Wintertime CO_2, CH_4 and CO emissions estimation for the Washington DC / Baltimore metropolitan area using an inverse modeling technique

Israel Lopez-Coto,[*,†] Xinrong Ren,[‡,¶] Olivia E. Salmon,[§,⊥] Anna Karion,[†] Paul B. Shepson,[§,‖] Russell R. Dickerson,[‡] Ariel Stein,[¶] Kuldeep Prasad,[†] and James R. Whetstone[†]

†*National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899*

‡*University of Maryland, 4254 Stadium Dr, College Park, MD 20742*

¶*Air Resources Laboratory, NOAA, 5830 University Research Court, College Park, MD 20740*

§*Purdue University, 610 Purdue Mall, West Lafayette, IN 47907*

‖*Stony Brook University, 100 Nicolls Rd, Stony Brook, NY 11794*

⊥*now at Lake Michigan Air Directors Consortium, 101 S Webster St, Madison, WI 53703*

E-mail: israel.lopezcoto@nist.gov

Abstract

Since greenhouse gas mitigation efforts are being mostly implemented in cities, the ability to quantify emission trends for urban environments is of paramount importance. However, previous aircraft work has indicated large daily variability in the results. Here we use measurements of CO_2, CH_4 and CO from aircraft over five days within an inverse model to estimate emissions from the D.C./Baltimore region. Results show good agreement with previous estimates in the area for all three gases. However, aliasing caused by irregular spatiotemporal sampling of emissions is shown to significantly impact both the emissions estimates and their variability. Extensive sensitivity tests allow us to quantify the contributions of different sources of variability and indicate that daily variability in posterior emissions estimates is larger than the uncertainty attributed to the method itself (i.e. 17% for CO_2, 24% for CH_4 and 13% for CO). Analysis of hourly reported emissions from power plants and traffic counts shows that 97% of the daily variability in posterior emissions estimates is explained by accounting for the sampling in time and space of sources that have large hourly variability and, thus, caution must be taken in properly interpreting variability that is caused by irregular spatiotemporal sampling conditions.

Introduction

As cities move toward mitigating their carbon footprints, estimating their emissions using atmospheric observations is a valuable way to assess the efficacy of mitigation policies. Recent work[1–7] has already demonstrated the capability of top-down (atmospheric measurement-based) estimation methods to inform bottom-up inventory methods for some greenhouse gases (GHGs). On regional and urban scales, top-down methods have been shown to be effective at estimating emissions using either tower-based or aircraft-based concentration measurements.[8–12]

Atmospheric trace gas concentration measurements from airborne platforms have been used extensively to estimate emissions from a region. Both oil and gas basins and urban regions have been studied using mass balance methods,[13–17] including the Washington D.C./ Baltimore metropolitan area.[11,12] Researchers have also used aircraft observations with transport models in an inversion framework to estimate emissions at regional,[18–21] urban[22,23] and local scales.[24]

Several studies have investigated the source of daily variability in aircraft-based top-down emissions estimates for a given region. Variability in estimated emission rates has previously been attributed to uncertainty in the mass balance methodology, which would confound or obscure real emissions changes.[25,26] More recent work using airborne measurements over oil and gas fields has shown that temporal variability in emissions must be considered when interpreting estimates from single-day flights, however. Lavoie et al.[27] found significant temporal variability in single source emissions of methane (CH_4) from the Eagle Ford oil and gas production basin in Texas, while Schwietzke et al.[28] investigated the effect of episodic CH_4 emissions from natural gas facilities on the regional mass balance estimates in the Fayetteville Shale.

In this study, we use observations collected during five aircraft flights over a two-week period in February 2016 within a Bayesian inversion framework to: 1) estimate emissions of CO_2, CH_4 and CO from the cities of Washington D.C. and Baltimore, MD, (Fig. 1), 2) quantify the uncertainty, and its sources, in each day's emissions estimate and, 3) explain the cause for the observed daily variability in the estimated emissions.

To this end, we use an ensemble of inversions where prior emissions, transport model and observation dataset were varied. Ensemble spread and correlations between six transport models were used to construct the full model-data mismatch covariance matrix, and the background mole fraction was first estimated by using sensitivities to nearby outside sources and then further optimized within the inversion. Additionally, sensitivity tests were conducted investigating the impacts of background choice, omitting correlations in the transport error covariance matrix and changing the magnitude of the prior emission errors. We use the inversion ensemble and sensitivity tests to quantify the different sources of variability and, thus, understand the uncertainty inherent in the inverse methodology. We then investigate daily variability in estimated emissions and to what extent this variability can be explained by aliasing caused by irregular sampling of spatial and temporal variability in large sources within the study domain.

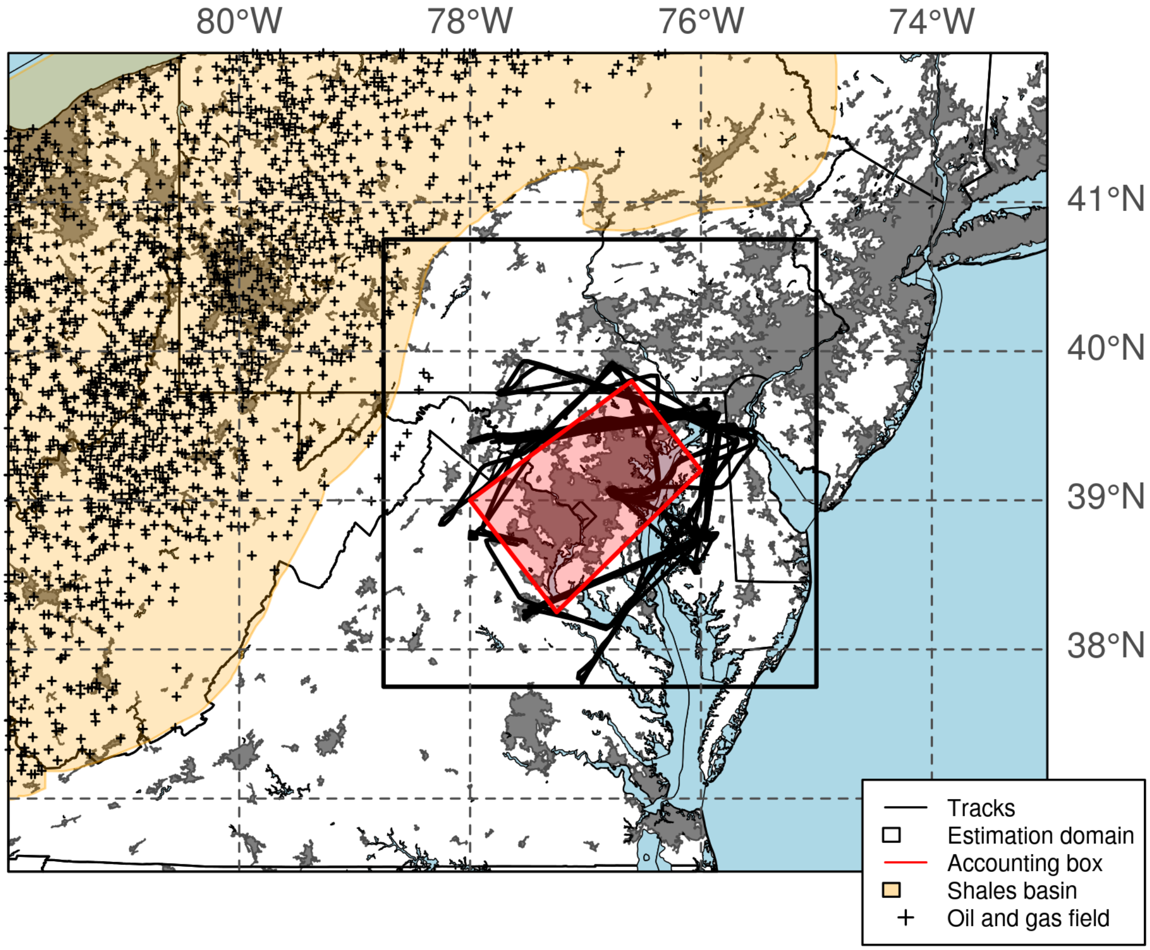


Figure 1: Computational domain (0.03° resolution) showing the inversion domain (black rectangle) and the outer domain (entire map) used to account for nearby outside sources. Flight tracks, Census-designated urban areas (gray shaded regions), the Marcellus, Devonian (Ohio) and Utica shale plays in the Appalachian basin and locations of the geometric center (centroid) of the oil and gas fields are also shown.[29] Total emissions are reported here within the accounting box (red polygon) defined by the corners: (39.80°N, 76.60°W), (39.00°N,78.00°W), (38.25°N, 77.25°W) and (39.20°N, 76.00°W).

Methods

Observations

Trace gas observations from two airborne platforms were used in this study: Purdue University's Beechcraft Duchess, housing the Airborne Laboratory for Atmospheric Research, or ALAR, (Purdue) and the University of Maryland's Cessna 402B research aircraft (UMD).

The two aircraft flew simultaneously for 5 days, mostly during afternoon hours, collecting trace gas mole fraction and meteorological data along transects at different altitudes that covered the full depth of the PBL (Fig. 1 and SI for further details). To determine the effect of withholding observations from the inversion system, we alternatively used CO_2 and CH_4 observations from both aircraft, the UMD aircraft alone, or the Purdue aircraft alone, as part of the ensemble of inversions. Purdue did not measure CO, thus the CO inversions used only UMD observations.

Bayesian Inversion Framework

We estimate trace gas emissions using a Bayesian inverse analysis[30,31] as in Lopez-Coto et al.[32] Optimum posterior estimates of fluxes are obtained by minimizing the cost function J:

$$J(\boldsymbol{x}) = \frac{1}{2}\left[(\boldsymbol{x}-\boldsymbol{x}_b)^T \mathbf{P}_b^{-1}(\boldsymbol{x}-\boldsymbol{x}_b) + (\mathbf{H}\boldsymbol{x}-\boldsymbol{y})^T \mathbf{R}^{-1}(\mathbf{H}\boldsymbol{x}-\boldsymbol{y})\right] \tag{1}$$

where $\boldsymbol{x}_b$ is the first guess or a priori state vector, $\mathbf{P_b}$ the *a priori* error covariance matrix which represents the uncertainties in our *a priori* knowledge about the fluxes and $\mathbf{R}$ the error covariance matrix, which represents the uncertainties in the observation operator $\mathbf{H}$ and the observations $\boldsymbol{y}$, also known as model-data mismatch. The observation operator $\mathbf{H}$ is constructed using the sensitivity of observations to surface fluxes, or footprints (units: ppm μmol^{-1} m^2 s) generated with a transport model. Here we modify the formulation to include optimization of the background in the inversion (see SI for details).

Transport Models

In order to generate an ensemble of transport models and therefore better represent the uncertainties, NOAA's Hybrid Single-Particle Lagrangian Integrated Trajectory dispersion model (HYSPLIT)[33] was driven with 5 different meteorological products: the High Resolution Rapid Refresh (HRRR) NOAA operational forecast product[34] and 4 configurations of the Weather Research and Forecasting model (WRF[35]) provided by the National Center for Atmospheric Research (NCAR) that included 4 different PBL parametrizations, 2 sources of initial and boundary conditions and the inclusion of the Building Energy Parameterization (BEP) urban canopy model in one of the configurations. In addition, the vertical mixing option in HYSPLIT also varied (Table S1 and SI for details).

Emissions Inventories

Nine CO_2 emissions inventories were used in the inversion to investigate the resultant variability in the posterior emissions (Table S2). Four of them (Vulcan (VU[36]), ODIAC (OD[37]), FFDAS (FF[38]) and ACES (AC[39]) are existing anthropogenic CO_2 inventories but for a different year; one provided only on-road emissions (DARTE (DA[40])); one is the mean of the previous five (EB); and the rest (flat (FL) and simple (SP[32])) are constructed here to complement the ensemble of prior fluxes. In addition, we use the ACES mean for February between 12 - 19 EST (AC2). Since DARTE only provided on-road emissions, a simple calculation of urban emissions was used to complement it. CH_4 prior emissions were represented using EPA's gridded inventory (EP) for 2012,[41] EDGAR v4.3.2[42] for 2012 (EG), the mean of the previous two (EB), and a flat prior (FL). For CO we use EDGAR v4.3.2 (EG),[43] the National Emissions Inventory (NEI) for 2011 from EPA (NI,[44]), the annual mean ACES inventory (AC as in the CO_2 case) scaled using the mean observed ΔCO:ΔCO_2 ratio (6.18 ppb/ppm) and, again, a flat prior (FL).

Background Determination

Properly accounting for the background is critical for the inversion as the flux correction is based on the observed enhancements above the background value. The impact of upwind sources can be important especially in areas such as the one under study here, where multiple sources exist in the surroundings (Fig. 1). Thus, we estimated the contribution from outside the domain using a Lagrangian approach by convolving footprints from a reduced set of our ensemble of transport models and prior fluxes. We extended the domain to the full extent shown in Fig. 1. The full background was then represented as the ensemble mean of the contribution from outside of the domain of interest ($\boldsymbol{y_{oc}}$, time-varying along the track) plus the long-range background ($\boldsymbol{y_{lr}}$, constant for a given flight). This methodology provided a time varying *a priori* background that included uncertainties that was then further optimized in the inversion (SI).

Error Covariances

i) Prior Flux Error Covariance

The prior flux error covariance represents the uncertainties in the prior estimation of the fluxes. Although bottom-up CO_2 emissions estimates are made on global and national scales with small uncertainties, considerable errors are introduced when the emissions are disaggregated to grid cells, due to the usage of proxies to spatially distribute emissions.[45] Reported errors at grid cell levels range from 4% to more than 190%, averaging about 120%.[46] For CH_4 and CO it is likely that the errors at grid cell levels are even larger than for CO_2 because of the less well-known characteristics of these species' sources. Given these reported uncertainties at grid cell levels, we use a value of 100% of the grid cell emissions as uncertainty in this work for all the prior inventories and gases with the exception of FFDAS where we use a scaled up version of the provided uncertainties and the EB case for CO_2 where we use the standard deviation of the ensemble at each pixel to represent the uncertainties. In all cases, a covariance exponential model in space was assumed. (See SI for details)

ii) Outside Contribution (background) Prior Error Covariance

We consider a double exponential model, in space and time, to represent the error covariance of the outside contribution ($\boldsymbol{y_{oc}}$) along the track. The diagonal is populated with the uncertainty of the initial guess outside contribution based on the variance from the different transport models and prior fluxes (SI).

iii) Model Error

The model-data mismatch error covariance was assumed to have three independent contributions: 1) uncertainty in the observations, 2) uncertainty in the long-range background concentration and 3) uncertainty in the transport model representation. The uncertainties in the observations have their origin in the measurement uncertainties and the representativity of the assigned mean to the averaging period (one minute in our case). This contribution is not correlated and thus the covariance was considered diagonal. The long-range background ($\boldsymbol{y_{lr}}$) determination also introduces uncertainty into the system. This contribution was also assumed to be uncorrelated. Lastly, the transport model uncertainty is complex with several previously published methods for its determination. Here we tested two methods, both based on the ensemble of transport models. First, we tested a diagonal covariance populated with the inter-model variance simulated using the same surface fluxes (the prior emissions in each inversion case) in all the transport models similar to Engelen et al.[47] and Desroziers et al.[48] As stated in Engelen et al.,[47] this estimate can be too large for some models and too small for other models, thus, in order to better represent the fidelity of each model and for each observation, we weighted the inter-model standard deviation with the relative error computed by using the wind measurements from the aircraft. This definition of the transport model error covariance assumes there are no correlations in space and time which is unlikely to be true. Therefore, for the second method, which was used in the main ensemble of inversions, we computed the correlations between the different transport models and included them in the covariance matrix, leaving the first method as a sensitivity test (see SI for details).

Sensitivity Analysis

As described in the previous sections, the main inversion ensemble was composed by different prior emissions (9 for CO_2, 4 for CH_4 and 4 for CO), 6 transport models and 3 combinations of the observations for 5 flight days, totaling 1,290 inversions (810 + 360 + 120). This inversion ensemble was configured with the background and prior and transport error covariances choices that are most reasonable for the analysis. However, in order to additionally test the sensitivity of the posterior estimates to inversion setup choices that might not be as appropriate, we also investigated the effects of changing the background determination, the transport error covariance, and the prior flux error covariance, separately from the main inversion ensemble. Specifically, for the background test, we performed the inversion 1) without optimizing the Lagrangian background, 2) scaling the Lagrangian background, and 3) selecting a single constant value along the track as background defined by the 1st, 5th or 10th percentile, to compare with our base case of optimizing the background (OBC1). For the scaled background case, a single scaling factor for each flight was applied to the background time series. This scaling factor was the ratio of posterior to prior emissions for the inversion case where the background was not optimized or scaled. We also tested the impact of using only a diagonal transport error covariance as well as reducing and increasing the uncertainty in the prior fluxes (50%, 100% and 200%). This sensitivity test resulted in a total of 12 cases with 15,480 individual inversions, (Table S3).

Both the main inversion ensemble and the sensitivity test were analyzed in the same fashion, grouping by cases (prior, transport, day, observation dataset or sensitivity case) and then computing the mean and quantiles as shown in Figs. 2, S7, S10, S13 and S16. The variability associated with each grouping was then computed as the standard deviation among each case's mean value.

Normalized Observed Emissions

We construct an analysis to investigate whether the hourly variability of the energy generation and traffic sectors' emissions, combined with the specific flight pattern on a given day, can explain the daily variability in the posterior CO_2 estimates. Both of these sources have publicly available data at the hourly level: Continuous Emissions Monitoring System (CEMS[49]) data for power plants and Travel Monitoring Analysis System (TMAS[50]) data for traffic counts. First, we sum all the power plant emissions and traffic counts within the footprint (we use the ensemble mean footprint as a mask) of each observation used in the inversion and within the defined accounting box. We match the hourly power plant emissions and traffic counts with the observation time, accounting for transport time to the point of the observation at hourly temporal resolution. Then we average this value (the sum of all traffic counts or powerplant emissions within each footprint) over all observations in each flight for each of the five flights. Using an average allows us to account for the difference in the number of observations per flight. Because traffic counts and power plant emission rates are in different units, we define the normalized observed emissions (nOE), allowing for the combination of the two sectors. We normalize counts and power plant emissions each to their respective campaign mean so that the campaign mean is equal to one. Furthermore, we use the relative contribution of the different sectors in the ACES 2011 annual mean [39] within the defined accounting box to construct the normalized observed emissions (nOE) for each flight as follows:

$$nOE_i = f_e \frac{CEMS_i}{\langle CEMS \rangle} + f_r \frac{TMAS_i}{\langle TMAS \rangle} + 1 - (f_e + f_r) \tag{2}$$

In the above definition, i is the index indicating the flight, f_e is the contribution of the electricity production sector (16%) and f_r is the contribution of the traffic emissions (46%) in ACES. The last term of Eq. 2 represents the remainder of anthropogenic CO_2 emission sectors. By this construction, the mean nOE for the campaign is also equal to 1.

Results

In the following subsections we present the main results of the analysis and discuss the variability and uncertainty of the emissions estimates. In this context, the terms variability and uncertainty are not used as synonyms. Rather, we use the term variability to describe how a property (posterior total emissions for the most part) changes (varies) with respect to different variables like time, space or model choices. The term uncertainty refers to the ability of the inverse method to represent the measurand, and it combines all sources of variability for a single day's estimate.

