate CO_2 and XCO_2 error structures. Most Ensemble Prediction Systems (EPS's) provide spatial error correlations for meteorological variables which could be used to construct error covariances for CO_2 and XCO_2. Error covariances of CO_2 mole fractions depend on the CO_2 fluxes, but error structures in the atmospheric models should remain independent of the CO_2 flux distribution. Díaz-Isaac et al. (2018a) show that first-order discrepancies in PBLH seem related to large errors in CO_2. Here, we investigate further the links between errors across different variables. We present the results in Section 3.4 for the variances and in section 4.3 for the error correlations.

3 Results

3.1 Sampling noise due to ensemble size

We computed the sample variances over the domain from the 5-, 8-, 10- and 25-member ensembles as shown in Fig. 1. The increase in variances and the presence of additional fine-scale structures are visible in small-size ensembles (5 to 10 members) compared to the 25-member ensemble. Fine-scale structures reflect the sampling noise in the small-size ensembles, reaching a maximum in the 5-member ensembles (cf. Fig. 1, panel d). These spurious structures appear with small-size ensembles and are to be filtered later. The range of values for error variances increases for small-size ensembles, independently of the calibration process. In addition, the variances in calibrated ensembles with more members are smaller because the inflation of the variance is a direct consequence of removing members. Hence, the calibration process better inflates the deviation of members from the mean for small ensembles. Díaz-Isaac et al. (2018a) have shown that the calibration process yields smaller-size ensembles to better represent model errors. Here, the variance from the 25-member ensemble remains harder to inflate by the calibration process, also less affected by sampling noise which is likely to be less representative of the actual transport model errors.

3.2 Filtering of sampling noise

We show here the values of the length scales in our filter resulting from the optimality criteria, applying both Gaussian (cf. Eq. 8) and non-Gaussian (cf. Eq. 10) filters to the raw variances. We implemented the dichotomy algorithm proposed in Ménétrier et al. (2015b) to obtain the optimal length scale of the filter, dividing (or multiplying) the length scale by a factor of 2 until convergence. The algorithm solves for the optimal length scale of the filter by scanning the space of
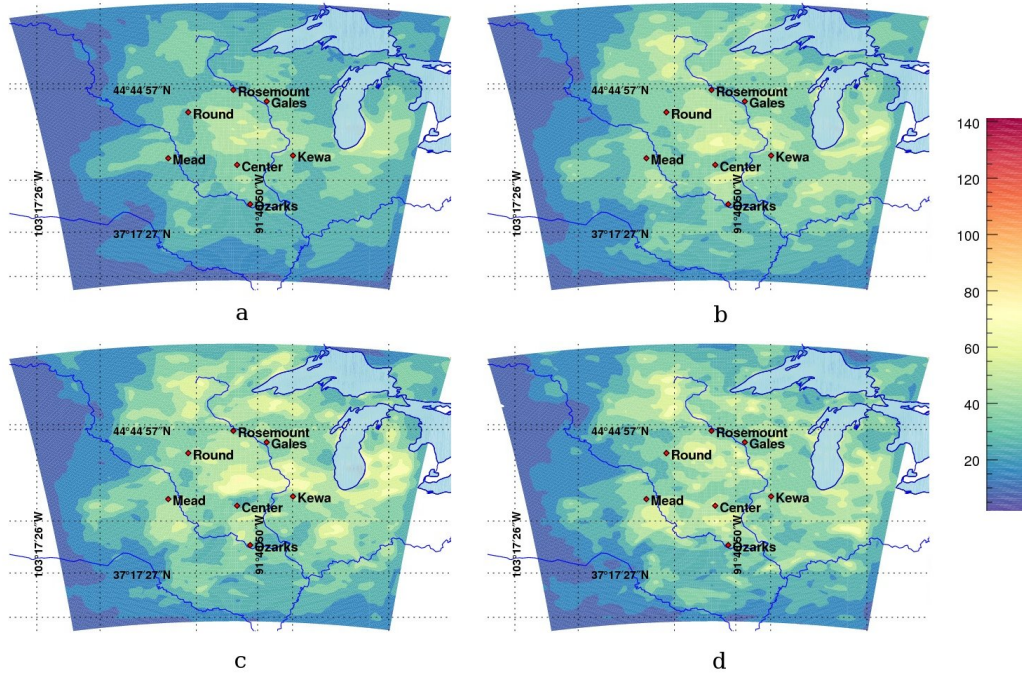
Figure 1. Variances of CO_2 in situ mole fractions (about 100m agl) (in ppm^2) using 25 members (upper left panel), 10 members (upper right panel), 8 members (lower left panel), and 5 members (lower right panel).

solutions iteratively (applying a multiplicative factor at each time step to minimize the cost function). For all our cases, we defined the upper bound of the diagnosed length scale at 750 km, to represent about half the size of our simulation domain (square of 1,600 km wide). This large value means that the extent of noisy structures would encompass the entire domain, and therefore would not be recoverable. Here, the sampling noise is characterized by length scales of sizes ranging from a few kilometers to several hundred kilometers. In practice, the algorithm always converges but length scale might be larger than the domain size, meaning that all spatial structures in the variances are considered as noise. This situation happens for two main reasons: spatial structures in the noise are similar to spatial gradients in the true variances, or the noise is larger than the true variances. Hence, the length scale falls beyond the limit of our simulation domain. In the method described by Ménétrier et al. (2015b), the ergodic assumption is necessary to diagnose robust estimators of the filter (cf. Section 2.2).

For CO_2 mole fractions (cf. Fig. 2), the algorithm for the calibrated 25-member ensemble systematically converges to small length scales (<50 km), indicating that noise structures are very small in our optimal ensemble. When using the Gaussian filter, the algorithm systematically converges below our 750-km threshold for all cases except for 30% of the days with the smallest ensemble (5 members). In the non-Gaussian case, the filter converges to larger length scales beyond our threshold with the 5- and 10-member ensembles for less than 30% of the days. Typically, length scales beyond our 750-km threshold are temporally coherent over periods of several days suggesting weather-related structures possibly inherited from synoptic-scale systems. These periods might be caused by high sampling noise compared to the true variances or by similar scales in spatial structures for both noise and true variances. Overall, the non-Gaussian filter shows a lower rate of convergence below 750 km compared to the Gaussian filter for CO_2 mole fractions.

For XCO_2 column mole fractions (cf. Fig. 3), the optimal length scales are larger and the non-Gaussian filter converges beyond our threshold more frequently (about 50% of the days for 8 members or less). Even with the optimal 25-member ensembles, error structures of about 50 to 200km are filtered out, significantly larger than for the CO_2 mole fractions. We discuss in Section 4.2 the possible physical reasons behind these larger length scales, possibly due to large-scale structures in the Free Troposphere or to the complexity in noise structures as XCO2 data integrate noises from different altitudes. Fig. 4 shows the results of the PBLH's for which both filters converges beyond our threshold for half of the days in the Gaussian case. However, the filter converges more often below our threshold with the non-Gaussian filter applied to 10-member ensembles. Variance noise for PBLH present skewed distributions (not shown here) requiring the use of a non-Gaussian filter. We conclude here that 8- and 10-member ensembles are the minimum sizes with which sub-threshold convergence can be obtained on most days for most vari-
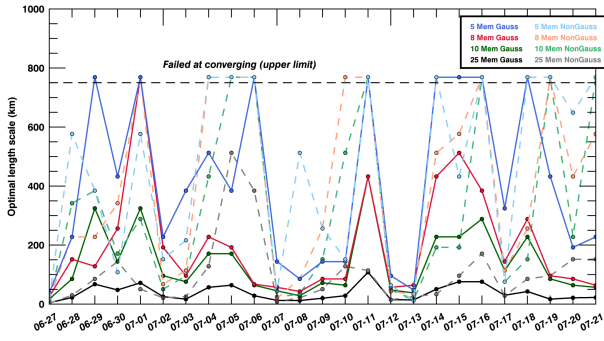
Figure 2. Length scale (in km) of the variance filter for in situ CO_2 mole fractions at about 100m agl using Gaussian (solid lines) and non-Gaussian (dash lines) equations from 27 June to 21 July, 2008.
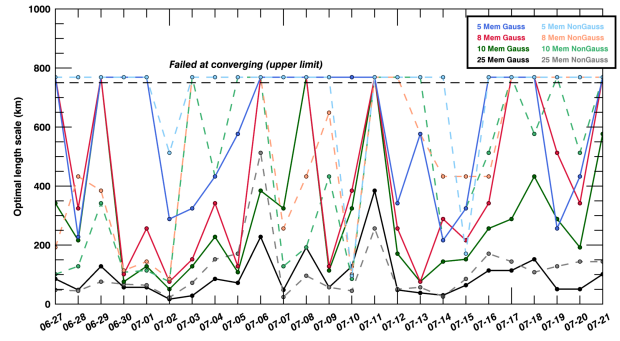
Figure 3. Length scale (in km) of the variance filter for XCO_2 total column dry air mole fractions using Gaussian (solid lines) and non-Gaussian (dash lines) equations from 27 June to 21 July, 2008.