Emissions Rates

Our mean estimates for the defined accounting box are 87 ± 28 kmol s^{-1} for CO_2, 0.42 ± 0.12 kmol s^{-1} for CH_4 and 0.59 ± 0.16 kmol s^{-1} for CO (mean ± 1-σ) where the bounds presented here represent the posteriors' daily variability. Ren et al.[11] using a mass balance method, estimated emission rates of 96 kmol s^{-1} for CO_2, 0.57 ± 0.28 kmol s^{-1} for CH_4 and 0.55 ± 0.27 kmol s^{-1} for CO using the same flight observations as this study. In addition, Salmon et al.[12] estimated a CO emission rate (also using a mass balance method) of 0.54 ± 0.47 kmol s^{-1} in February 2015. Our estimates are consistent with these within 1-σ uncertainties for both methods.

The applied inversion methodology corrected the prior inventories (Fig. 2a,c,e) by quite different amounts leading to consistent results in the posterior emissions, with variability due to choice of prior of 11%, 13% and 6% (or 9.6, 0.055 and 0.035 kmol s^{-1}) for CO_2, CH_4 and CO respectively (1-σ), significantly lower than the variability of the prior values themselves (flat prior included), 41%, 65% and 87% (or 20.8, 0.097 and 0.38 kmol s^{-1}). The flat (FL) prior led to the largest range and IQR for all of the three gases due to the loose constraint it imposed on the inversion. For CO_2, the FFDAS[38] prior (FF) resulted in the lowest posterior estimates as well as the lowest range and IQR due to the low prior uncertainty assigned, making it hard for the inversion to deviate from the prior values. For CH_4, the inversions using the 2012 EPA gridded inventory[41] (EP) as a prior provided the lowest estimates, probably due to the lower prior emissions allocated into the urban areas and, therefore, lower prior uncertainties, making it harder to correct those areas. For CO, the scaled ACES inventory (AC) led to the lowest estimates. Variability due to transport model choice was 15% for CO_2, 13% for CH_4 and 16% for CO (1-σ), (Figs. S7c, S10c and S13c). We note that HR and MY2 provided the highest and lowest estimates respectively, while MY and BL had the most variable results. The observation dataset choice impacted the results the least, with only a 6 % standard deviation of the mean for CO_2 and 10% for CH_4 with very similar range and IQR for each of the three cases (Figs. S7d, S10d). In contrast to the relatively small effect of varying these three model choices (prior, transport model, and observation dataset), the daily variability of the estimates was 33% for CO_2 and 28% for CH_4 and CO (1-σ) (Figs. 2b,d,f). The mean estimates for each day do not overlap with the IQR of the other days and while the CO_2 and CO estimates follow a very similar pattern (as they have similar sources), they differ from that of CH_4. In addition, the coefficient of determination between the daily emission estimates for the three gases is r^2=0.90 for CO vs CO_2, r^2=0.40 for CO_2 vs CH_4 and r^2=0.19 for CO vs CH_4. This suggests that the inversion is actually providing different estimates for each day, and that the posterior differences between days are not only the result of choices in the model set up.

The spatial distribution of the averaged CO_2 posterior emissions for each prior case shows that most of the emissions are coming from the urban areas, even in the flat prior case (Fig. S8). The results show that the roads (traffic emissions) and fine spatial scale features are only resolved in modeling results when high resolution inventories are used as the prior emissions. The inversion was able to spatially differentiate between the cities of Baltimore and Washington DC, correcting their emissions differently (Fig. S9): emissions from Washington, DC were adjusted upward in all cases while those from Baltimore were corrected downward in the cases of AC, AC2 and VU and only slightly upward for the rest.

The spatial distribution of the averaged CH_4 posterior emissions (Fig. S11) indicates that while some emissions are from urban areas, significant emissions occur NNE and NNW of the Washington - Baltimore metropolitan area as well, which is different than for CO_2. All the CH_4 priors were corrected upwards indicating an overall underestimation of emissions in the existing inventories (Fig. S12), with the strongest corrections applied to point sources outside urban areas. However, the urban areas were also corrected upward, with this correction being larger for EP than for EG or EB cases. Our posterior mean ratio to the 2012 EPA gridded

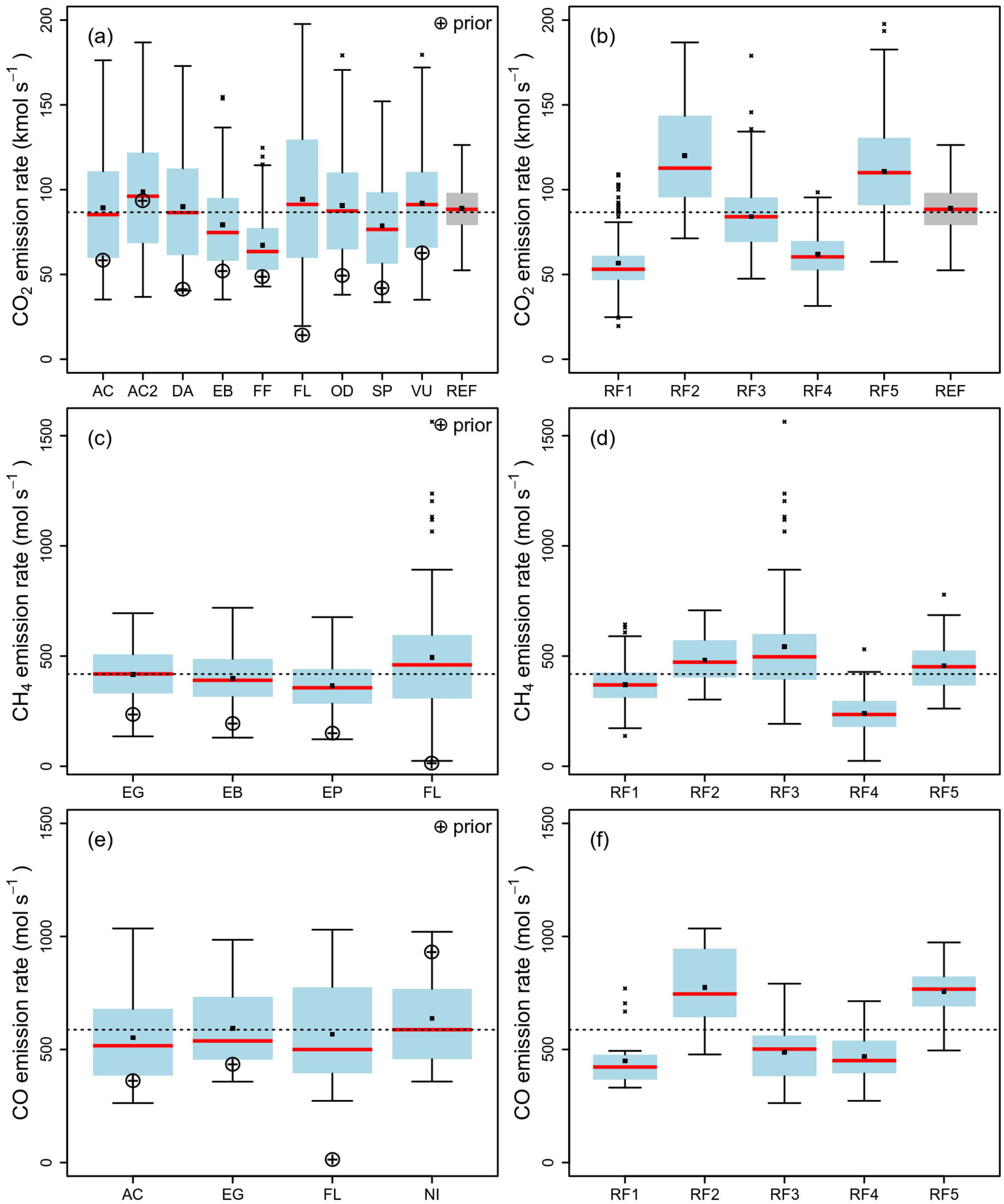


Figure 2: Boxplots of the total CO_2, CH_4 and CO estimated emission rate within the accounting box grouped by: (a,c,e) the different inventories used as priors and (b,d,f) the different research flights. The grey bar in panels (a,b) are the values provided by ACES, scaled to totals of 2016, for February between 12 - 19 EST (referred as REF). Blue bars indicate the 25^{th} to 75^{th} range, whiskers the range up to 1.5 times the IQR, x's the outliers ($>$ 1.5 x IQR), red line the median, square markers the mean and the dotted line the posterior mean. The circled pluses in panel (a,c,e) represent each prior's total emissions. (See methods section and Tables S1 and S2 for abbreviations)

inventory (EP),[41] 2.73 ± 0.76, is in very good agreement with Ren et al.'s[11] estimate of 2.8 times the EPA values for the same region.

The spatial distribution of the mean posterior CO fluxes (Fig. S14) indicates that the CO emissions largely originate in the urban areas, as they do for CO_2. In addition, the correction (Fig. S15) is mostly applied in the urban cores, increasing the fluxes for AC, FL and EG while strongly decreasing the emissions for NI case. Due to the construction of AC for CO (using the ACES CO_2 inventory scaled by mean observed $\Delta CO:\Delta CO_2$ ratio), power plant emissions were present in the prior, while we expect the power plants ratio to be small compared to other sources. The inversion was able to correct down at least a few of them (blue dots in Fig. S15a). The NEI CO prior case was strongly corrected down over all urban areas, even in Philadelphia, indicating that the inversion is able to correct underestimation as well as overestimation in the prior. The NEI CO overestimation has been extensively reported in the literature;[22,23,51] specifically in the DC/Baltimore region a close to 50 % overestimation of the NEI CO inventory has been reported,[11,12] similar to our result of 58%.

Sensitivity Analysis

For all three gases, the diagonal model-data mismatch error covariance (EDC1) provided larger emissions estimates than the equivalent full covariance case (C1) (Fig. S16). In addition, the range and IQR within each case was larger with the diagonal covariance indicating that the off-diagonal terms played an important role in limiting the number of possible solutions. The background selection impacted both the mean estimates and the range and IQR indicating that incorrect background specification can bias the estimation results. The prior flux error sensitivity test showed that posterior emissions estimates were larger when prior uncertainties were doubled, and the range and IQR within each case was also larger indicating a potential over-fitting problem. When prior uncertainties were halved from the base case, the estimates were lower and less variable, indicating the solutions were more constrained by the prior fluxes than by the observations. This effect was similar to the FF-

DAS prior case for CO_2, for which the prior uncertainties were likely too small. Despite the differences described above, the variability of the mean across the sensitivity analysis cases remained relatively low, at 11% for CO_2, 17% for CH_4 and 8% for CO.

Special Case: Flat Prior

The inversions using a spatially flat prior (FL) were able to provide mean totals close to those in which an inventory prior was used for all three gases (Fig. 2). This result demonstrates the potential for using aircraft measurements to estimate an overall city-wide emission rate for a location where a spatially explicit inventory or other emissions information is unavailable. However, we have also shown that the range and IQR in the flat prior case was the greatest among all the prior cases, implying that when using a flat prior sampling more time periods (i.e. using more observations) is required to provide confidence in the estimates. The spatial distribution of the campaign-averaged posterior fluxes for CO_2, CH_4 and CO (Fig. S17) is consistent with the results obtained with the other priors as well. For example, CO_2 and CO show very similar spatial distributions with most of the emissions originating in the urban areas while CH_4 shows a broader spatial distribution. Note that these spatial patterns are a result of a campaign of 5 days with winds coming from different directions (Fig. S1), leading to a good triangulation of the source locations.

Discussion: Uncertainty and Sources of Variability

Method Combined Uncertainty

We were able to disentangle and quantify the different sources of variability present in the inversion-based emissions estimates and found that the largest source of variability in the retrieved emissions is the daily variability. In the following analysis, we omit the daily variability because the goal is to understand the uncertainty we expect in each day's estimate and whether the daily variability is likely to be caused by general uncertainty in the method.

Here we assume each source of variability is independent of the others, so that the variances can be summed to estimate the method uncertainty in each day's estimate. We note that the assumption of independence is not likely to be true and therefore this uncertainty estimation might be biased due to not considering the correlations among them. In addition, the ensemble construction (transport, priors, observation dataset, covariances and background) might impact the ensemble spread and therefore might not be the true uncertainty in the method but it does, however, provide an indication of the likely variability introduced by the different model choices.

Three different cases are shown in Table 1 for estimating combined uncertainties. The Combined Uncertainty 1 case considers all sources of variability tested in the inversion. However, we believe that two transport models are outliers that suffer from improper mixing and resulted in biased estimations. The highest retrieved fluxes are obtained consistently using the HR configuration, suggesting that this configuration is too dispersive, although more research is needed to be more certain. The lowest posterior estimates consistently correspond to the configuration including the experimental vertical mixing parametrization (MY2), indicating that this method may under-predict vertical mixing. Removing these two outlier configurations reduces the variability due to transport model choice to 7% for CO_2, 10% for CH_4 and 8% for CO; these are used to calculate the Combined Uncertainty 2 case. Because the flight tracks are different for each aircraft, the variability due to the observation dataset may also be affected by the spatial and temporal distribution of the sources being measured, so we remove this variability to also calculate the Combined Uncertainty 3 case.

Daily Variability in Estimated Emissions: Aliasing

The daily variability in our posterior emissions from the inversion ensemble was 33% for CO_2 and 28% for CH_4 and CO (Table 1). This variability is larger than each individual source of variability as well as the three cases of the combined uncertainties as shown above, although for CH_4 the two are comparable. In order to better understand the origin of this variability

Table 1: Sources of variability and combined uncertainty.

Source of uncertainty	ϵ CO_2 (%)	ϵ CH_4 (%)	ϵ CO (%)
daily	33	28	28
prior	11	13	6
transport	15	14	16
transport no outliers	7	10	8
observation dataset	6	10	6[a]
sensitivity	11	17	8
Combined Uncertainty 1 (prior, transport, dataset and sensitivity)	22	27	20
Combined Uncertainty 2 (prior, transport no outliers, dataset and sensitivity)	18	26	14
Combined Uncertainty 3 (prior, transport no outliers and sensitivity)	17	24	13

[a] CO variability due to the observation dataset is assumed to be the same as for CO_2.

in the estimates, we conducted an analysis of the temporal variability and spatial sampling of the two largest sources of CO_2 in the accounting box, according to the ACES inventory: energy generation and on-road traffic.[39] Thirteen power plants and 87 counting stations were used within the accounting box (Fig. S18). Both of these sources have significant variability throughout a single day, with traffic counts in the area varying by up to a factor of 20 between night time and evening rush hour depending on the location (Fig. S19, S20), and individual power plant reported emissions varying up to a factor of two within a single day, but even more between days as they sometimes shut down completely (Fig. S21). If daily means of these emissions are investigated, neither the average daily mean of powerplant emissions nor the average daily mean of traffic counts correlates with the daily mean emissions from our inversion posterior. However, the daily variability in the posterior estimates can indeed be explained using an analysis that considers the hourly variability of these two sectors' emissions, combined with the specific flight pattern on a given day. We define each day's *normalized observed emissions* (nOE, Eq. 2) using powerplant and traffic count data to conduct this analysis.

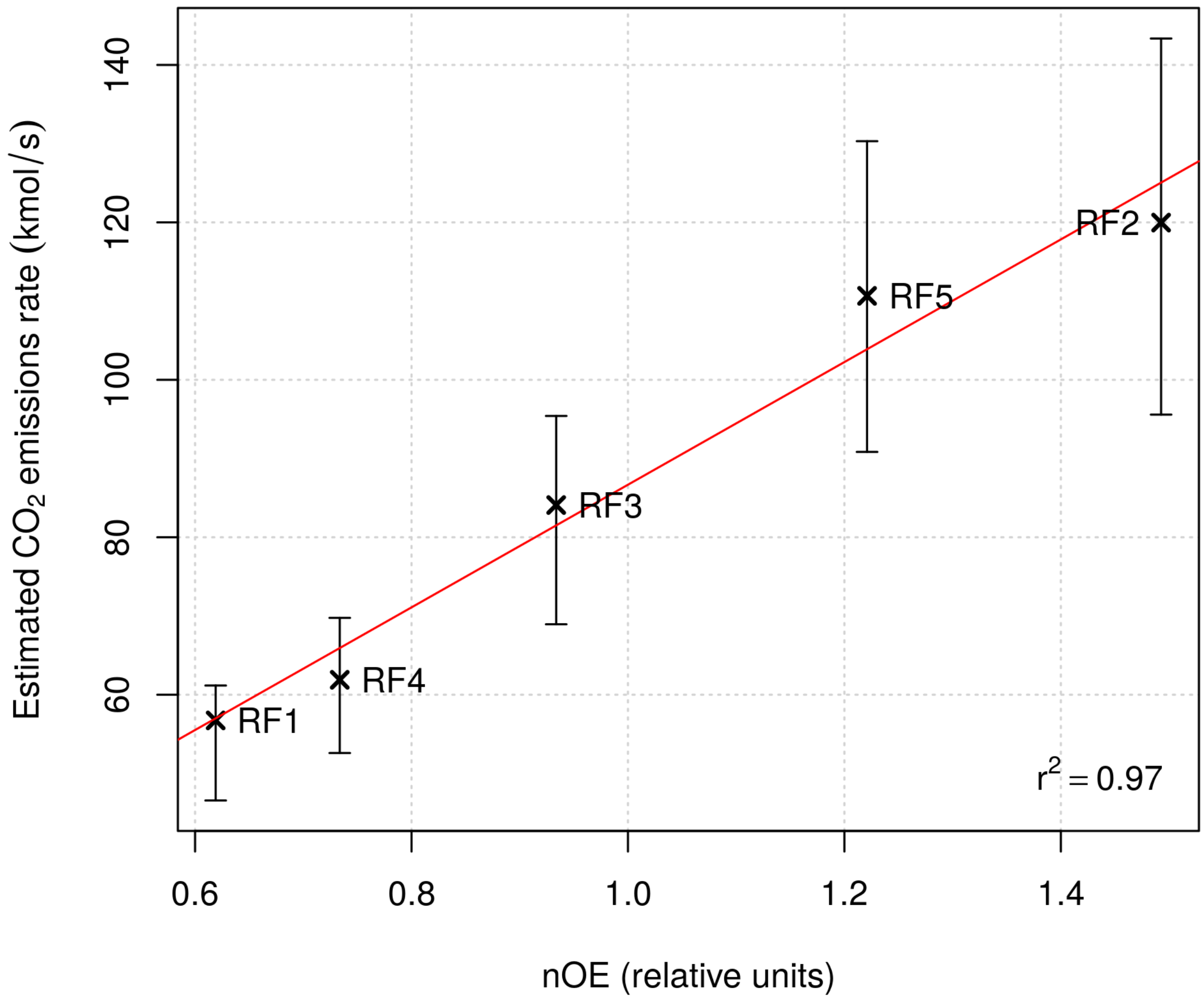


Figure 3: Estimated CO_2 emission rates (kmol s^{-1}) for each research flight as a function of the normalized observed emissions (nOE) computed using CEMS and TMAS hourly data. Errors bars correspond to 25^{th} and 75^{th} percentiles of the ensemble of inversions for each day. Red line indicates the linear fit.

Fig. 3 shows the daily mean estimated CO_2 emissions from the inversion as a function of the *normalized observed emissions* (nOE), with error bars representing the 25^{th} and 75^{th} percentiles of the ensemble of inversions for each day. The correlation between the two is nearly perfect ($r^2 = 0.97$), implying that the daily variability observed by the inversion is caused by irregular spatiotemporal sampling (aliasing) of the rapidly changing underlying emissions.

These results, showing that aliasing of large hourly variability in emissions from large

CO_2 sources can explain 97% of the variability in our CO_2 emissions estimate, suggest that similar spatiotemporal variability in CO and CH_4 sources could explain the variability in our estimates for those gases as well. We note that for CH_4 this is less clear due to the larger estimated uncertainty in the posterior emissions, but it is plausible given that large temporal variability in CH_4 source emissions has been reported in oil and gas production fields,[27,52] and likely exists in urban areas as well.

Path Forward

Flight campaigns are extremely useful for greenhouse gas (GHG) and pollutant emissions estimation because of the fast deployment and large spatial coverage that is provided by a moving platform. However, they are limited by the reduced temporal coverage as well as the difficulty of measuring all the areas at the same time. We have shown that this irregular sampling (in time and space) generates aliasing of the emissions impacting both the emissions estimates and the variability of those estimates. Therefore, moving forward, multiple flights over a region over different hours, days, months and seasons are recommended as well as multiple aircraft flying together with well-coordinated flight plans based on forecast back-trajectories so that the coverage of the cities can be maximized at all times. Addition of measurements from every platform (surface, aerial or from space) available should also help reduce the aliasing of emissions. This aliasing of emissions is likely not exclusive to aircraft campaigns but rather a ubiquitous problem to all monitoring systems based on spatiotemporally discrete sampling (aircraft, cars, polar orbiting satellites as well as sparse tower networks) and it must be considered when designing the measurements and interpreting the results.

Acknowledgement

The authors acknowledge Stuart McKeen from NOAA ESRL Chemical Sciences Division and Ravan Ahmadov NOAA ESRL Global Systems Division for providing the NEI-2011 data. Funding was provided by the NIST Greenhouse Gas measurements program.

Supporting Information Available

- supplemental-information.pdf: Detailed methodology and supplemental tables and figures

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) 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.; 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* **2015**, *112*, 1941–1946.

(2) Staufer, J.; Broquet, G.; Bréon, F.-M.; Puygrenier, V.; Chevallier, F.; Xueref-Rémy, I.; Dieudonné, E.; Lopez, M.; Schmidt, M.; Ramonet, M.; Perrussel, O.; Lac, C.; Wu, L.; Ciais, P. The first 1-year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion. *Atmospheric Chemistry and Physics* **2016**, *16*, 14703–14726.

(3) Henne, S.; Brunner, D.; Oney, B.; Leuenberger, M.; Eugster, W.; Bamberger, I.; Meinhardt, F.; Steinbacher, M.; Emmenegger, L. Validation of the Swiss methane emission

inventory by atmospheric observations and inverse modelling. *Atmospheric Chemistry and Physics* **2016**, *16*, 3683–3710.

(4) Brunner, D.; Arnold, T.; Henne, S.; Manning, A.; Thompson, R. L.; Maione, M.; O'Doherty, S.; Reimann, S. Comparison of four inverse modelling systems applied to the estimation of HFC-125, HFC-134a, and SF_6 emissions over Europe. *Atmospheric Chemistry and Physics* **2017**, *17*, 10651–10674.

(5) Sargent, M.; Barrera, Y.; Nehrkorn, T.; Hutyra, L. R.; Gately, C. K.; Jones, T.; McKain, K.; Sweeney, C.; Hegarty, J.; Hardiman, B.; Wang, J. A.; Wofsy, S. C. Anthropogenic and biogenic CO2 fluxes in the Boston urban region. *Proceedings of the National Academy of Sciences* **2018**, *115*, 7491–7496.

(6) Bergamaschi, P.; Karstens, U.; Manning, A. J.; Saunois, M.; Tsuruta, A.; Berchet, A.; Vermeulen, A. T.; Arnold, T.; Janssens-Maenhout, G.; Hammer, S.; Levin, I.; Schmidt, M.; Ramonet, M.; Lopez, M.; Lavric, J.; Aalto, T.; Chen, H.; Feist, D. G.; Gerbig, C.; Haszpra, L.; Hermansen, O.; Manca, G.; Moncrieff, J.; Meinhardt, F.; Necki, J.; Galkowski, M.; O'Doherty, S.; Paramonova, N.; Scheeren, H. A.; Steinbacher, M.; Dlugokencky, E. Inverse modelling of European CH_4 emissions during 2006–2012 using different inverse models and reassessed atmospheric observations. *Atmospheric Chemistry and Physics* **2018**, *18*, 901–920.

(7) Lamb, B. K.; Cambaliza, M. O. L.; Davis, K. J.; Edburg, S. L.; Ferrara, T. W.; Floerchinger, C.; Heimburger, A. M. F.; Herndon, S.; Lauvaux, T.; Lavoie, T.; Lyon, D. R.; Miles, N.; Prasad, K. R.; Richardson, S.; Roscioli, J. R.; Salmon, O. E.; Shepson, P. B.; Stirm, B. H.; Whetstone, J. Direct and Indirect Measurements and Modeling of Methane Emissions in Indianapolis, Indiana. *Environmental Science & Technology* **2016**, *50*, 8910–8917.

(8) 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.; 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* **2016**, *121*, 5213–5236.

(9) Mays, K. L.; Shepson, P. B.; Stirm, B. H.; Karion, A.; Sweeney, C.; Gurney, K. R. Aircraft-Based Measurements of the Carbon Footprint of Indianapolis. *Environmental Science & Technology* **2009**, *43*, 7816–7823.

(10) Bréon, F. M.; Broquet, G.; Puygrenier, V.; Chevallier, F.; Xueref-Remy, I.; Ramonet, M.; Dieudonné, E.; Lopez, M.; Schmidt, M.; Perrussel, O.; Ciais, P. An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements. *Atmospheric Chemistry and Physics* **2015**, *15*, 1707–1724.

(11) Ren, X.; Salmon, O. E.; Hansford, J. R.; Ahn, D.; Hall, D.; Benish, S. E.; Stratton, P. R.; He, H.; Sahu, S.; Grimes, C.; Heimburger, A. M. F.; Martin, C. R.; Cohen, M. D.; Stunder, B.; Salawitch, R. J.; Ehrman, S. H.; Shepson, P. B.; Dickerson, R. R. Methane Emissions From the Baltimore-Washington Area Based on Airborne Observations: Comparison to Emissions Inventories. *Journal of Geophysical Research: Atmospheres* **2018**, *123*, 8869–8882.

(12) Salmon, O. E.; Shepson, P. B.; Ren, X.; He, H.; Hall, D. L.; Dickerson, R. R.; Stirm, B. H.; Brown, S. S.; Fibiger, D. L.; McDuffie, E. E.; Campos, T. L.; Gurney, K. R.; Thornton, J. A. Top-Down Estimates of NOx and CO Emissions From Washington, D.C.-Baltimore During the WINTER Campaign. *Journal of Geophysical Research: Atmospheres* **2018**, *123*, 7705–7724.

(13) White, W.; Anderson, J.; Blumenthal, D.; Husar, R.; Gillani, N.; Husar, J.; Wilson, W.

Formation and transport of secondary air pollutants: ozone and aerosols in the St. Louis urban plume. *Science* **1976**, *194*, 187–189.

(14) Caulton, D. R.; Shepson, P. B.; Santoro, R. L.; Sparks, J. P.; Howarth, R. W.; Ingraffea, A. R.; Cambaliza, M. O. L.; Sweeney, C.; Karion, A.; Davis, K. J.; Stirm, B. H.; Montzka, S. A.; Miller, B. R. Toward a better understanding and quantification of methane emissions from shale gas development. *Proceedings of the National Academy of Sciences* **2014**, *111*, 6237–6242.

(15) Mehrotra, S.; Faloona, I.; Suard, M.; Conley, S.; Fischer, M. L. Airborne Methane Emission Measurements for Selected Oil and Gas Facilities Across California. *Environmental Science & Technology* **2017**, *51*, 12981–12987.

(16) Cambaliza, M. O. L.; Shepson, P. B.; Caulton, D. R.; Stirm, B.; Samarov, D.; Gurney, K. R.; Turnbull, J.; Davis, K. J.; Possolo, A.; Karion, A.; Sweeney, C.; Moser, B.; Hendricks, A.; Lauvaux, T.; Mays, K.; Whetstone, J.; Huang, J.; Razlivanov, I.; Miles, N. L.; Richardson, S. J. Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions. *Atmospheric Chemistry and Physics* **2014**, *14*, 9029–9050.

(17) Lavoie, T. N.; Shepson, P. B.; Cambaliza, M. O. L.; Stirm, B. H.; Karion, A.; Sweeney, C.; Yacovitch, T. I.; Herndon, S. C.; Lan, X.; Lyon, D. Aircraft-Based Measurements of Point Source Methane Emissions in the Barnett Shale Basin. *Environmental Science & Technology* **2015**, *49*, 7904–7913.

(18) Gerbig, C.; Lin, J. C.; Wofsy, S. C.; Daube, B. C.; Andrews, A. E.; Stephens, B. B.; Bakwin, P. S.; Grainger, C. A. Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research: Atmospheres* **2003**, *108*, 4757.

(19) Sheng, J.-X.; Jacob, D. J.; Turner, A. J.; Maasakkers, J. D.; Sulprizio, M. P.;

Bloom, A. A.; Andrews, A. E.; Wunch, D. High-resolution inversion of methane emissions in the Southeast US using SEAC^{4}RS aircraft observations of atmospheric methane: anthropogenic and wetland sources. *Atmospheric Chemistry and Physics* **2018**, *18*, 6483–6491.

(20) Cui, Y. Y.; Brioude, J.; McKeen, S. A.; Angevine, W. M.; Kim, S.-W.; Frost, G. J.; Ahmadov, R.; Peischl, J.; Bousserez, N.; Liu, Z.; Ryerson, T. B.; Wofsy, S. C.; Santoni, G. W.; Kort, E. A.; Fischer, M. L.; Trainer, M. Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin. *Journal of Geophysical Research: Atmospheres* **2015**, *120*, 6698–6711.

(21) Cui, Y. Y.; Brioude, J.; Angevine, W. M.; Peischl, J.; McKeen, S. A.; Kim, S.-W.; Neuman, J. A.; Henze, D. K.; Bousserez, N.; Fischer, M. L.; Jeong, S.; Michelsen, H. A.; Bambha, R. P.; Liu, Z.; Santoni, G. W.; Daube, B. C.; Kort, E. A.; Frost, G. J.; Ryerson, T. B.; Wofsy, S. C.; Trainer, M. Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The San Joaquin Valley. *Journal of Geophysical Research: Atmospheres* **2017**, *122*, 3686–3699.

(22) Brioude, J.; Kim, S.-W.; Angevine, W. M.; Frost, G. J.; Lee, S.-H.; McKeen, S. A.; Trainer, M.; Fehsenfeld, F. C.; Holloway, J. S.; Ryerson, T. B.; Williams, E. J.; Petron, G.; Fast, J. D. Top-down estimate of anthropogenic emission inventories and their interannual variability in Houston using a mesoscale inverse modeling technique. *Journal of Geophysical Research: Atmospheres* **2011**, *116*, D20305.

(23) Brioude, J.; Angevine, W. M.; Ahmadov, R.; Kim, S.-W.; Evan, S.; McKeen, S. A.; Hsie, E.-Y.; Frost, G. J.; Neuman, J. A.; Pollack, I. B.; Peischl, J.; Ryerson, T. B.; Holloway, J.; Brown, S. S.; Nowak, J. B.; Roberts, J. M.; Wofsy, S. C.; Santoni, G. W.; Oda, T.; Trainer, M. Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NO_x and CO_2 and their impacts. *Atmospheric Chemistry and Physics* **2013**, *13*, 3661–3677.

(24) Gourdji, S. M.; Yadav, V.; Karion, A.; Mueller, K. L.; Conley, S.; Ryerson, T.; Nehrkorn, T.; Kort, E. A. Reducing errors in aircraft atmospheric inversion estimates of point-source emissions: the Aliso Canyon natural gas leak as a natural tracer experiment. *Environmental Research Letters* **2018**, *13*, 045003.

(25) Heimburger, A. M. F.; Harvey, R. M.; Shepson, P. B.; Stirm, B. H.; Gore, C.; Turnbull, J.; Cambaliza, M. O. L.; Salmon, O. E.; Kerlo, A.-E. M.; Lavoie, T. N.; Davis, K. J.; Lauvaux, T.; Karion, A.; Sweeney, C.; Brewer, W. A.; Hardesty, R. M.; Gurney, K. R. Assessing the optimized precision of the aircraft mass balance method for measurement of urban greenhouse gas emission rates through averaging. *Elementa* **2017**, *5*, 1–15.

(26) Karion, A.; Sweeney, C.; Kort, E. A.; Shepson, P. B.; Brewer, A.; Cambaliza, M.; Conley, S. A.; Davis, K.; Deng, A.; Hardesty, M.; Herndon, S. C.; Lauvaux, T.; Lavoie, T.; Lyon, D.; Newberger, T.; Petron, G.; Rella, C.; Smith, M.; Wolter, S.; Yacovitch, T. I.; Tans, P. Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. *Environmental Science & Technology* **2015**, *49*, 8124–8131, PMID: 26148550.

(27) Lavoie, T. N.; Shepson, P. B.; Cambaliza, M. O. L.; Stirm, B. H.; Conley, S.; Mehrotra, S.; Faloona, I. C.; Lyon, D. Spatiotemporal Variability of Methane Emissions at Oil and Natural Gas Operations in the Eagle Ford Basin. *Environmental Science & Technology* **2017**, *51*, 8001–8009, PMID: 28678487.

(28) Schwietzke, S.; PÃľtron, G.; Conley, S.; Pickering, C.; Mielke-Maday, I.; Dlugokencky, E. J.; Tans, P. P.; Vaughn, T.; Bell, C.; Zimmerle, D.; Wolter, S.; King, C. W.; White, A. B.; Coleman, T.; Bianco, L.; Schnell, R. C. Improved Mechanistic Understanding of Natural Gas Methane Emissions from Spatially Resolved Aircraft Measurements. *Environmental Science & Technology* **2017**, *51*, 7286–7294, PMID: 28548824.

(29) EIA Detailed Oil and Gas Field Maps in the Appalachian Basin. https://www.eia.gov/maps/map_data/epca2_appal_gross_shp.zip, 2016; Accessed: 2018-09-10.

(30) Enting, I. *Inverse Problems in Atmospheric Constituent Transport*; Cambridge Atmospheric and Space Science Series; Cambridge University Press, 2002.

(31) Tarantola, A.; for Industrial, S.; Mathematics, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Other titles in applied mathematics; Society for Industrial and Applied Mathematics, 2005.

(32) Lopez-Coto, I.; Ghosh, S.; Prasad, K.; Whetstone, J. Tower-based greenhouse gas measurement network design—The National Institute of Standards and Technology North East Corridor Testbed. *Advances in Atmospheric Sciences* **2017**, *34*, 1095–1105.

(33) Stein, A. F.; Draxler, R. R.; Rolph, G. D.; Stunder, B. J. B.; Cohen, M. D.; Ngan, F. NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society* **2015**, *96*, 2059–2077.

(34) Benjamin, S. G.; Weygandt, S. S.; Brown, J. M.; Hu, M.; Alexander, C. R.; Smirnova, T. G.; Olson, J. B.; James, E. P.; Dowell, D. C.; Grell, G. A.; Lin, H.; Peckham, S. E.; Smith, T. L.; Moninger, W. R.; Kenyon, J. S.; Manikin, G. S. A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review* **2016**, *144*, 1669–1694.

(35) Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D.; Duda, M. G.; yu Huang, X.; Wang, W.; Powers, J. G. *A Description of the Advanced Research WRF Version 3*; 2008.

(36) Gurney, K. R.; Mendoza, D. L.; Zhou, Y.; Fischer, M. L.; Miller, C. C.; Geethakumar, S.; de la Rue du Can, S. High Resolution Fossil Fuel Combustion CO2 Emission Fluxes for the United States. *Environmental Science & Technology* **2009**, *43*, 5535–5541, PMID: 19708393.

(37) Oda, T.; Maksyutov, S. A very high-resolution (1 km x 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmospheric Chemistry and Physics* **2011**, *11*, 543–556.

(38) Rayner, P. J.; Raupach, M. R.; Paget, M.; Peylin, P.; Koffi, E. A new global gridded data set of CO2 emissions from fossil fuel combustion: Methodology and evaluation. *Journal of Geophysical Research: Atmospheres* **2010**, *115*, D19306.

(39) Gately, C. K.; Hutyra, L. R. Large Uncertainties in Urban-Scale Carbon Emissions. *Journal of Geophysical Research: Atmospheres* **2017**, *122*, 11,242–11,260.

(40) Gately, C. K.; Hutyra, L. R.; Sue Wing, I. Cities, traffic, and CO2: A multidecadal assessment of trends, drivers, and scaling relationships. *Proceedings of the National Academy of Sciences* **2015**, *112*, 4999–5004.

(41) Maasakkers, J. D.; Jacob, D. J.; Sulprizio, M. P.; Turner, A. J.; Weitz, M.; Wirth, T.; Hight, C.; DeFigueiredo, M.; Desai, M.; Schmeltz, R.; Hockstad, L.; Bloom, A. A.; Bowman, K. W.; Jeong, S.; Fischer, M. L. Gridded National Inventory of U.S. Methane Emissions. *Environmental Science & Technology* **2016**, *50*, 13123–13133, PMID: 27934278.

(42) Janssens-Maenhout, G.; Crippa, M.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Dentener, F.; Bergamaschi, P.; Pagliari, V.; Olivier, J. G. J.; Peters, J. A. H. W.; van Aardenne, J. A.; Monni, S.; Doering, U.; Petrescu, A. M. R. EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012. *Earth System Science Data Discussions* **2017**, *2017*, 1–55.

(43) Crippa, M.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Dentener, F.; van Aardenne, J. A.; Monni, S.; Doering, U.; Olivier, J. G. J.; Pagliari, V.; Janssens-Maenhout, G. Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. *Earth System Science Data* **2018**, *10*, 1987–2013.

(44) EPA National Emissions Inventory. https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data, 2011.

(45) Oda, T.; Maksyutov, S.; Andres, R. J. The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data* **2018**, *10*, 87–107.

(46) Andres, R. J.; Boden, T. A.; Higdon, D. M. Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example. *Atmospheric Chemistry and Physics* **2016**, *16*, 14979–14995.

(47) Engelen, R. J.; Denning, A. S.; Gurney, K. R. On error estimation in atmospheric CO2 inversions. *Journal of Geophysical Research: Atmospheres* **2002**, *107*, ACL 10–1–ACL 10–13.

(48) Desroziers, G.; Berre, L.; Chapnik, B.; Poli, P. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society* **2005**, *131*, 3385–3396.

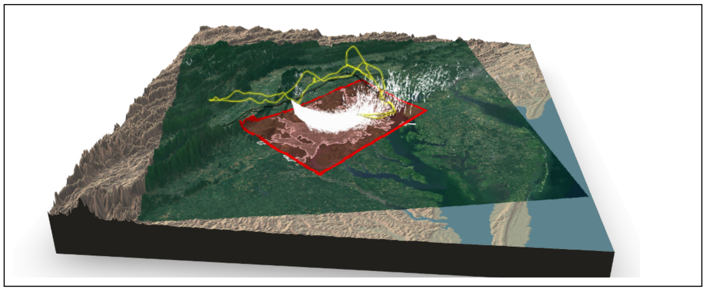
(49) EPA Continuous Emissions Monitoring System Data for 2016. ftp://newftp.epa.gov/DMDnLoad/emissions/hourly/monthly/2016/, 2016; Accessed: 2017-11-03.

(50) Travel Monitoring Analysis System. https://www.fhwa.dot.gov/policyinformation/tables/tmasdata/, 2016; Accessed: 2017-11-03.

(51) Kim, S. Y.; Millet, D. B.; Hu, L.; Mohr, M. J.; Griffis, T. J.; Wen, D.; Lin, J. C.; Miller, S. M.; Longo, M. Constraints on Carbon Monoxide Emissions Based on Tall Tower Measurements in the U.S. Upper Midwest. *Environmental Science & Technology* **2013**, *47*, 8316–8324, PMID: 23844675.