ables. 5-member ensembles will still be studied later on for covariances, as the localization of covariances does not depend on the filtered variances, but the low rate of convergence below 750 km might limit the use of the filtered variances.

3.3 Variance filtering and ensemble sizes

The filtered variances shown in Fig. 5, here with the Gaussian filter, for the different ensemble sizes are in better agreement with the variances of the 25-member ensemble both in term of spatial patterns and magnitudes among the different ensembles. The filter successfully removed noisy structures, therefore decreasing the dependence on the number of members used in each case. Despite length scales beyond our threshold for 30% of the days with the 5-member ensemble, filtered variances at the monthly time scale show similar structures than 8- or 10-member ensembles, with nearly all of the noisy structures being removed by the filter. Compared to earlier results, the ensemble size does not seem to fundamentally limit the capacity of the filter to remove the noise, despite days with convergence beyond 750 km. The variance magnitude remains slightly larger for 10 members or less, with a relative over-estimation of about 15%. Our 5-member ensemble provided the best match with only 10% higher than the 25-member filtered variances. The averaging over a whole month compensates for days with convergence beyond 750-km, producing reasonable estimates of the optimal variance even with 5 members. This results suggests that climatological error variances from small-size ensembles can be a good first approximation of the true variance when filtered correctly over most days.

3.4 Error variances in CO_2, XCO_2, and PBLH

We show in Fig. 6 the spatial distribution of error variances from the 25-member calibrated ensemble for in situ CO_2 mole fractions in the PBL (100m agl), in situ CO_2 mole fractions in the Free Troposphere (about 5km agl), total col-
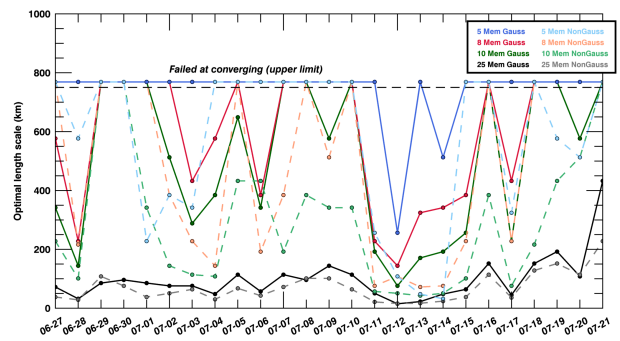
Figure 4. Length scale (in km) of the variance filter for Planetary Boundary Layer heights using Gaussian (solid lines) and non-Gaussian (dash lines) equations from 27 June to 21 July, 2008.

umn of XCO_2 dry air mole fractions, and PBLH (in meter agl). The four variables display very distinct spatial patterns. XCO_2 variance spatial patterns (cf. Fig. 6, panel c) exhibit distinct maximum values located in the eastern and southeastern part of the domain, whereas high CO_2 variances are observed in the northeastern part of the domain for free tropospheric CO_2 (cf. Fig. 6, panel b) or centrally located for CO_2 variances in the PBL (cf. Fig. 6, panel a). Finally, PBLH variances (cf. Fig. 6, panel d) show no indication of direct relationship between large errors in the western part of the domain and the other three CO_2 variables. We conclude here that no direct relationship can be utilized to construct CO_2 variances based on PBLH. Similarly, maximum variances among the three CO_2 variables are also significantly different in distribution and magnitudes.

3.5 Covariance Localization: Schur and Wiener filters

Error covariances in CO_2 mole fractions scale with the magnitude of the surface CO_2 fluxes. They are therefore difficult to interpret. Instead, we present here the error correlations
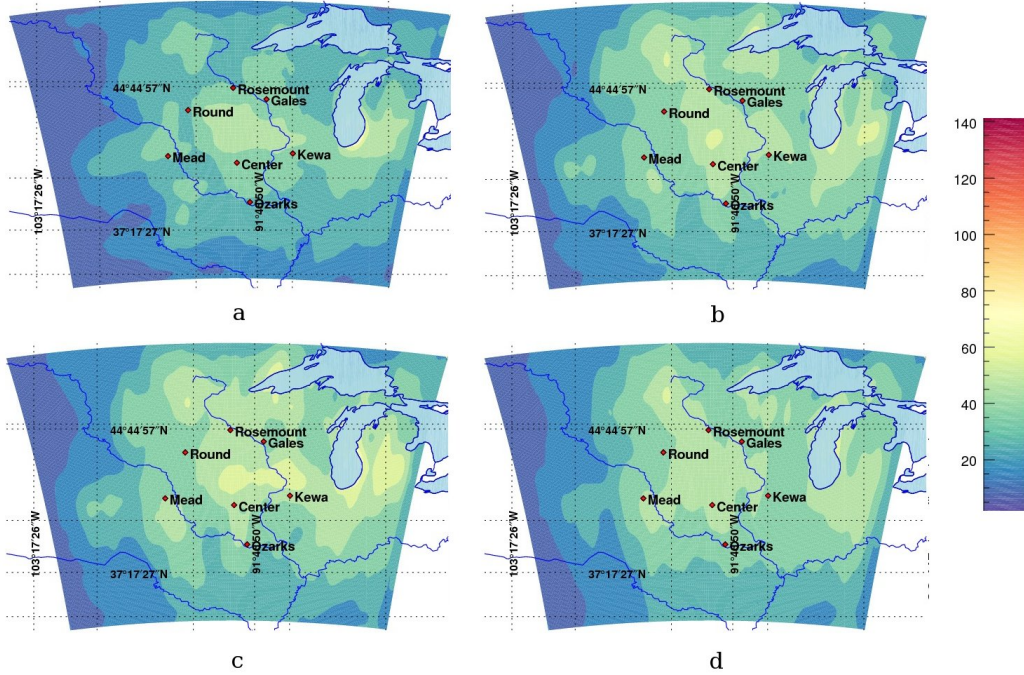
Figure 5. Monthly averages of the filtered variances of in situ CO_2 mole fractions (about 100m agl) (in ppm^2) using 25 members (upper left panel), 10 members (upper right panel), 8 members (lower left panel), and 5 members (lower right panel).
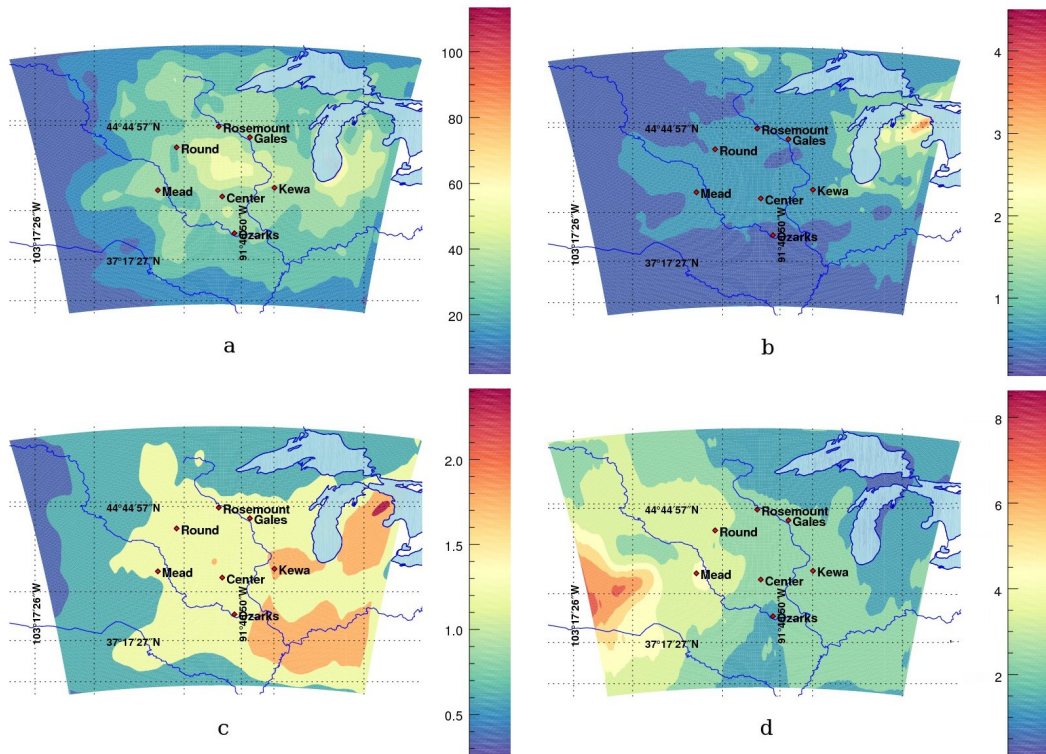
Figure 6. Filtered variances using the calibrated 25-member ensemble of in situ CO_2 mole fractions at 100m agl (in ppm^2) (a), in situ CO_2 mole fractions at 5km agl (in ppm^2) (b) XCO_2 total dry air mole fractions (in ppm^2) (c), and PBLH (in m^2) (d).