(52) Vaughn, T. L.; Bell, C. S.; Pickering, C. K.; Schwietzke, S.; Heath, G. A.; Pétron, G.; Zimmerle, D. J.; Schnell, R. C.; Nummedal, D. Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a natural gas production region. *Proceedings of the National Academy of Sciences* **2018**, *115*, 11712–11717.

Graphical TOC Entry



Supporting information for "Wintertime CO_2, CH_4 and CO emissions estimation for the Washington DC / Baltimore metropolitan area using an inverse modeling technique"

Israel Lopez-Coto,[*,†] Xinrong Ren,[‡,¶] Olivia E. Salmon,[§,⊥] Anna Karion,[†] Paul B. Shepson,[§,‖] Russell R. Dickerson,[‡] Ariel Stein,[¶] Kuldeep Prasad,[†] and James R. Whetstone[†]

†*National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899*

‡*University of Maryland, 4254 Stadium Dr, College Park, MD 20742*

¶*Air Resources Laboratory, NOAA, 5830 University Research Court, College Park, MD 20740*

§*Purdue University, 610 Purdue Mall, West Lafayette, IN 47907*

‖*Stony Brook University, 100 Nicolls Rd, Stony Brook, NY 11794*

⊥*now at Lake Michigan Air Directors Consortium, 101 S Webster St, Madison, WI 53703*

E-mail: israel.lopezcoto@nist.gov

- Number of pages: 49
- Number of figures: 21
- Number of tables: 3

Contents

1 Methods

1.1 Observations

Two airborne platforms were used to quantify trace gas emissions from the Baltimore-Washington area: Purdue University's Beechcraft† Duchess housing the Airborne Laboratory for Atmospheric Research, or ALAR, (Purdue) and the University of Maryland's Cessna† 402B research aircraft (UMD). Both planes flew simultaneously for 5 days, mostly during the afternoon hours, collecting trace gas mole fraction and meteorological data. Figure 1 shows the flight paths of both aircraft for the flights conducted over the Baltimore-Washington area in February 2016. A typical flight experiment includes transects at different altitudes to capture trace gas enhancement at the downwind side and spirals, en route vertical profiles generally exceeding the PBL, and missed approaches at regional airports to capture vertical gradients.

The equipment on the Purdue aircraft included a global positioning and inertial navigation system (GPS/INS), a Best Air Turbulence (BAT) probe for wind measurements, a cavity ring-down spectroscopy (CRDS) analyzer (Picarro† Model G2301-m) for CH_4, CO_2, and H_2O measurements. Details about the instrumentation on the Purdue aircraft are described elsewhere.[1,2]

The UMD Cessna was equipped with an instrument package to measure gaseous and particulate air pollutants, including a CRDS (Picarro†, Model G2401-m) analyzer to measure CO_2, CH_4, CO, and H_2O. The instrument package has been described in detail elsewhere.[3] Calibrations for CO_2, CH_4 (from both aircraft) and CO (UMD) were conducted both in-flight and on the ground using NOAA/WMO-traceable standards. Observations, originally collected at 0.5 Hz, were averaged at 1 minute resolution and the standard deviation of the averaging period was computed in order to assess the representativity of the mean for each

†Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

particular minute.

In order to ensure well-mixed conditions, the correlation of CO_2 concentration with altitude was computed from the bottom up (with a floor of 150 magl) until the correlation between concentration and altitude was significant, as we expect concentrations in well-mixed conditions not to be correlated with altitude. Then, the highest altitude with no significant correlation (correlation near to zero with p-values > 0.5) was used as the top of the mixed layer. Observations outside this mixed layer were excluded.

To determine the effect of withholding observations from the inversion system, we alternatively used CO_2 and CH_4 observations from both aircraft, the UMD aircraft alone, or the Purdue aircraft alone, as part of the ensemble of inversions. Purdue did not measure CO, so the CO inversions all used the UMD observations alone.

1.2 Bayesian Inversion Framework

We estimate trace gas emissions from measured atmospheric mole fractions using a Bayesian inverse analysis[4,5] as in Lopez-Coto et al.[6]

The measurements model can be written as follows:

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{\varepsilon}_r \tag{1}$$

where $\boldsymbol{y}$ is the observations vector (n x 1, where n is the number of observations), here the tracer mole fractions measured along the track; $\boldsymbol{x}$ is the state vector (m x 1, where m is the total number of pixels in the domain) which we aim to optimize, here the tracer fluxes; $\mathbf{H}$ is the observation operator (n x m) which converts the model state to observations, constructed by using the footprints computed by the transport model, and $\boldsymbol{\varepsilon}_r$ is the uncertainty in the measurements and in the modeling framework (model-data mismatch). Fluxes are assumed to be static in time for a given flight.

Optimum posterior estimates of fluxes are obtained by minimizing the cost function J:[4,5]

$$J(\boldsymbol{x}) = \frac{1}{2}\left[(\boldsymbol{x}-\boldsymbol{x}_b)^T \mathbf{P}_b^{-1}(\boldsymbol{x}-\boldsymbol{x}_b) + (\mathbf{H}\boldsymbol{x}-\boldsymbol{y})^T \mathbf{R}^{-1}(\mathbf{H}\boldsymbol{x}-\boldsymbol{y})\right] \tag{2}$$

where $\boldsymbol{x}_b$ is the first guess or a priori state vector, $\mathbf{P_b}$ the *a priori* error covariance matrix which represents the uncertainties in our *a priori* knowledge about the fluxes and $\mathbf{R}$ the error covariance matrix, which represents the uncertainties in the observation operator $\mathbf{H}$ and the observations $\boldsymbol{y}$, also known as model-data mismatch.

The analytical solution for the posterior state vector, $\boldsymbol{x}_a$, can be written as:

$$\boldsymbol{x}_a = \boldsymbol{x}_b - \mathbf{K}(\mathbf{H}\boldsymbol{x}-\boldsymbol{y}) \tag{3}$$

$$\mathbf{K} = \mathbf{P}_b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_b\mathbf{H}^T+\mathbf{R}\right)^{-1} \tag{4}$$

In the equations above, $\boldsymbol{y}$ usually represents the enhancement over the background, i.e. the mole fraction enhancement at an observation location and time that is attributable to emissions in the region of interest. However, here we split the background ($\boldsymbol{y_{bg}}$) into two terms: the outside contribution from nearby sources ($\boldsymbol{y_{oc}}$) and the long range background ($\boldsymbol{y_{lr}}$), so that the total mole fraction measured by the aircraft, $\boldsymbol{y}_{\mathbf{T}}$, is:

$$\boldsymbol{y}_{\mathbf{T}} = \boldsymbol{y_{ic}} + \boldsymbol{y_{bg}} = \boldsymbol{y_{ic}} + \boldsymbol{y_{oc}} + \boldsymbol{y_{lr}} \tag{5}$$

Therefore, the observations vector $\boldsymbol{y}$ considered in this work contains both the inside contribution ($\boldsymbol{y_{ic}}$) due to the emissions that we aim to estimate plus the outside contribution ($\boldsymbol{y_{oc}}$) from nearby sources.

In this case, the state vector contains additional parameters, similarly to,[7–9] characterizing the outside contribution from nearby sources for each observation ($\boldsymbol{x_{oc}}$) that are the $\boldsymbol{y_{oc}}$ computed as described in Section **1.5** ***Background determination***. Therefore, the observations operator $\mathbf{H}$ (n x (m+n)) is composed of the transport operator, $\mathbf{T}$ (n x m), and the outside contribution operator that is, in fact, the identity matrix, $\mathbf{I}$ (n x n).

$$\mathbf{H} = \begin{bmatrix} \mathbf{T}_{\text{nxm}} & \mathbf{I}_{\text{nxn}} \end{bmatrix} \tag{6}$$

With this formulation, the prior error covariance $\mathbf{P_b}$ gets modified in a similar manner to represent the uncertainty in the emissions and the outside contribution parts of the state vector:

$$\mathbf{P_b} = \begin{bmatrix} \mathbf{E}_{\text{mxm}} & \mathbf{0} \\ \mathbf{0} & \mathbf{O}_{\text{nxn}} \end{bmatrix} \tag{7}$$

where $\mathbf{E}$ is the portion of $\mathbf{P_b}$ associated with the error in the emissions priors, and $\mathbf{O}$ is the error on the prior estimate of the outside contribution.

1.3 Transport Models

The transport model used in this work was the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT[10]). The HYSPLIT model was used in a mode that allows it to emulate the Stochastic Time Inverted Lagrangian Transport model[11] and then compute sensitivity of observations to surface fluxes, or footprints (units: ppm μmol^{-1} m^2 s). Fig. S1 shows the total observations' sensitivity (footprints) for both aircraft only for the data within the well-mixed layer used in the inversion for the five different days (a-e) and the campaign mean (f). The footprints show that the Washington, DC - Baltimore metropolitan area was well-covered during the campaign as well as during each of the individual flights. The sensitivity to nearby outside sources is also apparent, as previously mentioned.

In order to generate an ensemble of transport models and therefore better represent the uncertainties, HYSPLIT was driven with five different meteorological products: the High Resolution Rapid Refresh (HRRR) NOAA operational forecast product[12] provided in the proper format by the NOAA Air Resources Laboratory (ARL) and four configurations of the National Center for Atmospheric Research (NCAR) Weather Research and Forecasting model (WRF[13]).

Four Planetary Boundary Layer (PBL) parameterizations were used in WRF along with two sources of initial and boundary conditions to drive it. The local PBL scheme MYNN[14] and the non-local scheme YSU[15] were used along with the North American Regional Reanalysis product (NARR[16]) provided by the National Center of Environmental Prediction (NCEP) as initial and boundary conditions. On the other hand, the QNSE scheme[17] and BOULAC[18] scheme with the Building Energy Parameterization (BEP[19]) were driven by HRRR.[20] The rest of the parameterizations were kept constant between WRF configurations being: RRTMg for the radiation scheme,[21] Thompson microphysics scheme[22,23] and Noah land surface model.[24]

WRF used a configuration with 3 nested domains (9, 3 and 1 km horizontal resolution) and 60 vertical levels with 34 below 3000 m. The temporal resolution of the output was set to 1 hour for the 9 and 3 km domains and 15 minutes for the 1 km domain, which covers most of the flight tracks. NARR has 32 km horizontal resolution, 30 vertical levels and 3 hours temporal resolution while HRRR has 3 km horizontal resolution, 51 vertical levels and 1 hour temporal resolution. When driven by HRRR, only the 3 and 1 km domains were used in WRF.

HYSPLIT was configured to use Planetary Boundary Layer Heights (PBLH) and Turbulent Kinetic Energy (TKE) from the meteorological models with the exception of YSU, which does not produce TKE due to the non-local nature of this PBL parameterization, and HRRR. In these cases, HYSPLIT used the Kantha-Clayson parametrization to diagnose the turbulence. In addition, an experimental vertical mixing parametrization where the eddy diffusivity for scalars, K_z, exported directly from the underlying WRF model is used in HYSPLIT to compute the vertical velocity variances, was used with WRF-MYNN driven by NARR. Table S1 summarizes the six transport configurations.

The HYSPLIT computation domain for the inversion was set to 100 x 125 grid cells (lat x lon) at 0.03° spatial resolution in order to fully cover the flight tracks and the area of interest. In addition, a secondary domain of 300 x 183 grid cells at the same spatial resolution was used to cover the outer region of influence (Fig. 1). Footprints were computed for both aircraft every minute following the flight tracks.

Here we were after uncertainties that come from the fact that we do not know the right physics and from the initial and boundary conditions. We wanted to have an ensemble of plausible solutions covering that spectrum.

The PBL parametrization drives the vertical mixing of mass, heat and momentum in the planetary boundary layer (PBL)[25] and therefore directly impacts the prediction of temperature, winds and of course planetary boundary layer height (PBLH). The TKE is also different for each scheme, specially between MYNN and QNSE since the latest is based on a substantially different theory.[17] In addition to the 4 PBL schemes, the WRF ensemble consisted of 2 sets of initial and boundary conditions (HRRR and NARR), which also impact winds, temperatures, PBLH and other parameters. Also, the surface layer parametrization was different in the models and one version had the BEP urban canopy model, directly impacting the heat and latent fluxes, which act as the surface boundary condition for the PBL scheme and strongly influence the near surface variables and PBL mean properties.[26] The winds drive the advection in the Lagrangian model but the dispersion is driven by the velocity variances which are parametrized in different ways, also making a big impact.[27] In this work, we used 3 mixing parametrizations in HYSPLIT: KC which depends on the friction velocity and the PBLH, one based on the TKE from WRF and one experimental parametrization that uses directly the eddy diffusivity from WRF (computed by the PBL scheme) to derive the velocity variances. In addition, the footprints, by definition, depend inversely on the PBLH.[11]

We believe that all these choices and options generate enough (plausible) differences on the ensemble of footprints. Nevertheless, to study the similarities among configurations, we applied an agglomerative hierarchical clustering method.[28–30]

The algorithm consists of an iterative process which looks for the smallest dissimilarities between elements, based on the selected distance metric. Once the first 2 elements are clustered, the algorithm computes distances (similarities) between this new cluster and each of the former clusters using the linkage criterion. This process is repeated until all elements are clustered into just one. To cluster N elements, N-1 iterations are required.

Each model is represented by a vector of the wind components for each minute along the flight tracks for the 5 days, $\mathbf{X} = (u_1 \dots u_N, v_1 \dots v_N)$, where N is the total number of minutes of the campaign.

The comparison metric is the Euclidean distance between models and the linkage criterion is the "average" criterion, which is based on the average distance between pairs, i.e., the link between two clusters contains all element pairs, and the distance between clusters equals the average distance between the two elements. "Ward" and "complete" criteria were tested as well with identical results in the resulting grouping.

Figure S2 shows a graphical representation of the clustering results, a dendogram, where the dissimilarities (distance) between models are shown in the y axis. The top hierarchy is split in two branches that are distinguished by the initial conditions (left branch contains only configurations driven by NARR while the right branch contains configurations driven by HRRR or HRRR itself). Thus, the most important choice generating variability in the winds is in fact the initial conditions. The right branch is split further in two, with QS on the left and HR and BL clustered together on the right. This indicates that HR and BL are more similar to each other than they are to QS. This result is reasonable considering that HR and BL use PBL schemes that follow a very similar theory (using different constants and length scale formulations), while QS uses a substantially different theory for parametrizing the turbulence. It is interesting to note that the difference between QS and the cluster BL-HR is also larger than the differences between MY and YU. However, the difference between MY and YU is larger than the difference between HR and BL. In conclusion, BL and HR are the most similar configurations with respect to the wind prediction along the flight tracks.

However, the differences between them are not small, being about 65 % of the most different cluster.

In addition, we analyzed the relative ensemble spread (RES) with respect to enhancements using AC2 emissions inventory, (Fig. S3). The relative ensemble spread of the enhancements had a campaign average of 50 % and was similar for all days, although modest differences exist between days. This might also be due to the fact that different emissions are being sampled each day due to the different flight patterns.

1.4 Emissions Inventories

In addition to the ensemble of transport models, we also used an ensemble of prior fluxes to represent the *a priori* knowledge about the emissions in the area (summarized in Table S2). All the inventories were re-gridded to the inversion domain (0.03°) using a geographical re-projection with bilinear interpolation method.

1.4.1 CO_2

Nine CO_2 emissions inventories were used in the inversion to investigate the resultant variability in the posterior emissions. Four of them (Vulcan, ODIAC, FFDAS and ACES) are existing anthropogenic CO_2 inventories but for a different year; one provides only on-road emissions (DARTE); one is the mean of the previous five (Ensemble); and the rest (Flat and Simple) are constructed here to complement the ensemble of prior fluxes (Fig. S4, Table S3).

Vulcan[31] is a 10x10 km fossil fuel emissions dataset for the United States for the year 2002. The Open-source Data Inventory for Anthropogenic CO_2 (ODIAC[32]) is a dataset with a horizontal resolution of ~1 km based on total emissions estimated by the Carbon Dioxide Information and Analysis Center (CDIAC) at the US Department of Energy's Oak Ridge National Laboratory. Here we use the ODIAC monthly average for February 2015. The Fossil Fuel Data Assimilation System (FFDAS[33]) is a global product with a horizontal grid of 0.1° x 0.1°. The Database of Road Transportation Emissions (DARTE[34]) is a data set that provides a 33-year, 1 km resolution inventory of annual on-road CO_2 emissions for the conterminous United States. Since DARTE only provides road emissions, the National Land Cover Database (NLCD, 2011) was used to compute an urban fraction and assign an emission value of 5 μmol m^{-2} s^{-1} (multiplied by the urban fraction) for urban areas to complement this prior. This value is derived based on the other priors' values for urban areas. The Anthropogenic Carbon Emissions System (ACES) provides estimates of annual and hourly CO_2 emissions from the combustion of fossil fuels for 13 states across the Northeastern United States on a 1 x 1 km spatial grid, for the year 2011.[35] For the ensemble of inversions, we use two different versions of ACES as priors: first, the 2011 annual mean (AC), and second, the mean over the Februaries of 2013 and 2014 (AC2) during the afternoon hours to be consistent with the hours that the flights were conducted.

Using the five inventories described above (Vulcan, ODIAC, DARTE, FFDAS and ACES) we also computed their mean (Ensemble thereafter) and used it as an additional prior in our inverse analysis. Furthermore, we constructed a flat prior that is constant for the whole domain with a value of 1 μmol m^{-2} s^{-1}. This value is arbitrary, as this prior is designed to represent the case of zero prior knowledge about emissions. Lastly, we constructed a simple inventory following the methodology in Lopez-Coto et al.[6] where the land use emissions are considered to be the urban fraction multiplied by 5 μmol m^{-2} s^{-1}, the road emissions to be 2 μmol m^{-2} s^{-1}, and the point sources emissions from the EPA GHGRP. The value assigned to on-road emissions is low due to the large area that one road pixel represents at our resolution (~ 9 km^2). In addition, this value is close to the mean value across the inversion domain for the road emissions provided by DARTE (1.9 μmol m^{-2} s^{-1}). All emissions priors were constant in time. Fig. S4 shows the nine CO_2 prior emissions used in the inversions.

1.4.2 CH_4

Methane prior emissions were represented using the EPA gridded inventory for 2012,[36] EDGAR v4.3.2[37] for 2012, the mean of the previous two, and a flat prior. Both the EPA and

EDGAR inventories are provided at 0.1° × 0.1° and were re-gridded to the 0.03° resolution of this study. The flat prior was chosen to be 1 nmol m^{-2} s^{-1} (Fig. S5, Table S3).

1.4.3 CO

For CO we use EDGAR v4.3.2[38] at 0.1°, the National Emissions Inventory (NEI) for 2011 at 4 km resolution from EPA,[39] the annual mean ACES inventory (AC as in the CO_2 case) scaled using the mean observed ΔCO:ΔCO_2 ratio (6.18 ppb/ppm) and, again, a flat prior (Fig. S6, Table S3).

1.5 Background Determination

Properly accounting for the background is critical for the inversion as the flux correction is based on the observed enhancements above the background value. The impact of upwind sources can be important especially in areas such as the one under study here, where multiple urban areas and oil and gas fields exist around the area (Figure 1).

Measurements along an upwind flight transect often do not properly represent the background in the downwind transect because of differences in timing of both transects plus the differences induced by the transport of air masses itself, such as flow convergence or divergence and differences in the mixing layer height.[3,40,41]

Here we choose to optimize the contribution to the background of sources nearby but outside our domain, $\boldsymbol{y_{oc}}$ (Eq. 5). First, we estimate this contribution as a first guess using a Lagrangian approach by convolving footprints from a reduced set of our ensemble of transport models and with prior fluxes. We extend the domain to the full extent shown in Fig. 1, to account for the contribution of large nearby sources, including the cities of Philadelphia, New York and Pittsburgh as well as the gas operations in the Marcellus shale. The full background is then represented as the ensemble mean of the contribution from outside of the domain of interest ($\boldsymbol{y_{oc}}$, time-varying along the track) plus the long-range background ($\boldsymbol{y_{lr}}$, constant for a given flight) (Eq. 5). $\boldsymbol{y_{lr}}$ is defined here as a reference value measured along the track, the 5th percentile, minus the mean contribution from inside and outside of the inversion domain for all the locations measuring below the specified reference value (Eq. 8).

$$y_{\mathrm{lr}}{=}\mathrm{p_{5th}}\,(\mathrm{obs}) - \frac{1}{N}\sum_{\mathrm{obs}<\mathrm{p_{5th}}(\mathrm{obs})}^{N}(y_{\mathrm{ic}}{+}y_{\mathrm{oc}}) \tag{8}$$

Selecting the 5th percentile as reference value shields the background from abnormally low values that may occur due to non-representative situations. On the other hand, the specific inside contribution of the reference value along the track might be misrepresented due to transport model and emissions errors. To alleviate this situation, we consider that contribution to be the mean value of all the measurements below the reference value, as indicated in Eq. 8. This methodology yields to a time varying background and the associated uncertainties, as described below.

The uncertainty in the background ($\boldsymbol{\sigma_{bg}}$) is composed of 3 terms: 1) the uncertainty in the outside contribution ($\boldsymbol{\sigma_{oc}}$) due to the transport models and prior fluxes, 2) the uncertainty in the inside contribution (σ_{ic}) due to the ensemble of transport models and prior fluxes and 3) the uncertainty in the determination of the inside and outside contribution due to the potential mis-location of the reference value picked along the track (σ_{mis}). All the uncertainties are computed as the standard deviation of the respective set of data used in the calculation.

Because the outside contribution determined in this way depends on the priors used in the computation, it might underestimate or overestimate the values if the ensemble of priors do. To address this problem, we optimize the outside contribution of the background along with the fluxes in the same inversion as described above, Eq. 5 to 7.

In addition, we also performed a sensitivity test (separately from the main ensemble of inversions) to specifically determine the impact of 1) not optimizing the outside contribution, 2) scaling the outside Contribution, and 3) using a less sophisticated approach of selecting a single constant value along the track as background defined by the 1st, 5th or 10th percentile. For the scaled background case, a single scaling factor for each flight was applied to the background time series. This scaling factor was the ratio of posterior to prior emissions for the inversion case where the background was not optimized or scaled (case 1 in the above text).

1.6 Error Covariances

1.6.1 Prior Flux Error Covariance

The prior flux error covariance, **E** in Eq. 7, represents the uncertainties in the prior estimation of the fluxes. It is commonly assumed to follow an exponential model where the correlation between two points decays as the distance between them increases.[6,42–44]

$$E_{ij} = \sigma_i \sigma_j e^{-d_{ij}/L} \tag{9}$$

where σ_i represents the uncertainty for the pixel i, d_{ij} represents the distance between the pixels i and j and L is the correlation length of the spatial field.

A wide range of correlation lengths is found in the literature from less than 10 km to hundreds or thousands of kilometers.[8,42,44] Typically, small values of the correlation length are associated with high-resolution studies conducted in small domains, as for Indianapolis,[44] while large correlation length values are seen in low-resolution inversions in regional to global domains.[42] In this work, the correlation length was assumed to be 10 km, consistent with Lopez-Coto et al.,[6] where the authors found this value to be appropriate for studies at urban scales.

Although bottom-up CO_2 emissions estimates are made on global and national scales with small uncertainties, considerable errors are introduced when the emissions are disaggregated due to the usage of proxies to spatially distribute emissions.[45] Reported errors at grid cell levels range from 4% to more than 190%, averaging about 120%.[46] These errors depend on the inventory disaggregation methodology as well as on the resolution that the error evaluation is performed. For CH_4 and CO it is likely that the errors at grid cell levels are even larger than for CO_2 because of the less well-known characteristics of these species' sources.

Given the aforementioned reported uncertainties at grid cell levels, here we use a value of 100% of the grid cell emissions as uncertainty for all the prior inventories and gases with the exception of FFDAS and the ensemble cases for CO_2. FFDAS provides uncertainties at grid cell level that are very small as compared to the other uncertainty estimates. Because we use the FFDAS annual mean for 2010 to represent a few days in February 2016, the FFDAS provided uncertainties probably do not represent the real errors in our application; therefore we multiplied the provided annual uncertainty by $12^{1/2}$ to try to get a monthly uncertainty estimate, assuming the annual uncertainty is provided as the uncertainty of the annual mean. The uncertainties still remained low compared to the uncertainty estimates for the rest of priors and the impact on the inversion will reflect that. For the ensemble mean prior, we used the standard deviation of the ensemble at each pixel to represent the uncertainties.

For the flat prior cases, we assigned uncertainty values of 10 μmol m^{-2} s^{-1} for CO_2, 30 nmol m^{-2} s^{-1} for CH_4 and 50 nmol m^{-2} s^{-1} for CO. This choice was based on the 90th percentile of the ensemble of prior emissions within the accounting box for CO_2 (8.9 μmol m^{-2} s^{-1}) and CH_4 (32.5 nmol m^{-2} s^{-1}). For CO we used the CO_2 value scaled by the ΔCO:ΔCO_2 ratio.

Because the inventories used here represent only anthropogenic emissions, pixels with low or zero fossil fuel emissions will have a very low uncertainty value making it difficult for the inversion to correct those areas, for example in cases where there may be non-reported emissions such as fugitive emissions or even wintertime biogenic respiration. To address this, we set a floor in the prior uncertainties of 1 μmol m^{-2} s^{-1} for CO_2, 3 nmol m^{-2} s^{-1} for CH_4 and 5 nmol m^{-2} s^{-1} for CO.

1.6.2 Outside Contribution (background) Prior Error Covariance

We consider a double exponential model, in space and time, to represent the error covariance of the outside contribution along the track (**O** in Eq. 7). The diagonal is populated with the uncertainty of the initial guess outside contribution ($\boldsymbol{\sigma_{oc}}$) based on the variance from the different transport models and prior fluxes.

Because the only constraint imposed on the outside contribution during the inversion comes through the covariance, we impose very large correlation length (L=10^4 km) and correlation time (τ = 8760 h). These choices are based on the assumption that the error structure in the outside contribution is similar across large scales in space and time, meaning that if an underestimation/overestimation exists in a region, would likely occur in the nearby areas even if they are very far apart due to the nature of the construction.

Making the correlation equal to zero would allow each individual point to be corrected independently leading to a general over-fitting. Correlations equal to one would force the entire time series for one flight to be scaled up or down together, not allowing for any additional correction in time and space of this background. The selected correlation model allows the inversion to coherently adjust the time series while retaining some flexibility to adjust each point independently based on the specific errors assigned along the diagonal.

1.6.3 Model Error

The model-data mismatch error covariance (**R**) was assumed to have three independent contributions: 1) uncertainty in the observations ($\mathbf{R_{obs}}$), 2) uncertainty in the long range background mole fraction ($\mathbf{R_{lrbg}}$) and 3) uncertainty in the transport model representation ($\mathbf{R_{transport}}$). The uncertainties in the observations are assigned as the measurement uncertainties (0.2 ppm for CO_2, 2 ppb for CH_4 and 2 ppb for CO, obtained from the calibrations and comparisons between measurements from the two aircraft) and the representativity of the assigned mean to the whole averaging period (one minute in our case). This contribution is not correlated and thus the covariance was considered diagonal, where the diagonal was populated with the maximum of the measurement variance and the variance of the averaging period (1 minute). The long-range background determination also introduces uncertainty into the system. This contribution was also assumed to be uncorrelated and the covariance diagonal populated with the sum of the variances due to the inside contribution and the mis-location errors ($\sigma_{ic}^2 + \sigma_{mis}^2$). Lastly, the transport model uncertainty is complex with several previously published methods for its determination. Here we tested two methods, both based on the ensemble of transport models. First, we tested a diagonal covariance populated with the inter-model variance simulated using the same surface fluxes (the prior emissions in each inversion case) in all the transport models similar to[47] and.[48] As stated in,[47] this estimate can be too large for some models and too small for other models, thus, in order to better represent the fidelity of each model and for each observation, we weighted the inter-model standard deviation ($\boldsymbol{\sigma_e}$) with the relative error ($\boldsymbol{\epsilon}$) computed by using the wind measurements from the aircraft as follows:

$$\boldsymbol{\sigma^2} = \boldsymbol{\sigma_e^2 \epsilon^2} = \boldsymbol{\sigma_e^2}(\boldsymbol{\epsilon_{ws}^2} + \boldsymbol{\epsilon_{wd}^2}) \tag{10}$$

where $\boldsymbol{\epsilon_{ws}}$ is the relative error for wind speed and $\boldsymbol{\epsilon_{ws}}$ is the normalized absolute error for the wind direction. Due to the circular nature of the wind direction, the absolute difference is kept between 0 and π by measuring the absolute differences larger than π in the opposite direction ($2\pi - \Delta$). Then we normalized the error to the maximum range, π.

This definition of the transport model error covariance assumes there are no correlations in space and time which is unlikely to be true. Therefore, for the second method, which was used in the main ensemble of inversions, we computed the correlations between the different transport models and included them into the covariance as follows:

$$\mathbf{R}_{transport} = \boldsymbol{\sigma} \otimes \boldsymbol{\sigma} \cdot cor(\mathbf{T}) \tag{11}$$

where $\boldsymbol{\sigma}$ is the weighted inter-model standard deviation computed as in the previous case,

$\otimes$ is the outer product, and $\mathbf{T}$ is a matrix constructed with the simulated observations using the same surface fluxes (the prior emissions in each inversion case) in all transport models.

2 Figures

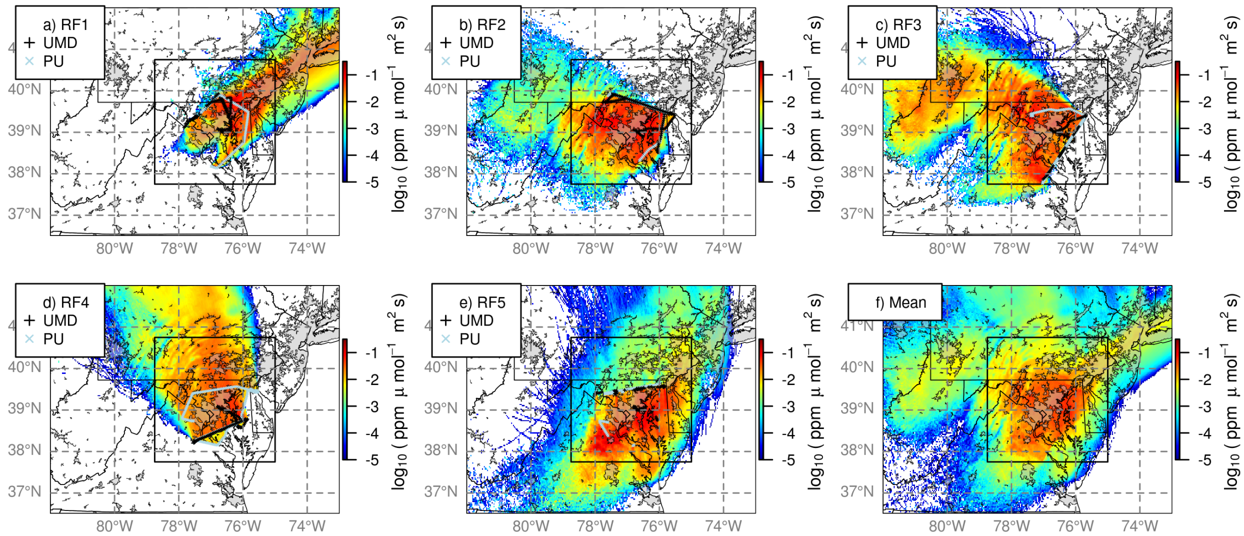


Figure S1: Total observations' sensitivity for both aircraft only for the data within the well-mixed layer used in the inversion for a) 02/08/2016 (RF1), b) 02/12/2016 (RF2), c) 02/17/2018 (RF3), d) 02/18/2016 (RF4), e) 02/19/2016 (RF5) and f) the campaign average.

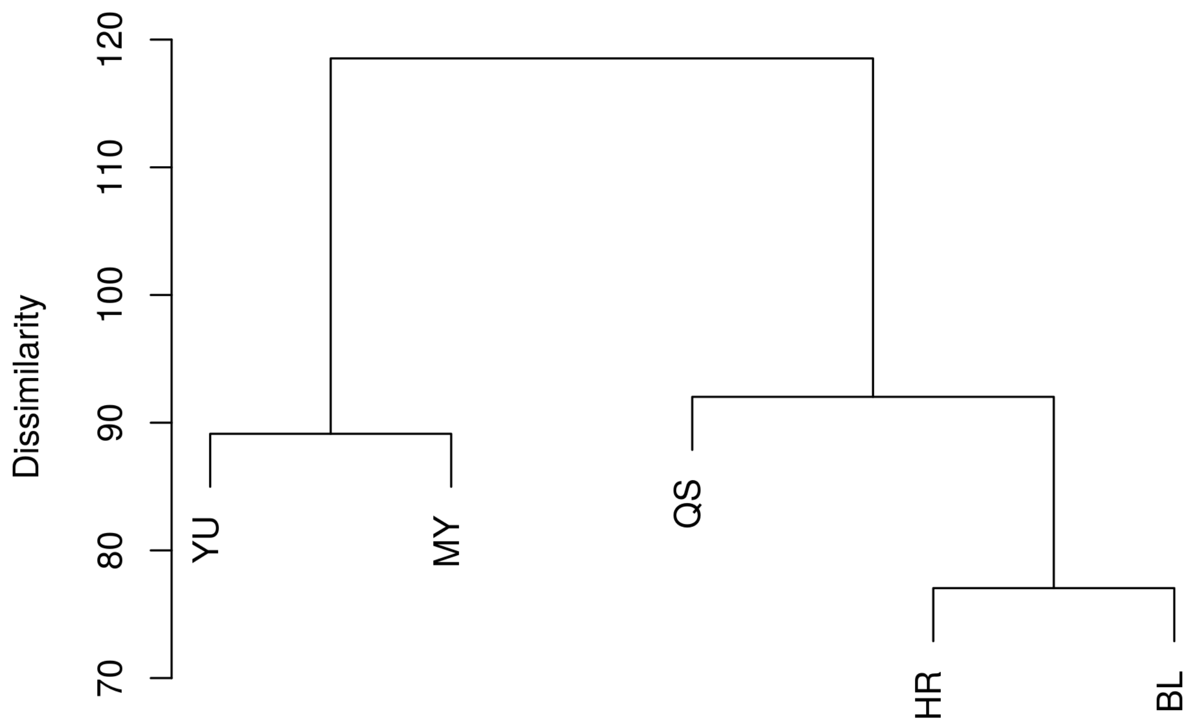


Figure S2: Dendogram computed using agglomerative hierarchical clustering with euclidean distance as similarity metric and the "average" method as the linkage criterion.

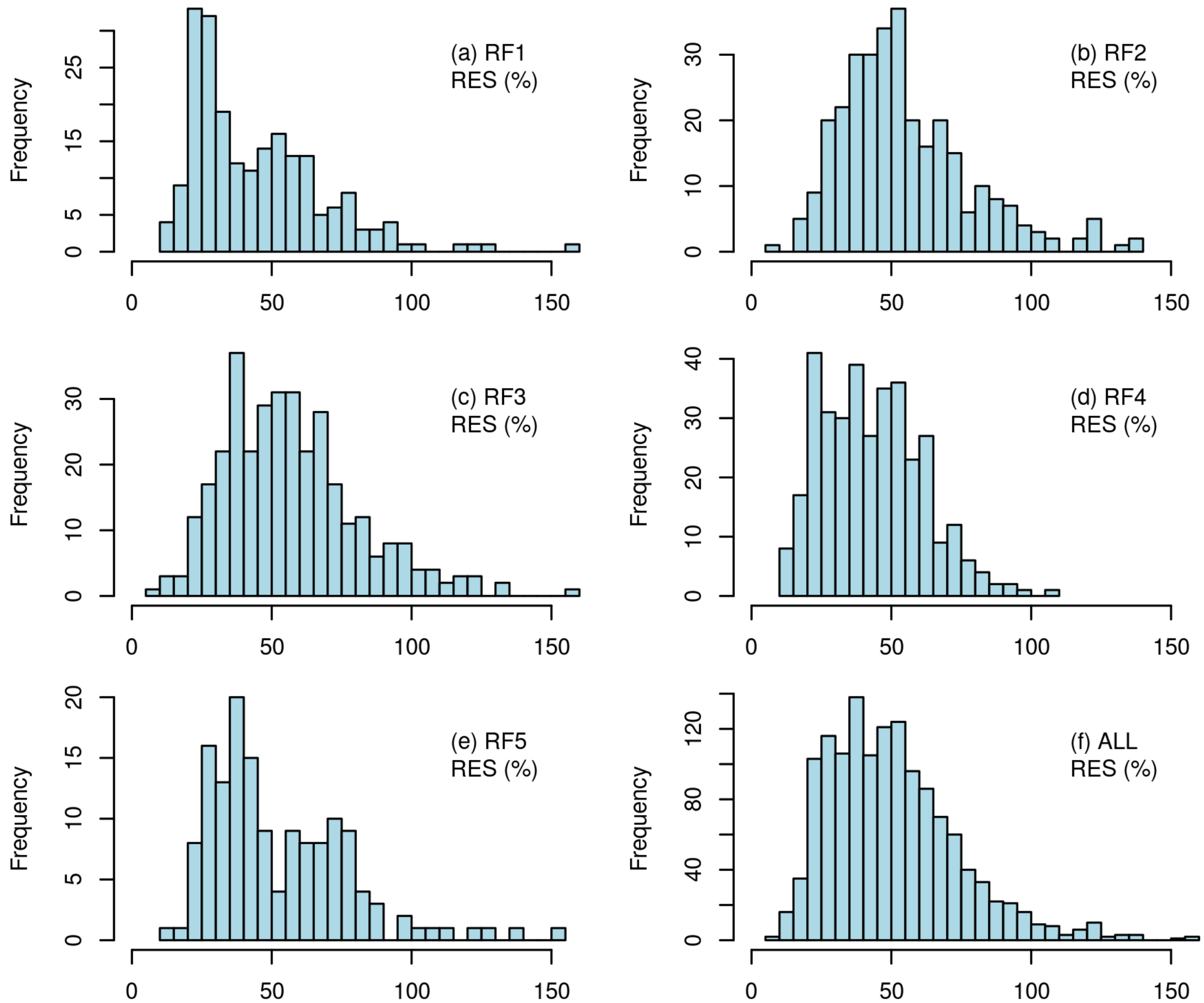


Figure S3: Histograms of the Relative Ensemble Spread (RES, %) for the different days (a-e) and for the entire campaign (f) using AC2 emissions inventory.