to highlight the spatial structures inherited from the transport models, independent of the magnitude of the underlying CO_2 surface fluxes. We show in Fig. 7 the hourly correlation structures from the full 45-member ensemble (top row) and our 25-member calibrated ensemble (lower row) at one of the instrumented towers in the US Midwest, *i.e.* Centerville, Iowa. We applied both Schur (cf. Fig. 7, middle column) and Wiener (cf. Fig. 7, right column) filters to compare the impact of both filters on the raw correlations (cf. Fig. 7, left column). The Schur filter has less impact on the correlations compared to the Wiener filter which attenuates significantly the magnitude of the correlations for both ensembles. In Section 2.2, the local averaging of the optimal length scale assumes that the sampling noise is spatially homogeneous (third ingredient of the methods). This homogeneity assumption is required for the ergodicity assumption to apply, therefore yielding a domain-averaged filtering approach. The sub-domain used for the covariance filtering (here 400x400 km^2) limits the spatial extent to 200 km around the observation location. The size of the domain was defined primarily for computational efficiency and based on the size of correlation structures, usually of about 100-200 km in length scale. To evaluate this assumption, we compared the size of the sub-domain to filter the covariances with a large area of 900x900 km^2 for the 45-member ensemble to a smaller area of 400x400 km^2 for the 25-member ensemble. Filtered correlations show similar results for Schur and slightly larger values for the smaller sub-domain when applying the Wiener filter. We conclude here that the spatial local averaging has a minor impact on the results, and that our 25-member ensemble has similar spatial structures to the original 45-member ensemble, with larger correlations at short distances. We extend this analysis to the monthly time scale by showing monthly averaged error correlations, super-imposed from different tower locations on the same map to aggregate the results at multiple locations (cf. Fig. 8). When averaged over longer time scales (cf. Fig. 8), the filtered correlations become isotropic, distributed around each location. The magnitudes remain larger with the Schur filter (cf. Fig. 8, middle panel) compared to the Wiener filter (cf. Fig. 8, right panel) but the differences are noticeably smaller. The unfiltered correlations (cf. Fig. 8, left panel) are noticeably larger due to noisy structures). After filtering, the spatial structures are distributed around the observation locations following a pseudo-Gaussian pattern. The magnitude of the error correlations, *i.e.* the length scale of the errors, is reduced in both cases compared to the raw correlations (cf. Fig. 8, left panel). This result confirms the sub-domain used here (400x400 km^2) is sufficient to represent the error correlation structures around each measurement location and describes fully the error structures.

We show in Fig. 9 the results for the different ensemble sizes using the Schur filter. The 10- and 8-member ensembles show similar magnitude and patterns for the different sites, but the correlations are smaller than with the original ensemble. In comparison, the Wiener filter (cf. Fig. 10) generates consistent patterns with 25-, 10- and 8-member ensembles. In both cases, the filters decreases significantly the correlations in the 5-member ensemble, revealing the inability of the filter to separate the noise from the actual error correlations. We present the localized correlation length scales for each tower and for each day in Fig. 12 (upper left panel). For both Center and Mead, length scales are noticeably larger than for the other towers and decrease rapidly until July 2nd, before converging back to the same values diagnosed for other measurement sites. The differences across towers suggest local differences in error correlations, even across the same region for a single day (up to 70 km across our sites). These differences correspond to the beginning of summer, when both weather and ecosystem fluxes vary rapidly especially in agricultural areas. We discuss in Section 4.3 the relationship between meteorological and CO_2 variables for both error variances and error correlations.