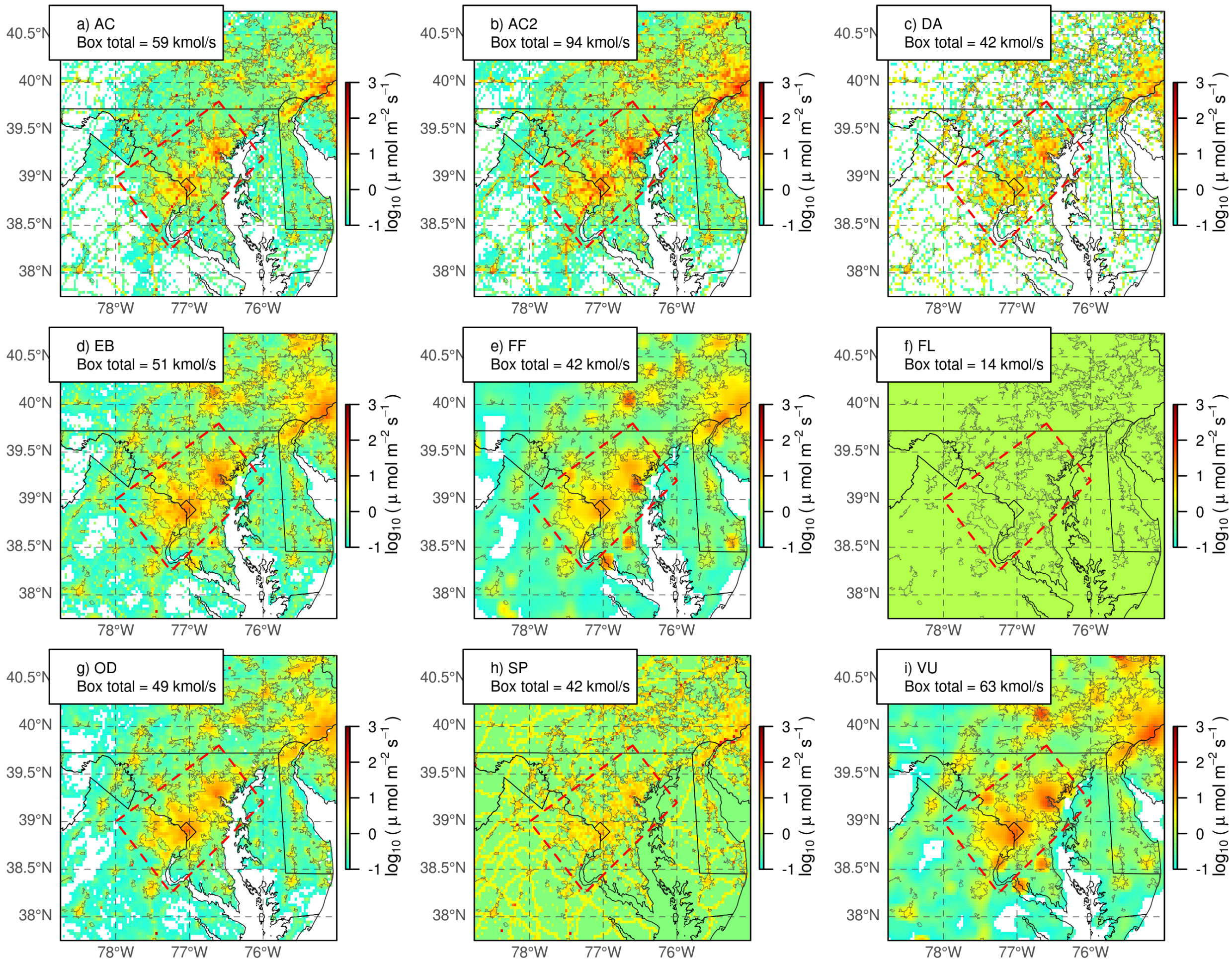


Figure S4: Prior CO_2 emission rate spatial distribution (a) AC is ACES inventory annual mean, (b) AC2 is the mean for February between 12 - 19 EST, (c) DA is the DARTE inventory, (d) EB is the ensemble mean inventory, (e) FF is FFDAS inventory, (f) FL is the Flat inventory, (g) OD is ODIAC, (h) SP is the simple inventory and (i) VU is VULCAN.

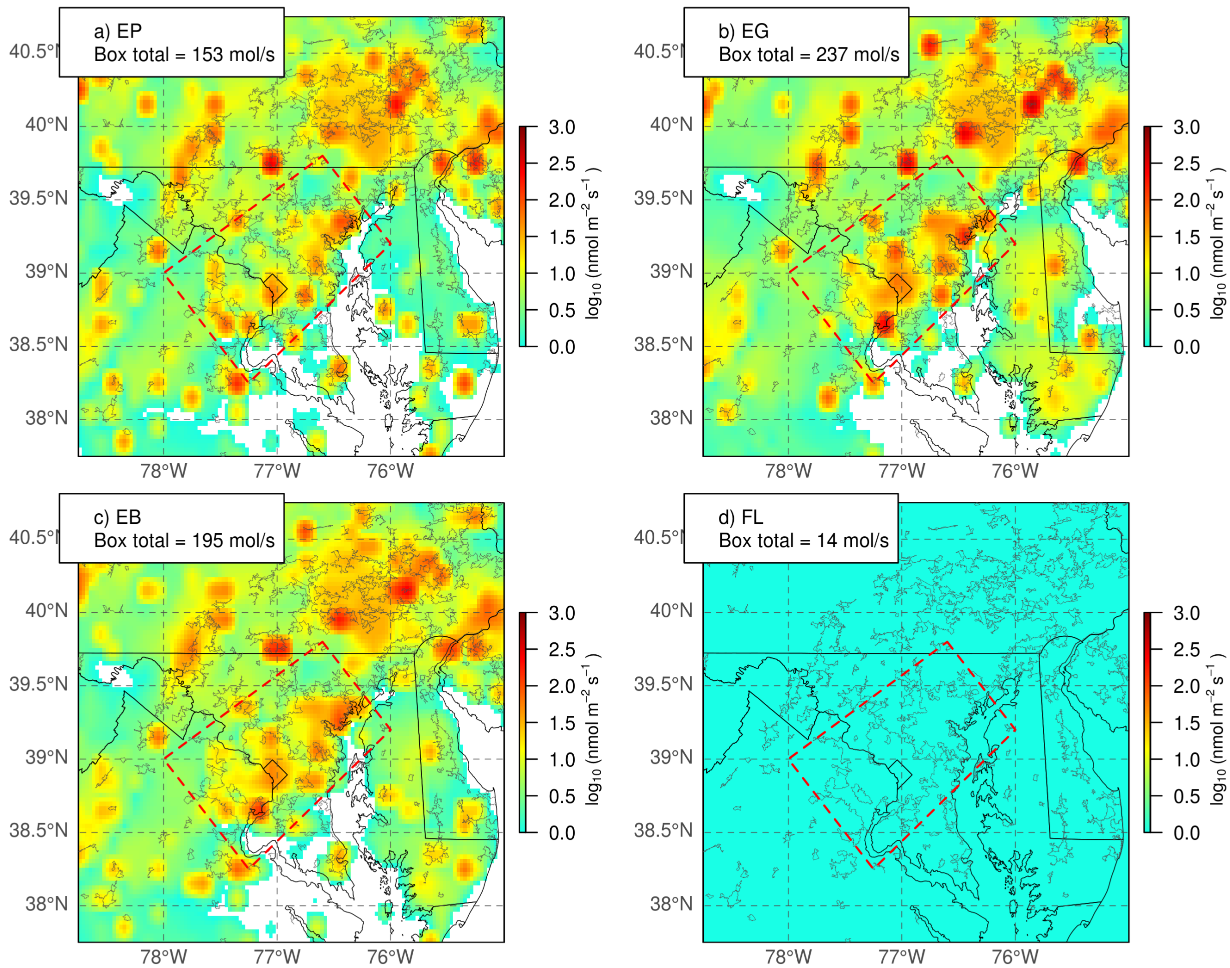


Figure S5: CH_4 prior emissions (a) EP is EPA inventory, (b) EG is EDGAR (v4.3.2), (c) EB is the ensemble mean and (d) FL is the flat inventory

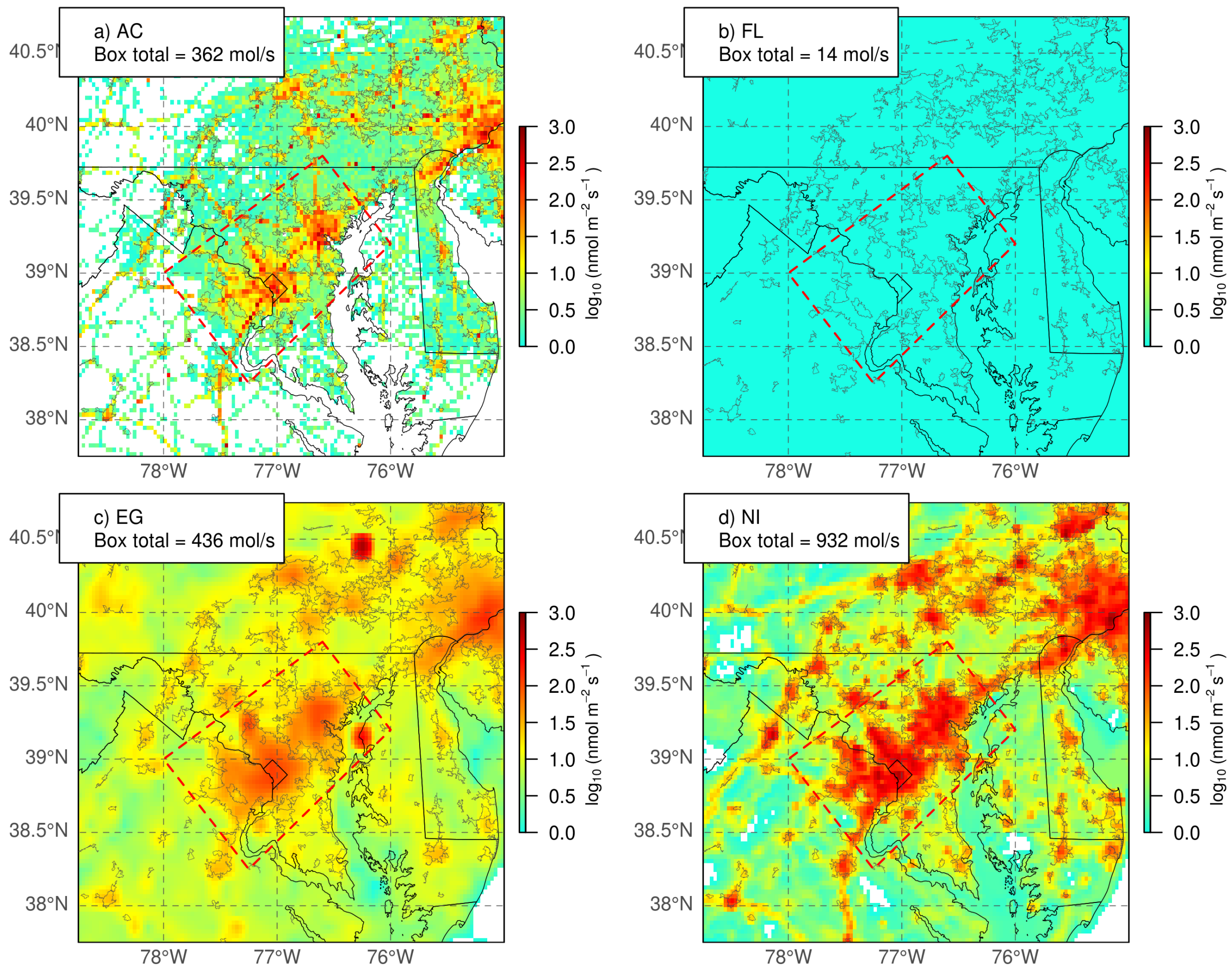


Figure S6: CO prior emissions (a) AC is ACES inventory scaled using the observed $\Delta CO:\Delta CO_2$ ratio, (b) FL is Flat inventory, (c) EG is EDGAR (v4.3.2) and (d) NI is NEI-2011 inventory

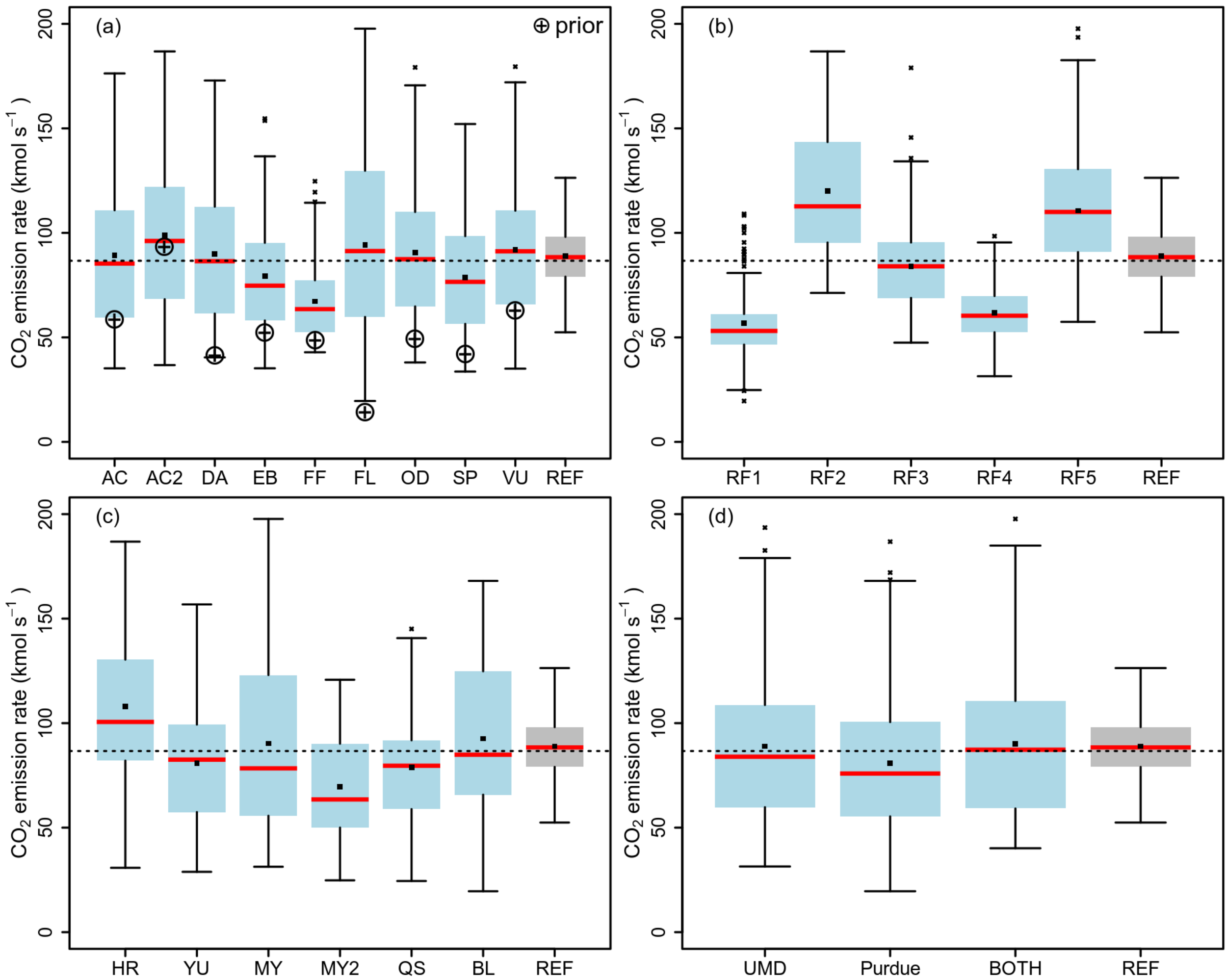


Figure S7: Boxplots of the total estimated CO_2 emission rate within the accounting box compared to the values provided by ACES, scaled to totals of 2016, for February between 12 - 19 EST (referred as REF in the four panels) grouped by: (a) the different inventories used as priors where AC is ACES inventory annual mean, AC2 is the mean for February between 12 - 19 EST, DA is the DARTE inventory, EB is the ensemble mean inventory, FF is FFDAS inventory, FL is the Flat inventory, OD is ODIAC, SP is the simple inventory and VU is VULCAN; (b) the different research flights; (c) the different transport model configurations where HR is HRRR, YU is YSU, MY is MYNN, MY2 is MYNN with HYSPLIT using the WRF eddy diffusivities to compute the mixing, QS is QNSE and BL is BouLac; (d) the observation dataset choice using observations from only the UMD Cessna, Purdue Duchess, or both. Blue bars indicate the 25th and 75th quantiles, whiskers the range, x's the outliers (1.5 times the IQR), red line the median, square markers the mean and the dotted line the posterior mean. The circled pluses in panel (a) represent each prior's total emissions.

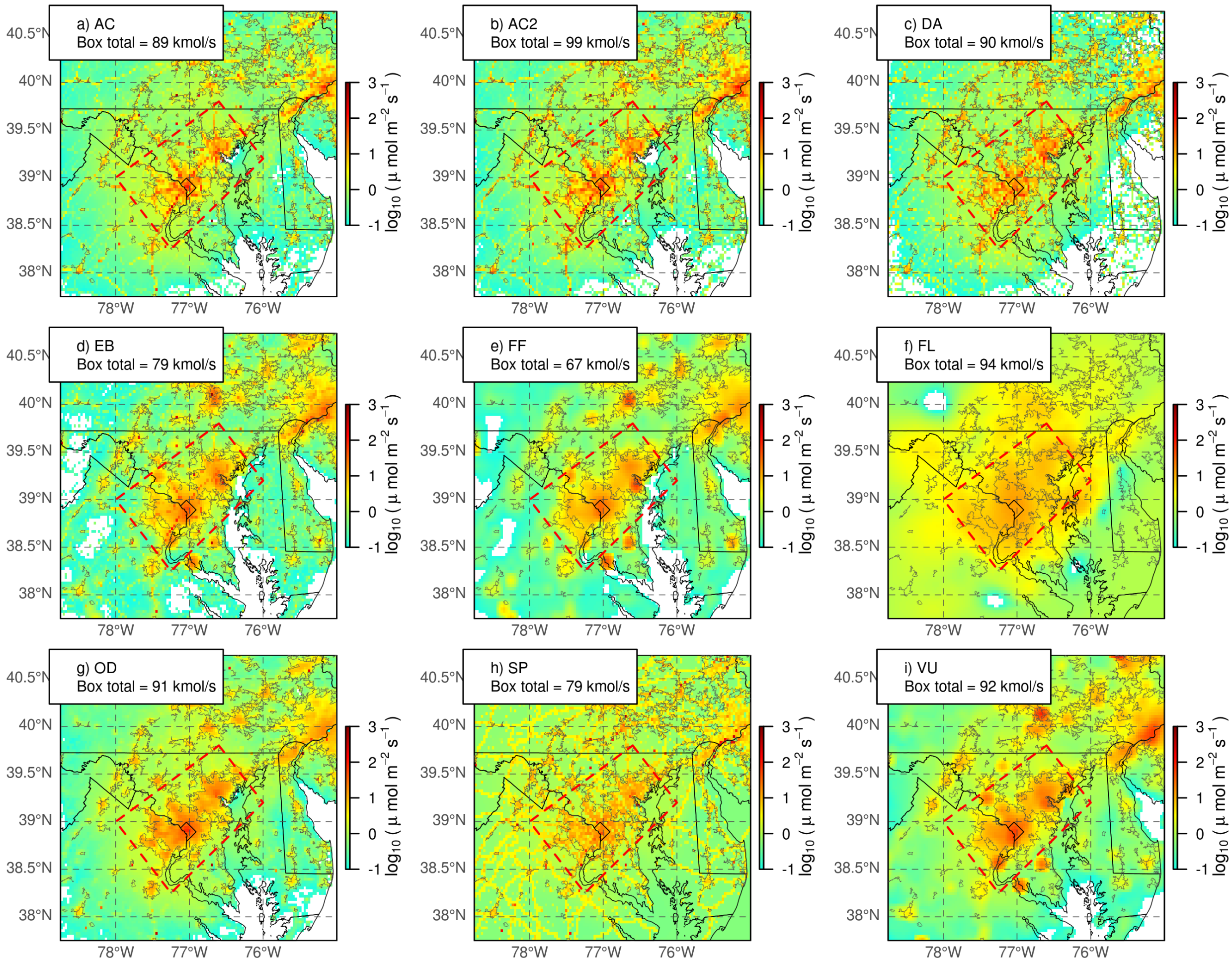


Figure S8: Mean estimated CO_2 emission rate spatial distribution for all days and transport models using the different priors (a) AC is ACES inventory annual mean, (b) AC2 is the mean for February between 12 - 19 EST, (c) DA is the DARTE inventory, (d) EB is the ensemble mean inventory, (e) FF is FFDAS inventory, (f) FL is the Flat inventory, (g) OD is ODIAC, (h) SP is the simple inventory and (i) VU is VULCAN.