4 Discussion

4.1 Minimum size of calibrated ensembles

We discussed in Section 3.3 the dependence of the success of the variance filtering on the ensemble size. From these results, an ensemble of at least 8 to 10 members seems required to reach convergence below our 750-km convergence and hence filter the error variances. However, the spatial representation of the averaged filtered variances using a calibrated 5-member ensemble (cf. Fig. 5) indicates a reasonable recovery of the error variances at the monthly time scale but not at the daily time scale. Theoretically, the minimum number of members for the covariance filtering is four, based on the Eq. 11 with a factor $(N-3)$ in the denominator, which shows that 5-member ensembles are close to this limit and are not recommended in a more general context. The application here suggests 5-member ensembles are acceptable but 8- to 10-member ensembles would be a minimum both in practice and in theory over different seasons and regions. Convergence beyond our threshold over single days has a limited impact on the monthly mean filtered variances. We conclude here that the filter produces satisfactory results to generate first-order estimates of the CO_2 mole fraction errors. To achieve a systematic daily convergence below threshold, we recommend a larger number of members in the ensemble. One important point here is the calibration step performed before filtering, which optimizes the information content in each member relative to the other members. Therefore, a randomly-generated ensemble may require additional members in order to represent the actual error variances. We tested the filtering technique on random ensemble (*i.e.* uncalibrated) and found that beyond-threshold convergence is more frequent with 10 or less members (not shown here).
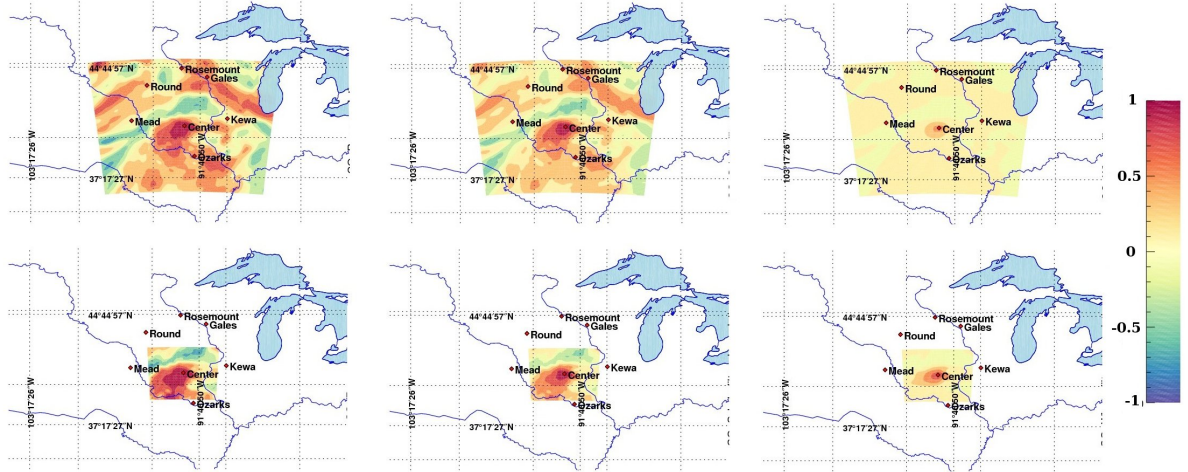
Figure 7. Hourly raw (left column) and filtered correlations using Schur (middle column) and Wiener (right column) filtered correlations of in situ CO_2 mole fractions at the Centerville tower (at about 100m agl) on June 26, 2008. The correlations are based on the original 45-member ensemble (top row) with a large subdomain (900x900 km^2) and the 25-member ensemble (bottom row) over the reference subdomain (400x400 km^2). The highlighted domain also corresponds to the domain used for the spatial averaging of the correlations in the optimality conditions.
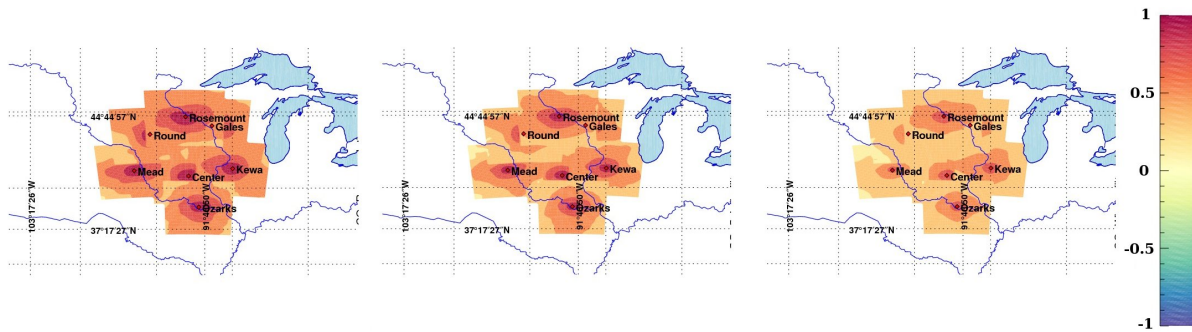
Figure 8. Raw (left panel) and filtered correlations of CO_2 in situ mole fractions at about 100m agl based on 25 members using the Schur localization (middle panel) and Wiener and Schur filter (right panel).

4.2   Impact of calibration on small-size ensembles

We have explored the impact of the calibration process on the error variances and covariances by filtering non-calibrated sub-ensembles of 8 and 10 members (cf. Fig. 11). These random ensembles have no member in common with their calibrated counter-part, and are composed of simulations using various physics configuration randomly selected among the 45 original model configurations. The optimal length scale of the variance filter is systematically lower for non-calibrated ensembles (cf. Fig. 11, in grey and light blue) compared to calibrated ensembles (cf. Fig. 11, in black and royal blue), suggesting lower levels of noise with larger spatial structures, similar to larger ensemble sizes (25 members or more). Because members of the calibrated ensembles were selected to maximize the information content, calibrated-ensemble members differ more from each other than non-calibrated en-

semble members. As shown in Díaz-Isaac et al. (2018a), one member with higher or lower PBLH statistics (usually the monthly mean model estimate) is systematically selected in order to generate calibrated ensembles with enough variance, and therefore capture the spatial and temporal variability in observed PBLH's from 14 radiosondes. Calibration might suggest different combinations of model physics for every single time period. In future studies, we recommend a combination of several pre-selected model physics with added perturbations to produce a sufficient ensemble spread but not perform the calibration for every time period. The calibration procedure might still be applied to existing ensembles but remain insufficient to sample the full spatial and temporal variability as discussed in Díaz-Isaac et al. (2018a). Some of these members introduce different spatial structures compared to the original ensemble, increasing the spread significantly. However, the small size of our ensembles with a
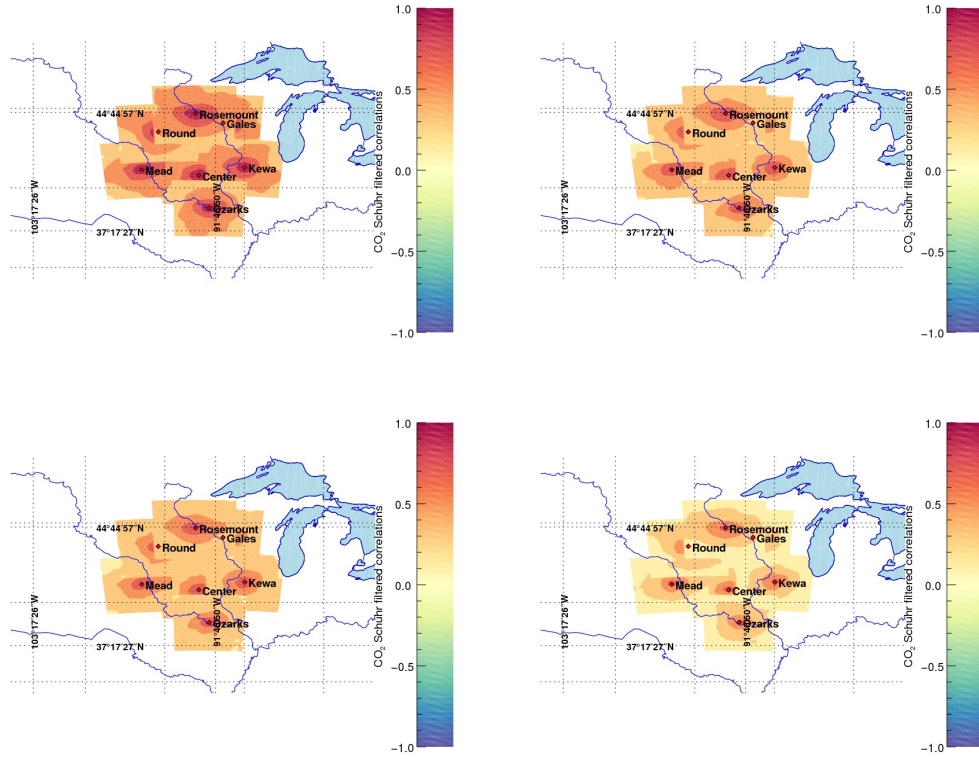
Figure 9. Filtered correlations (using Schur product only) of CO_2 in situ mole fractions (about 100m agl) (in ppm^2) (about 100m agl) (in ppm^2) using 25 members (upper left panel), 10 members (upper right panel), 8 members (lower left panel), and 5 members (lower right panel).

larger variance may affect the sampling of these structures and hence our ability to differentiate noise from actual error structures. The approach proposed by Ménétrier et al. (2015a) was initially designed for ensembles with independently distributed members, which is not the case here. Some true correlation structures in the calibrated ensemble may be considered as noise by the filter. Another approach would generate an ensemble based on the localized ensemble-based error covariance matrix. This approach will be developed further in future studies but is beyond the scope of this paper.

4.3   Spatial structures in CO_2 and meteorological errors

In Fig. 6, significant differences in filtered error variances were observed between in situ CO_2 mole fractions (at different altitudes), XCO_2 mole fractions, and PBLH. Therefore, we conclude here that transport errors of meteorological variables are not transferable to CO_2 and XCO_2 error variances. This finding is in agreement with Miller et al. (2015) who found no direct relationship between errors in the meteorology and in situ CO_2 mole fractions. However, considering the covariances, the spatial structures in CO_2 mole fraction errors inherited from transport model errors exhibit

well-defined patterns (e.g. Fig. 10). By fitting a simple Gaussian function in the form of $e^{-\frac{x}{L}}$ to the filtered covariance fields, we diagnosed the characteristic length scale of the spatial error structures for the different variables, here CO_2 100m-high mole fractions, the mean horizontal zonal wind component, and PBLH. Fig. 12 shows the daily length scales at the seven measurement locations over the simulation period (*i.e.* 27 June to 21 July). The length scales L for the three variables increase rapidly between 27 June and 4 July from less than 100 km to 150 km or higher. As there is no long-term spin-up (re-initialization of the perturbations every 5 days), the asymptotic behavior is most likely due to the seasonality of the errors from early to late summer. The seasonal changes in the atmospheric dynamics impact the spatial structures in the errors for the three variables. Across the seven sites, the characteristic length scale of spatial error structures also vary significantly, in particular for PBLH with large differences across sites. Both CO_2 mole fractions and the mean zonal wind component reach a maximum value over July, respectively 140 km and 120 km. The comparison of the mean length scales (lower right panel in Fig. 12) highlight the differences between the three variables. Both the CO_2 100m-high mole fractions and the mean horizontal
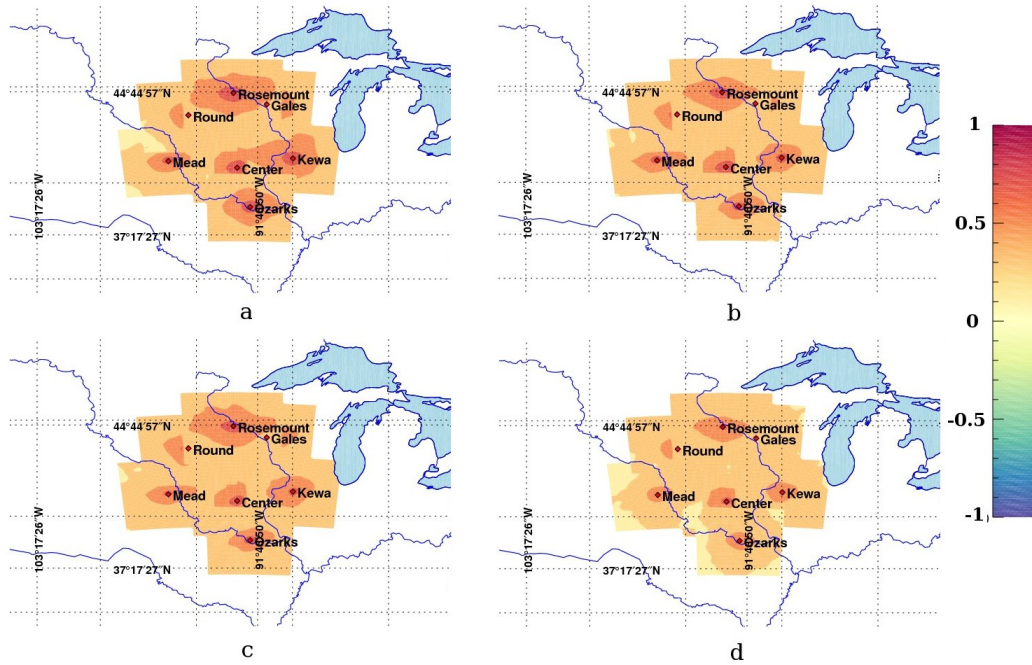
Figure 10. Filtered correlations (using Wiener & Schur) of CO_2 in situ mole fractions (about 100m agl) (in ppm^2) (about 100m agl) (in ppm^2) using 25 members (upper left panel), 10 members (upper right panel), 8 members (lower left panel), and 5 members (lower right panel).
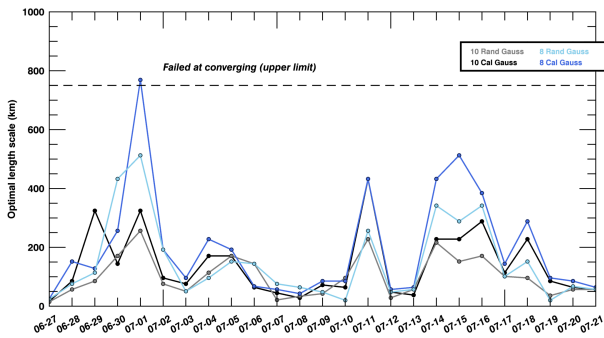
Figure 11. Length scale (in km) of the Gaussian variance filter for 10-member random (in grey) and calibrated (in black) ensembles, and 8-member random (in light blue) and calibrated (royal blue) ensembles, for in situ CO_2 mole fractions at about 100m agl using Gaussian equations from 27 June to 21 July, 2008.

zonal wind component show similar variations and converge towards the same values, but differ significantly from 29 June to July 10. We conclude here that first-order estimates for CO_2 spatial error correlations may be derived from meteoro-
5 logical error structures, in particular from the mean horizontal wind speed. But these approximations may be valid only for specific time periods. As presented in Section 3.4, error

variances in CO_2 and XCO_2 mole fractions are decoupled from PBLH errors. Here, we suggest that error correlations may be derived from wind errors but error variances should 10 still be computed independently.

4.4 Evaluation and modeling of error correlations

This study presents a methodology to filter the noise in error structures from a small-size ensemble. The evaluation of the filtered structures would benefit from dense mea- 15 surement campaigns sampling spatial structures across large domains, such as the Atmospheric Carbon and Transport (ACT)-America campaigns[2]. Previous studies have shown the utility of aircraft measurements to diagnose error correlations (Gerbig et al., 2003) but the separation of spatial struc- 20 tures induced by surface flux errors and atmospheric transport errors remains challenging in order to construct observation error covariance matrices. The combination of ensemble systems such as Ensemble Kalman Filter (EnKF) systems and intensive aircraft campaigns will provide additional in- 25 sights to evaluate filtering approaches (e.g. Chen et al., 2019). To introduce the findings of our study into an atmospheric inversion system, an additional step would be required in order to construct a regularized error covariance matrix. In this study, we acknowledge here that we have applied Schur and 30 Wiener filters using the raw filtering matrices (Eqs. 9, 11, and