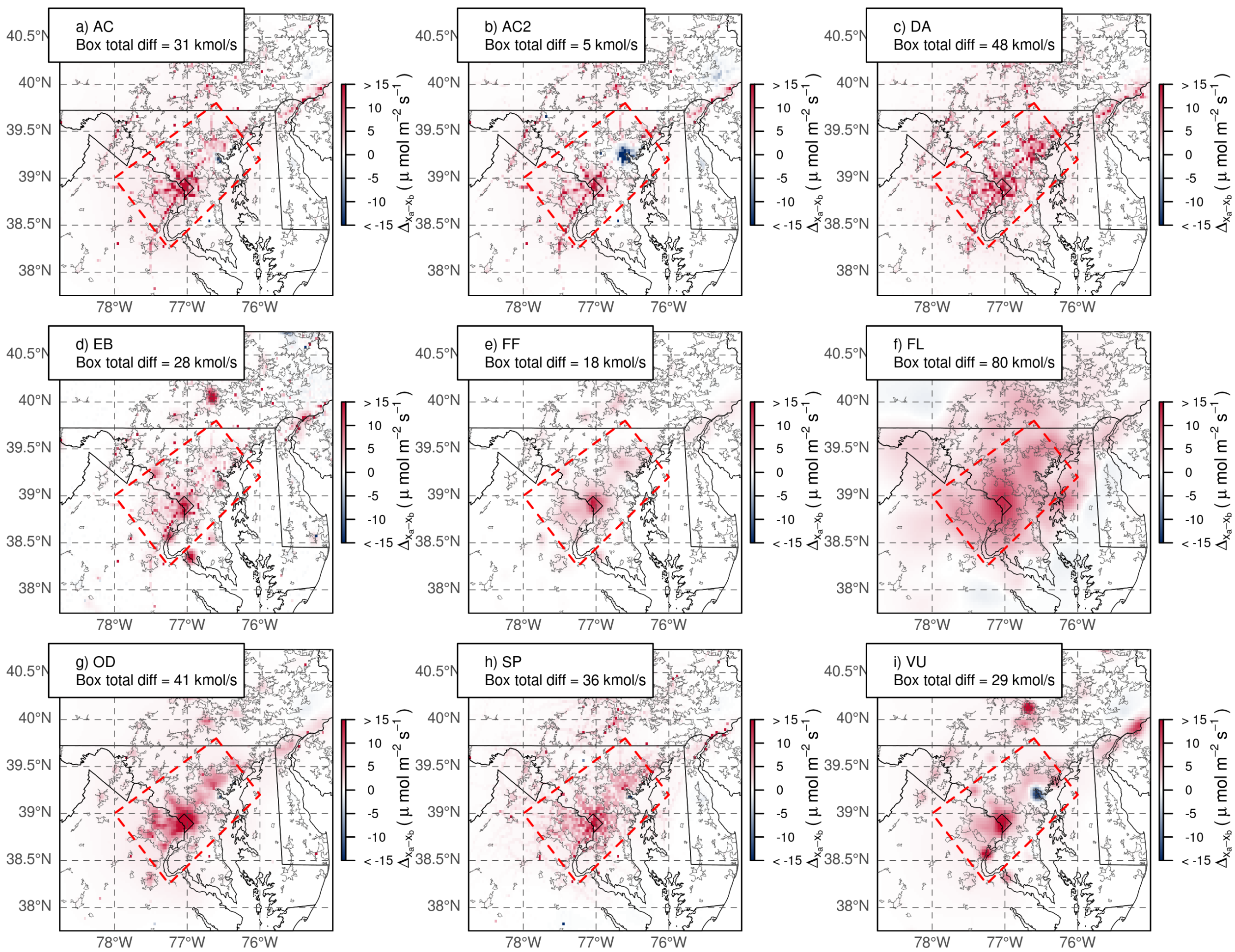


Figure S9: Spatial distribution of differences between the mean estimated CO_2 emission rate and the prior emissions for all days and transport models using the different priors: (a) AC is ACES inventory annual mean, (b) AC2 is the mean for February between 12 - 19 EST, (c) DA is the DARTE inventory, (d) EB is the ensemble mean inventory, (e) FF is FFDAS inventory, (f) FL is the Flat inventory, (g) OD is ODIAC, (h) SP is the simple inventory and (i) VU is VULCAN (Table S2). The legend also indicates the total difference inside the accounting box (dashed red).

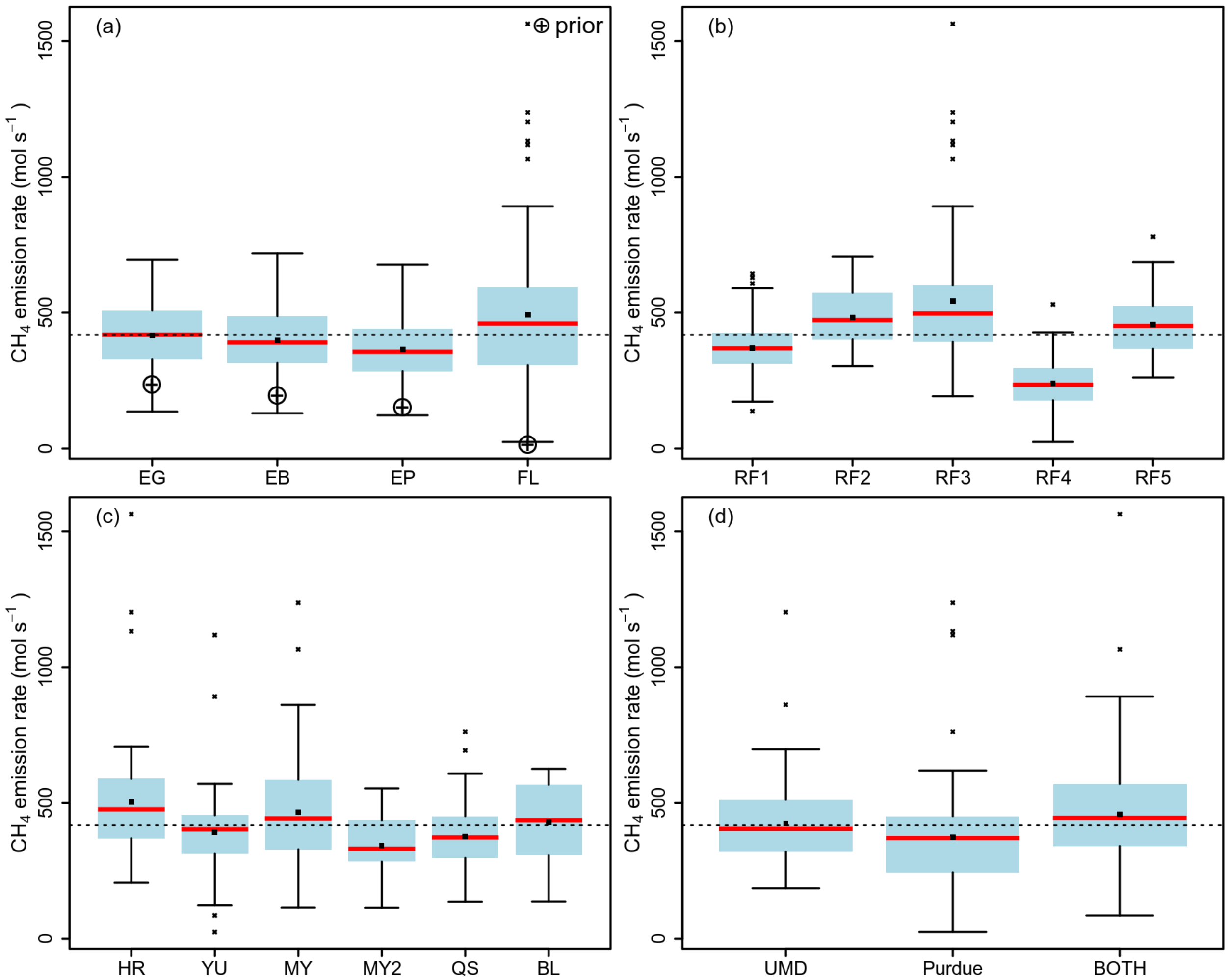


Figure S10: Total estimated CH_4 emission rate within the accounting box grouped by: (a) the different inventories used as priors where EG is EDGAR, EP is EPA, EB is the ensemble and FL is the Flat inventory; (b) the different days; (c) the different transport model configurations (as in Fig. S7); (d) the observation dataset choice. Markers as in Fig. S7

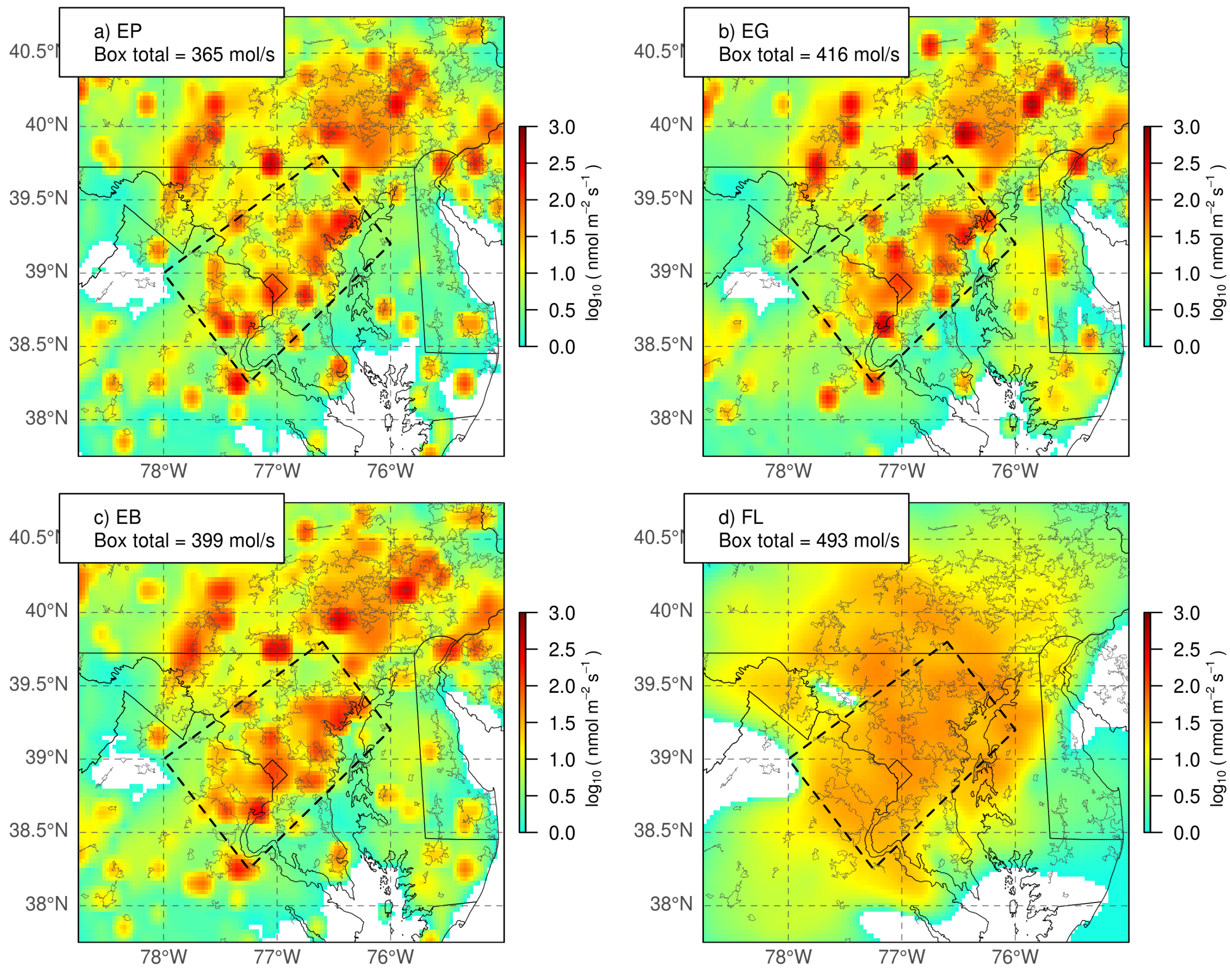


Figure S11: Mean estimated CH_4 emission rate spatial distribution for all days and transport models using the different priors (a) EP is EPA inventory, (b) EG is EDGAR, (c) EB is the ensemble mean inventory and (d) FL is the Flat inventory

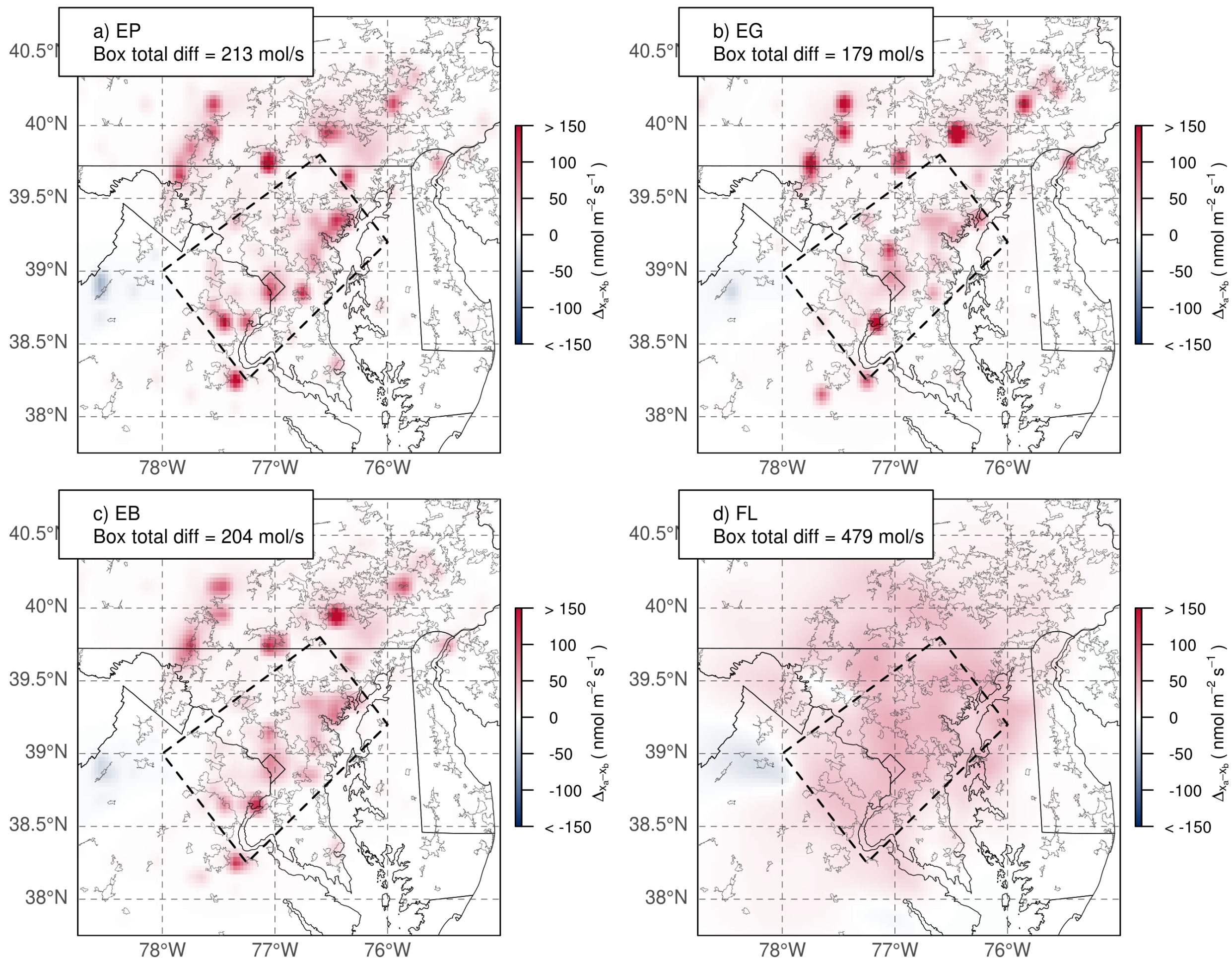


Figure S12: Spatial distribution of differences between the mean estimated CH_4 emission rate and the prior emissions for all days and transport models using the different priors (a) EP is EPA inventory, (b) EG is EDGAR, (c) EB is the ensemble mean inventory and (d) FL is the Flat inventory

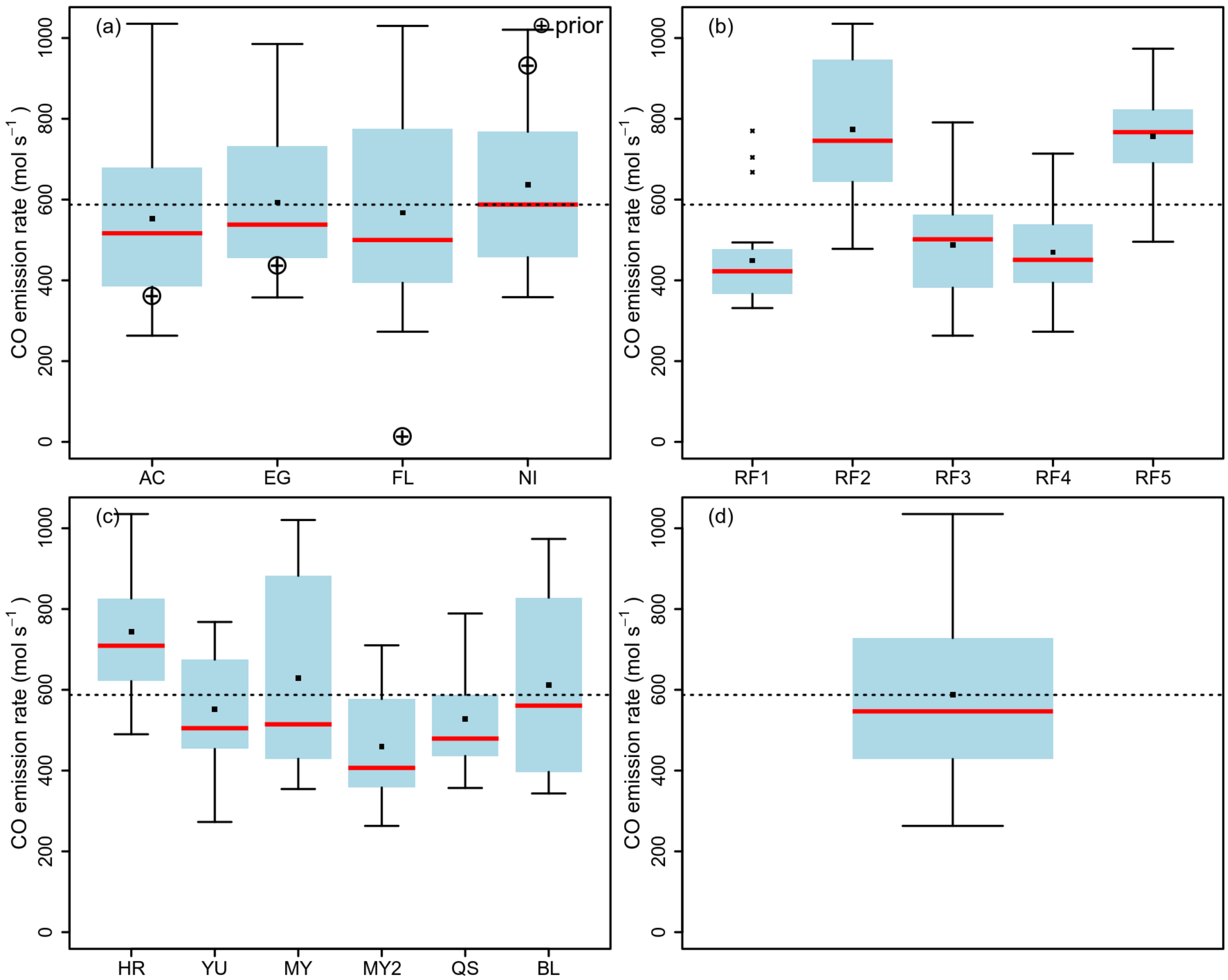


Figure S13: Total estimated CO emission rate within the accounting box grouped by: (a) the different inventories used as priors where AC is ACES inventory annual mean scaled using the mean observed $\Delta CO:\Delta CO_2$ ratio, EG is EDGAR inventory, FL is the Flat inventory and NI is the NEI inventory; (b) the different days; (c) the different transport model configurations (as in Fig. S7); (d) the observation dataset selection using only UMD plane because no CO measurements were made with the Purdue plane. Markers as in Fig. S7