[2]https://www-air.larc.nasa.gov/missions/ACT-America/
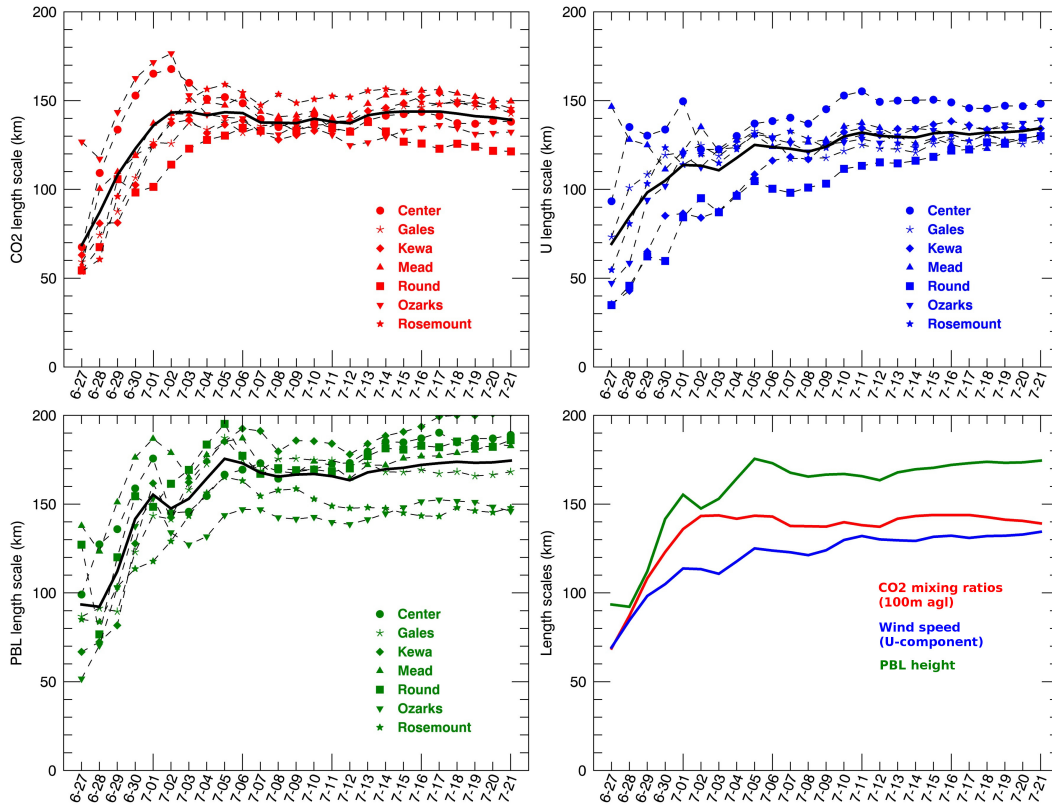
Figure 12. Characteristic length scales L of filtered error correlations (from the 25-member ensemble) using an exponential function ($e^{-\frac{x}{L}}$) for CO_2 100m-high mole fractions (upper left panel), the mean horizontal zonal wind component (U; upper right panel), PBLH (lower left panel), and the mean values for the three variables (lower right panel) for every day of the simulation period (27 June to 21 July).

6), that may not be semi-positive definite, requirement for getting semi-positive definite regularized covariances. Future studies should include an additional step by adding a regularization of the covariances before filtering. For example, Lauvaux et al. (2009) proposed to model the error structures using a diffusion equation able to represent anisotropic structures, following the methodology described in Pannekoucke et al. (2008). Here, we presented a local filter able to remove the noise in the error structures. The diagnosed error structures can be approximated by different spatial functions of varying degree of complexity, further regularized to generate positive-definite error covariance matrices. This next step is beyond the scope of this paper but will be conducted in the future to generate an efficient model of the corresponding error covariances based on our current results.

5 Conclusions

We have diagnosed the error variances and the spatial error structures from our mesoscale transport models at daily and monthly time scales. Applied to both CO_2 mole fractions and meteorological variables, we implemented a cost-effective filtering technique currently used in meteorological data assimilation systems (Ménétrier et al., 2015b) to describe spatial error structures using a small-size ensemble. The approach remains affordable for multi-year inversions of sources and sinks at continental or regional scales. The removal of noisy structures in our small-size ensembles is evaluated by comparison to larger-size ensembles, both the original 45-member ensemble and our optimal calibrated sub-ensemble of 25 members. A second filtering approach for error covariances was successfully applied using the Wiener filter, producing similar results compared to the Schur filter over the 1-month simulation period. Differences were noticeable at shorter time scales (*i.e.* daily). The spatial distribution of error variances and spatial error structures are recoverable from small-size ensembles of 8 to 10 members, daily for in situ CO_2 mole fractions and monthly for total column XCO_2, providing a more realistic representation of transport errors in future mesoscale inversions of CO_2 fluxes. We noted that error variances of in situ CO_2 mole fractions and total column XCO_2 differ significantly, even when varying the altitudes or considering PBLH error structures. We conclude that error variances for remote sensing observations need to be quantified independently of PBL or free tropospheric mole fractions.

We have discussed the potential use of meteorological error structures such as the mean horizontal wind to approximate spatial error correlations of in situ CO_2 mole fractions. The seasonal variations in wind, PBLH, and in situ CO_2 mole fractions are highly correlated, while the typical length scales in error structures vary from 100km to 150km in the middle of summer depending on the variable. We conclude here that meteorological error structures may provide a first-order estimation of correlation length scales in CO_2 inversions when no ensemble of CO_2 simulations is available.

6 Code availability

The code is accessible under request by contacting the corresponding author (thomas.lauvaux@lsce.ipsl.fr).

7 Data availability

The model simulation outputs are available under request by contacting the corresponding author (thomas.lauvaux@lsce.ipsl.fr).

Appendix A: Calibration of the reference 25-member ensemble

In this study, we generate a reference ensemble to evaluate the sampling noise in the ensembles of smaller sizes. The original 45-member ensemble, uncalibrated, cannot be used as a reference as it under-estimates model errors. Yet, the reference ensemble needs to include enough members to limit the sampling noise. Based on the same original ensemble of 45 members as in Díaz-Isaac et al. (2018a), we reduce it to a calibrated 25-member ensemble using the Simulated Annealing (SA) algorithm. The selection of the optimal calibrated ensemble is based on three meteorological variables (i.e., wind speed, wind direction and PBLH) and follows the same procedure described in Díaz-Isaac et al. (2018a). The SA for 25 members uses 40,000 iterations to reach convergence, significantly larger than the 20,000 iterations for 10-, 8- and 5-member ensembles. The same criteria used by Díaz-Isaac et al. (2018a) was applied to the selection process of the calibrated 25-member ensemble, improving the flatness of the rank histograms (Fig. A1). This selection is based on two criteria. First, we selected all the 25-member sub-ensembles with a rank histogram score smaller than six for each individual meteorological variable. In a second step, we filtered out the 25-member ensembles accepted by the SA algorithm but corresponding to a bias (i.e. mean model-data mismatch over 25 days) larger than the bias in the original 45-member ensemble. These criteria are applied to the three meteorological variables. This procedure is described in more details in Díaz-Isaac et al. (2018a). The rank histograms in Fig. A1 show a limited under-dispersion of the 25-member ensemble, significantly reduced after calibration from 6.1, 6.2, 3.2 to 5.1, 4.9, and 2.5 for wind speed, wind direction, and PBLH respectively.

Author contributions. The WRF-Chem simulations were performed by L.I. Díaz-Isaac and T. Lauvaux; the filtering technique was coded by M. Bocquet and T. Lauvaux based on the work of B. Ménétrier; the concept and ideas were designed by M. Bocquet, T. Lauvaux, and N. Bousserez; the manuscript was prepared by T. Lauvaux, L.I. Díaz-Isaac, N. Bousserez and M. Bocquet.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. This research was supported by National Aeronautics and Space Administration (NASA) Terrestrial Ecosystem and Carbon Cycle Program (grant #NNX11AE79G), by the NASA's Earth Venture Program Atmospheric Carbon and Transport (ACT)- America (grant #NNX15AG76G), the Alfred P. Sloan Graduate Fellowship, by the NASA Carbon Monitoring System program (grant #NNX13AP34G), and by the National Oceanic and Atmospheric Administration (grant #NA14OAR4310136). CEREA is a member of *Institut Pierre-Simon Laplace* (IPSL).

References

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Mon. Wea. Rev., 129, 2884–2903, 2001.

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmospheric Measurement Techniques, 7, 647–687, doi:10.5194/amt-7-647-2014, https://www.atmos-meas-tech.net/7/647/2014/, 2014.

Baker, D. F., Doney, S. C., and Schimel, D. S.: Variational data assimilation for atmospheric CO_2, Tellus B, 58, 359–365, doi:10.1111/j.1600-0889.2006.00218.x, http://dx.doi.org/10.1111/j.1600-0889.2006.00218.x, 2006.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., , and Zhu, Z.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988-2003, Global Biogeochem. Cycles, 20, 439, doi:0.1029/2004GB002, 2007.

Brousseau, P., Berre, L., Bouttier, F., and Desroziers, G.: Flow-dependent background-error covariances for a convective-scale data assimilation system, Quarterly Journal of the Royal Meteorological Society, 138, 310–322, doi:10.1002/qj.920, http://dx.doi.org/10.1002/qj.920, 2012.
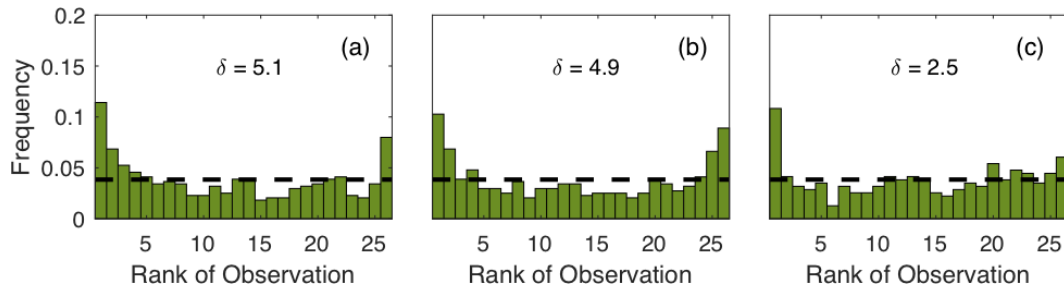
Figure A1. Rank histograms of the 25-member ensemble after calibration for wind speed (a), wind direction (b) and Planetary Boundary Layer Height (c) over 35 days (June 18 to July 21, 2008) using 14 rawinsonde sites available over the US Midwest. The horizontal dashed line corresponds to the ideal value for a flat rank histogram with respect to the number of members. The rank histogram score (δ) defines the flatness of the rank histogram, 1 being the ideal value.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmospheric Chemistry and Physics, 5, 2691–2702, doi:10.5194/acp-5-2691-2005, https://www.atmos-chem-phys.net/5/2691/2005/, 2005.

Chen, H. W., Zhang, F., Lauvaux, T., Davis, K. J., Feng, S., Butler, M. P., and Alley, R. B.: Characterization of Regional-Scale CO_2 Transport Uncertainties in an Ensemble with Flow-Dependent Transport Errors, Geophysical Research Letters, 46, 4049–4058, doi:10.1029/2018GL081341, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081341, 2019.

Chevallier, F.: Impact of correlated observation errors on inverted CO_2 surface fluxes from OCO measurements, Geophysical Research Letters, 34, doi:10.1029/2007GL030463, http://dx.doi.org/10.1029/2007GL030463, l24804, 2007.

Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, Geophysical Research Letters, 33, doi:10.1029/2006GL026496, http://dx.doi.org/10.1029/2006GL026496, l13802, 2006.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2010JD013887, http://dx.doi.org/10.1029/2010JD013887, 2010.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Quéré, C. L., Myneni, R., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2015.

Crisp, D., Atlas, R., Breon, F.-M., Brown, L., Burrows, J., Ciais, P., Connor, B., Doney, S., Fung, I., Jacob, D., Miller, C., O'Brien, D., Pawson, S., Randerson, J., Rayner, P., Salawitch, R., Sander, S., Sen, B., Stephens, G., Tans, P., Toon, G., Wennberg, P., Wofsy, S., Yung, Y., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, Advances in Space Research, 34, 700 – 709, doi:https://doi.org/10.1016/j.asr.2003.08.062, http://www.sciencedirect.com/science/article/pii/S0273117704003539, 2004.

De La Chevrotière, M. and Harlim, J.: A Data-Driven Method for Improving the Correlation Estimation in Serial Ensemble Kalman Filters, Mon. Wea. Rev., 145, 985–1001, doi:10.1175/MWR-D-16-0109.1, 2017.

Díaz-Isaac, L. I., Lauvaux, T., Davis, K. J., Miles, N. L., Richardson, S. J., Jacobson, A. R., and Andrews, A. E.: Model-data comparison of MCI field campaign atmospheric CO_2 mole fractions, Journal of Geophysical Research: Atmospheres, 119, 10 536–10 551, doi:10.1002/2014JD021593, http://dx.doi.org/10.1002/2014JD021593, 2014JD021593, 2014.

Díaz-Isaac, L. I., Lauvaux, T., Bocquet, M., and Davis, K. J.: Calibration of a multi-physics ensemble for greenhouse gas atmospheric transport model uncertainty estimation, Atmospheric Chemistry and Physics Discussions, 2018, 1–40, doi:10.5194/acp-2018-1117, https://www.atmos-chem-phys-discuss.net/acp-2018-1117/, 2018a.

Díaz-Isaac, L. I., Lauvaux, T., and Davis, K. J.: Impact of physical parameterizations and initial conditions on simulated atmospheric transport and CO_2 mole fractions in the US Midwest, Atmospheric Chemistry and Physics, 18, 14 813–14 835, doi:10.5194/acp-18-14813-2018, https://www.atmos-chem-phys.net/18/14813/2018/, 2018b.

Engelen, R. J., Denning, A. S., and Gurney, K. R.: On error estimation in atmospheric CO_2 inversions, Journal of Geophysical Research: Atmospheres, 107, ACL 10–1–ACL 10–13, doi:10.1029/2002JD002195, http://dx.doi.org/10.1029/2002JD002195, 2002.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, 2002.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast er-

ror statistics, Journal of Geophysical Research: Atmospheres, 99, 10 143–10 162, doi:10.1029/94JC00572, http://dx.doi.org/10.1029/94JC00572, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, doi:10.1007/s10236-003-0036-9, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Feely, R. A., Wanninkhof, R., Takahashi, T., and Tans, P.: Influence of El Nino on the equatorial Pacific contribution to atmospheric CO_2 accumulation, Nature, 398, 597–601, doi:10.1038/19273, 1999.

Flowerdew, J.: Towards a theory of optimal localisation, Tellus A, 67, 25 257, doi:10.3402/tellusa.v67.25257, 2015.

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K.-R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmospheric Chemistry and Physics, 14, 3855–3864, doi:10.5194/acp-14-3855-2014, https://www.atmos-chem-phys.net/14/3855/2014/, 2014.

Garaud, D. and Mallet, V.: Automatic calibration of an ensemble for uncertainty estimation and probabilistic forecast: Application to air quality, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD015780, http://dx.doi.org/10.1029/2011JD015780, d19304, 2011.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2002JD003018, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD003018, 2003.

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.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter, Monthly Weather Review, 129, 2776–2790, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2, https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2, 2001.

Hilton, T. W., Davis, K. J., Keller, K., and Urban, N. M.: Improving North American terrestrial CO_2 flux diagnosis using spatial structure in land surface model residuals, Biogeosciences, 10, 4607–4625, doi:10.5194/bg-10-4607-2013, https://www.biogeosciences.net/10/4607/2013/, 2013.

Houtekamer, P. L. and Mitchell, H. L.: Data Assimilation Using an Ensemble Kalman Filter Technique, Monthly Weather Review, 126, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2, 1998.

Houtekamer, P. L. and Mitchell, H. L.: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 129, 123–137, doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2, 2001.

Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E. J., Feng, L., Ganshin, A., Hasekamp, O., Jones, D., Maksyutov, S., Marshall, J., Oda, T., O'Dell, C. W., Oshchepkov, S., Palmer, P. I., Peylin, P., Poussi, Z., Reum, F., Takagi, H., Yoshida, Y., and Zhuravlev, R.: An intercomparison of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements, Journal of Geophysical Research: Atmospheres, 120, 5253–5266, doi:10.1002/2014JD022962, http://dx.doi.org/10.1002/2014JD022962, 2014JD022962, 2015.

Jankov, I., Berner, J., Beck, J., Jiang, H., Olson, J. B., Grell, G., Smirnova, T. G., Benjamin, S. G., and Brown, J. M.: A Performance Comparison between Multiphysics and Stochastic Approaches within a North American RAP Ensemble, Monthly Weather Review, 145, 1161–1179, doi:10.1175/MWR-D-16-0160.1, https://doi.org/10.1175/MWR-D-16-0160.1, 2017.

Keenan, T. F., Prentice, I. C., Canadell, J. G., Williams, C. A., Wang, H., Raupach, M., and Collatz, G. J.: Recent pause in the growth rate of atmospheric CO_2 due to enhanced terrestrial carbon uptake, Nature Communications, 7, 13 428–, doi:10.1038/ncomms13428, http://dx.doi.org/10.1038/ncomms13428, 2016.

Kim, J.-S., Kug, J.-S., Yoon, J.-H., and Jeong, S.-J.: Increased Atmospheric CO_2 Growth Rate during El Niño Driven by Reduced Terrestrial Productivity in the CMIP5 ESMs, Journal of Climate, 29, 8783–8805, doi:10.1175/JCLI-D-14-00672.1, https://doi.org/10.1175/JCLI-D-14-00672.1, 2016.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO_2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, doi:10.5194/bg-6-1089-2009, http://www.biogeosciences.net/6/1089/2009/, 2009.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmospheric Chemistry and Physics, 12, 337–354, doi:10.5194/acp-12-337-2012, http://www.atmos-chem-phys.net/12/337/2012/, 2012.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Journal of Geophysical Research: Atmospheres, 121, 5213–5236, doi:10.1002/2015JD024473, http://dx.doi.org/10.1002/2015JD024473, 2016.

Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., Manning, A. C., Boden, T. A., Tans, P. P., Houghton, R. A., Keeling, R. F., Alin, S., Andrews, O. D., Anthoni, P., Barbero, L., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Currie, K., Delire, C., Doney, S. C., Friedlingstein,

P., Gkritzalis, T., Harris, I., Hauck, J., Haverd, V., Hoppema, M., Klein Goldewijk, K., Jain, A. K., Kato, E., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Melton, J. R., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., O'Brien, K., Olsen, A., Omar, A. M., Ono, T., Pierrot, D., Poulter, B., Rödenbeck, C., Salisbury, J., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Sutton, A. J., Takahashi, T., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2016, Earth System Science Data, 8, 605–649, doi:10.5194/essd-8-605-2016, https://www.earth-syst-sci-data.net/8/605/2016/, 2016.

Lei, L. and Anderson, J. L.: Comparisons of Empirical Localization Techniques for Serial Ensemble Kalman Filters in a Simple Atmospheric General Circulation Model, Mon. Wea. Rev., 142, 739–754, doi:10.1175/MWR-D-13-00152.1, 2014.

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., Herndon, S. C., Nehrkorn, T., Zahniser, M. S., Jackson, R. B., Phillips, N., and Wofsy, S. C.: Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts, Proceedings of the National Academy of Sciences, 112, 1941–1946, doi:10.1073/pnas.1416261112, http://www.pnas.org/content/112/7/1941.abstract, 2015.

Ménétrier, B., Montmerle, T., Michel, Y., and Berre, L.: Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variance Filtering and Covariance Localization, Mon. Wea. Rev., 143, 1622–1643, doi:10.1175/MWR-D-14-00157.1, http://dx.doi.org/10.1175/MWR-D-14-00157.1, 2015a.

Ménétrier, B., Montmerle, T., Michel, Y., and Berre, L.: Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part II: Application to a Convective-Scale NWP Model, Mon. Wea. Rev., 143, 1644–1664, 2015b.

Miles, N. L., Richardson, S. J., Davis, K. J., Lauvaux, T., Andrews, A. E., West, T. O., Bandaru, V., and Crosson, E. R.: Large amplitude spatial and temporal gradients in atmospheric boundary layer CO_2 mole fractions detected with a tower-based network in the U.S. upper Midwest, Journal of Geophysical Research: Biogeosciences, 117, n/a–n/a, doi:10.1029/2011JG001781, http://dx.doi.org/10.1029/2011JG001781, 2012.

Miller, S. M., Hayek, M. N., Andrews, A. E., Fung, I., and Liu, J.: Biases in atmospheric CO_2 estimates from correlated meteorology modeling errors, Atmospheric Chemistry and Physics, 15, 2903–2914, doi:10.5194/acp-15-2903-2015, https://www.atmos-chem-phys.net/15/2903/2015/, 2015.

Pannekoucke, O., Berre, L., and Desroziers, L.: Background error correlation length-scale estimates and their sampling statistics, Q. J. R. Meteorol. Soc., 134, 497–508, 2008.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, 104, 18 925–18 930, doi:10.1073/pnas.0708986104, http://www.pnas.org/content/104/48/18925.abstract, 2007.

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, https://www.biogeosciences.net/10/6699/2013/, 2013.

Raynaud, L. and Pannekoucke, O.: Sampling properties and spatial filtering of ensemble background-error length-scales, Quarterly Journal of the Royal Meteorological Society, 139, 784–794, doi:10.1002/qj.1999, http://dx.doi.org/10.1002/qj.1999, 2013.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Let., 28, 175–178, 2001.

Schimel, D., Stephens, B. B., and Fisher, J. B.: Effect of increasing CO2 on the terrestrial carbon cycle, Proceedings of the National Academy of Sciences, 112, 436–441, doi:10.1073/pnas.1407302112, http://www.pnas.org/content/112/2/436.abstract, 2015.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, National Center of Atmospheric Research, Tech. Note, NCAR/TN-475+STR, 113pp, 2008.

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO_2, Science, 316, 1732–1735, doi: 10.1126/science.1137004, 2007.

Viatte, C., Lauvaux, T., Hedelius, J. K., Parker, H., Chen, J., Jones, T., Franklin, J. E., Deng, A. J., Gaudet, B., Verhulst, K., Duren, R., Wunch, D., Roehl, C., Dubey, M. K., Wofsy, S., and Wennberg, P. O.: Methane emissions from dairies in the Los Angeles Basin, Atmospheric Chemistry and Physics, 17, 7509–7528, doi:10.5194/acp-17-7509-2017, https://www.atmos-chem-phys.net/17/7509/2017/, 2017.

Wu, L., Bocquet, M., Chevallier, F., Lauvaux, T., and Davis, K.: Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions, Tellus B: Chemical and Physical Meteorology, 65, 20 894, doi:10.3402/tellusb.v65i0.20894, http://dx.doi.org/10.3402/tellusb.v65i0.20894, 2013.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO_2 and CH_4 Retrieved from GOSAT: First Preliminary Results, SOLA, 5, 160–163, doi:10.2151/sola.2009-041, 2009.