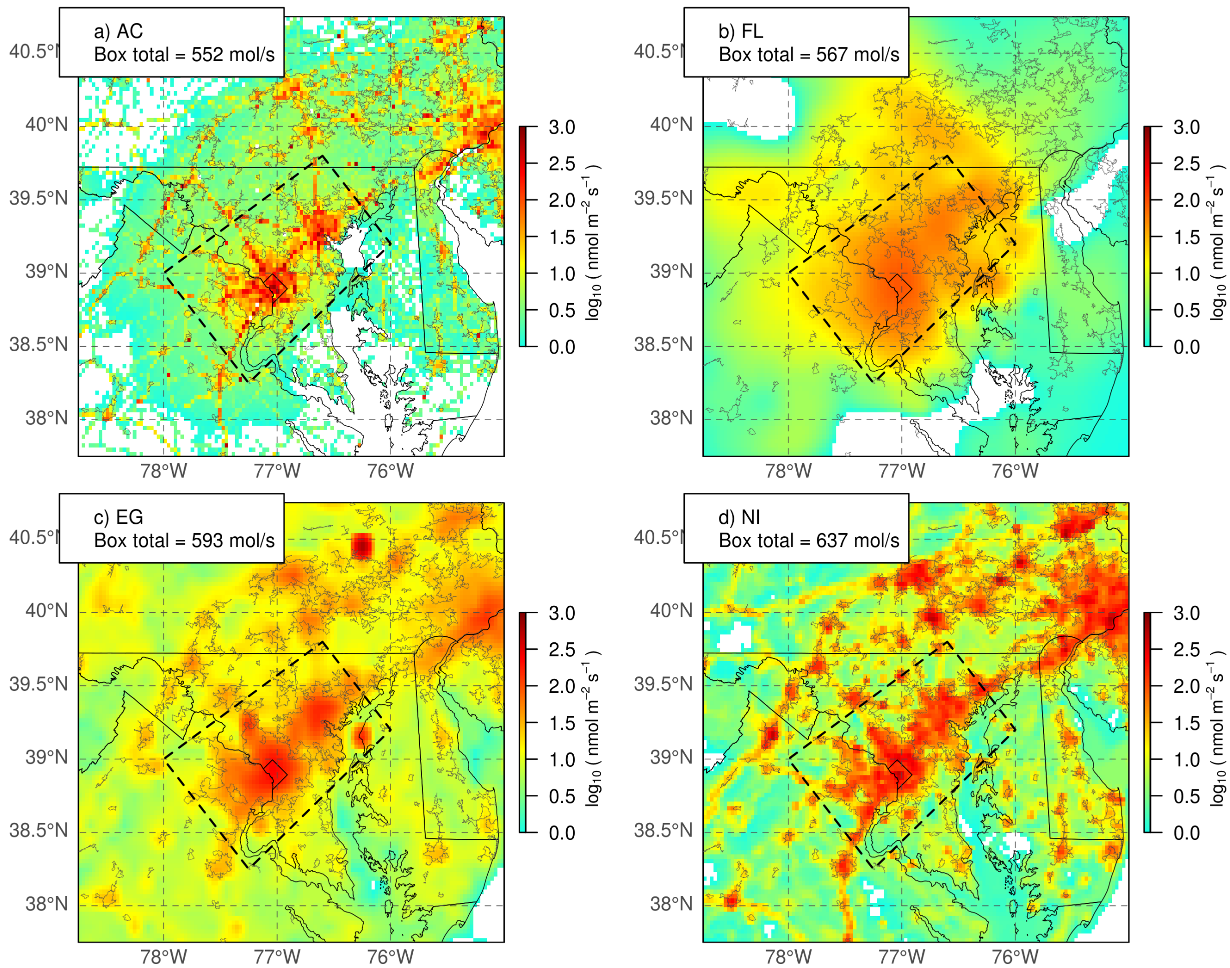


Figure S14: Mean estimated CO emission rate spatial distribution for all days and transport models using the different priors (a) AC is ACES inventory annual mean scaled using the mean observed ΔCO:ΔCO_2 ratio, (b) FL is the Flat inventory, (c) EG is EDGAR inventory and (d) NI is the NEI inventory

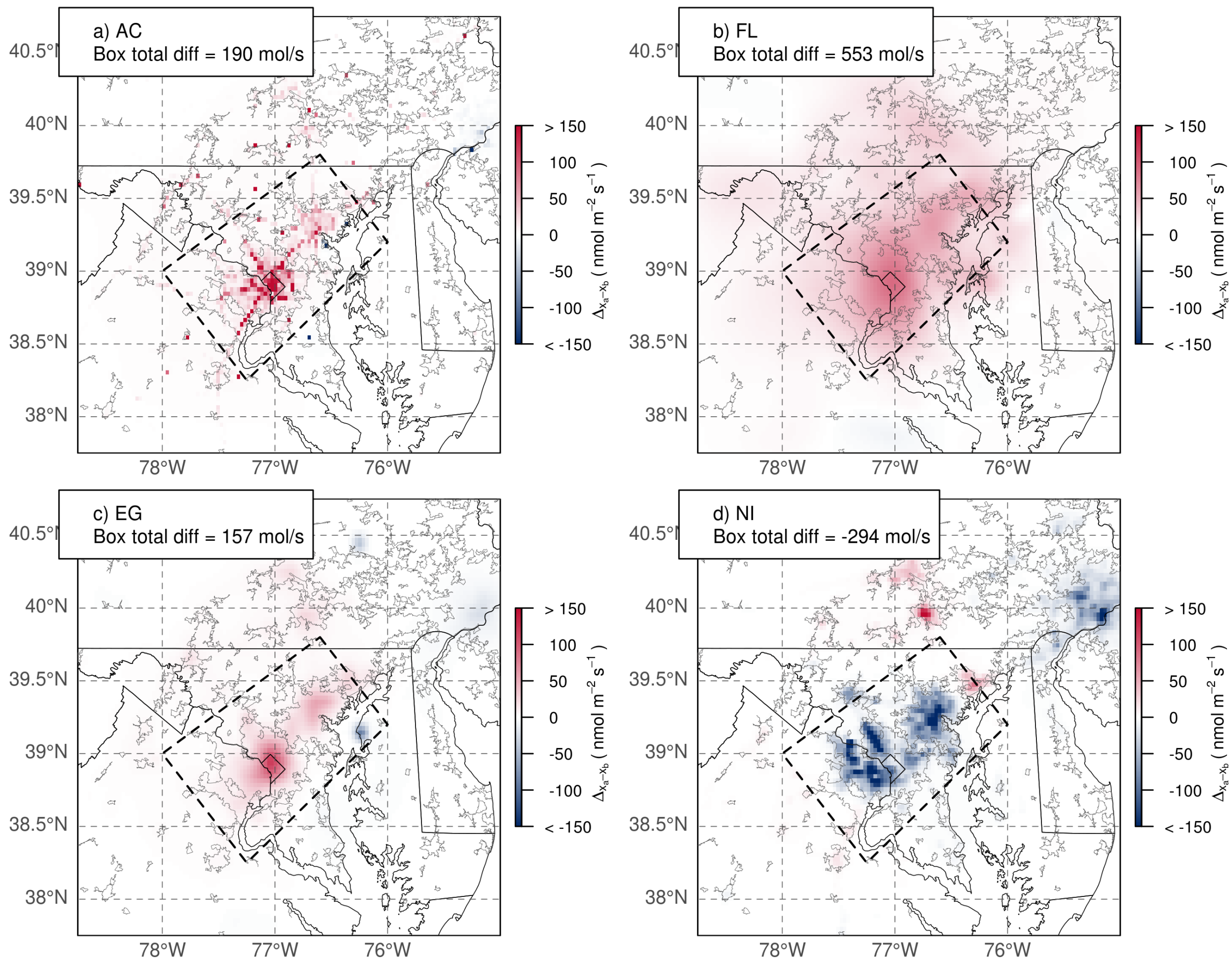


Figure S15: Spatial distribution of differences between the mean estimated CO emission rate and the prior emissions for all days and transport models using the different priors (a) AC is ACES inventory annual mean scaled using the mean observed $\Delta CO{:}\Delta CO_2$ ratio, (b) FL is the Flat inventory, (c) EG is EDGAR inventory and (d) NI is the NEI inventory

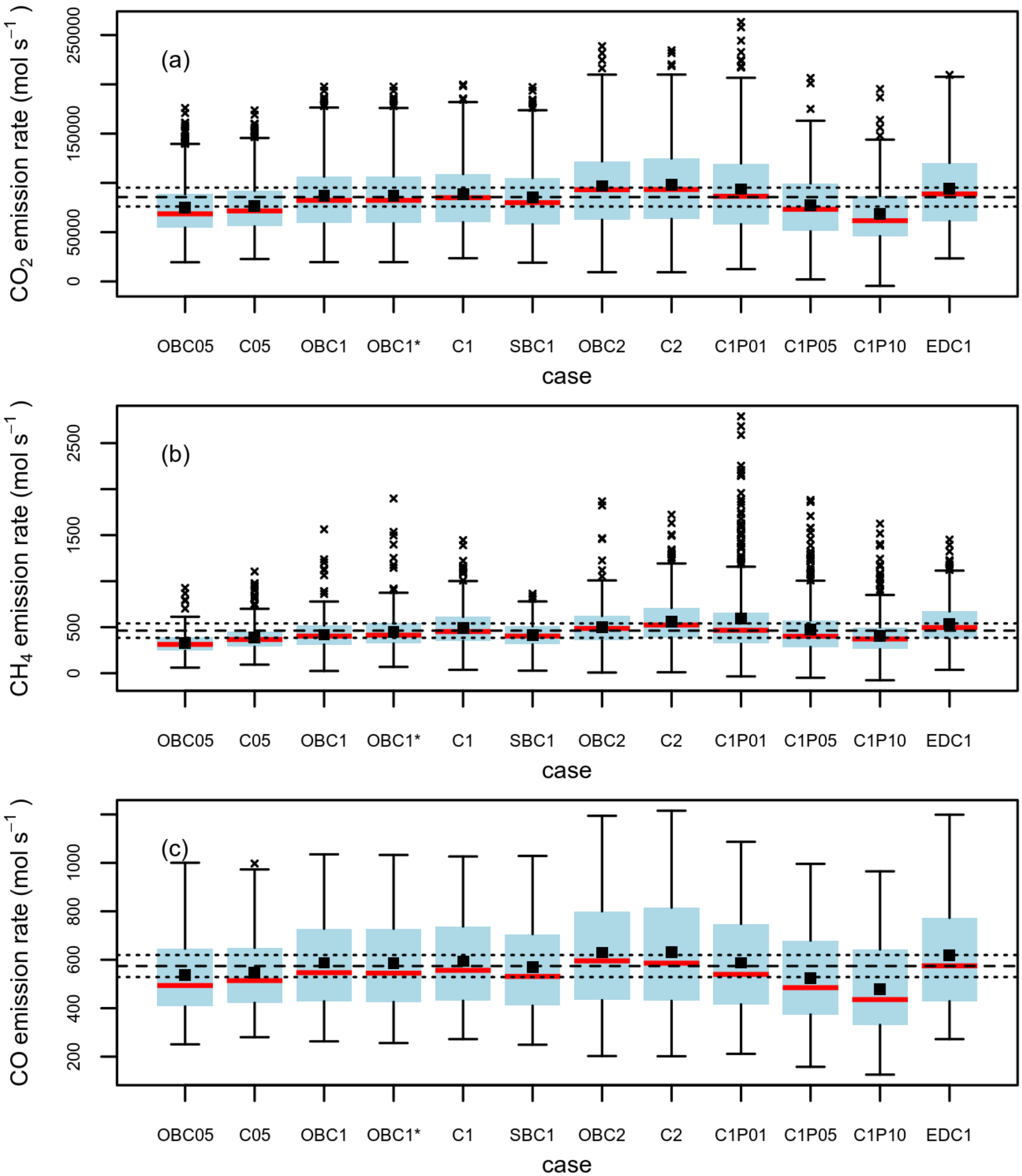


Figure S16: Boxplots of the sensitivity analysis for a) CO_2 (N = 9720), b) CH_4 (N = 4320) and c) CO (N = 1440) for the 12 cases where the covariances and background choice were changed: OB are cases with optimized Lagrangian background, SB is scaled Lagrangian background, cases C05, C1 and C2 are non-optimized Lagrangian background and C1P01, C1P05 and C1P10 are using a constant background determined by the quantile 1st, 5th or 10th respectively. Case EDC1 refers to diagonal transport error covariance. The C in all cases refers to the prior flux error covariance being 50%, 100% or 200%. Blue bars indicate the 25th and 75th quantiles, whiskers the range, x's the outliers (1.5 times the IQR), red line the median, square markers the mean, the dashed line the mean and the dotted lines the range $\pm$ 1-σ.

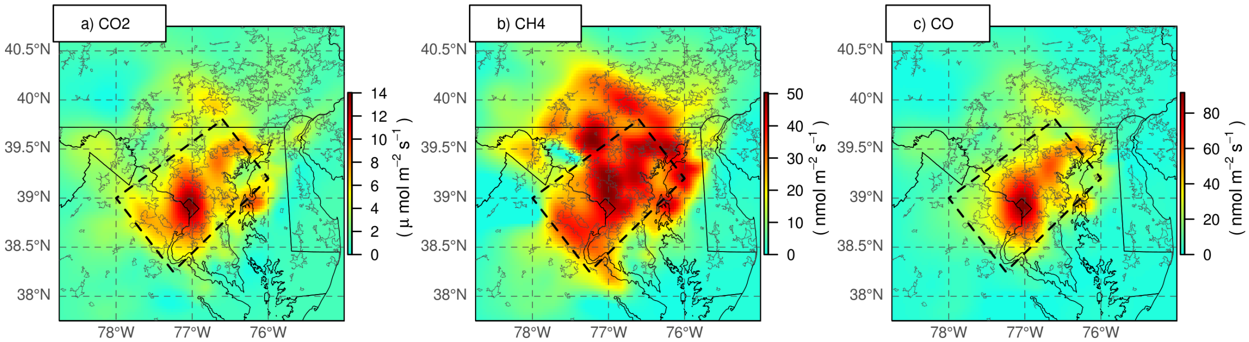


Figure S17: Posterior fluxes obtained using the flat prior averaged across the five days of the campaign.

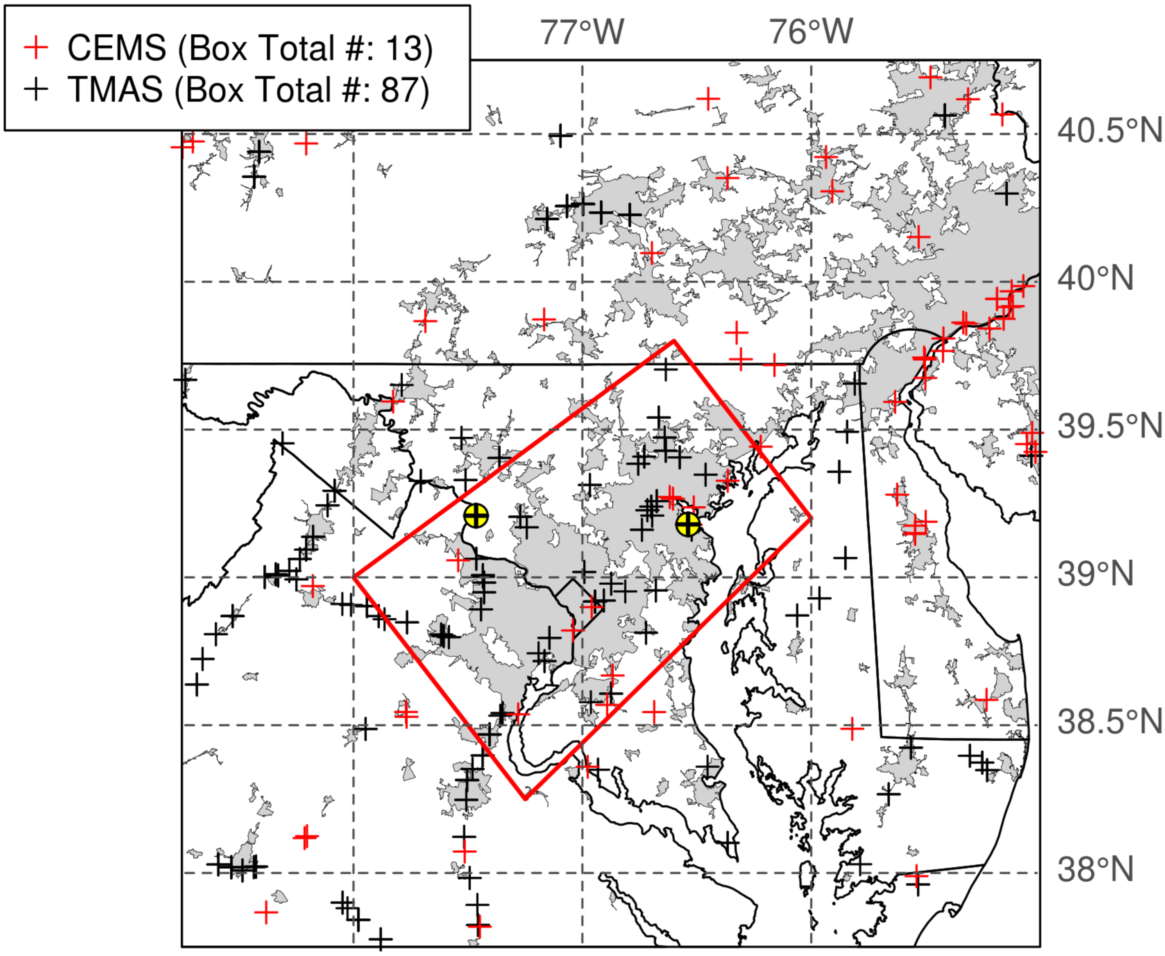


Figure S18: Location of the CEMS power plants and TMAS counting stations within the inversion domain. Accounting box also shown. The circled black crosses with yellow background are the Dickerson power plant (left) and Brandon Shores power plant (right).

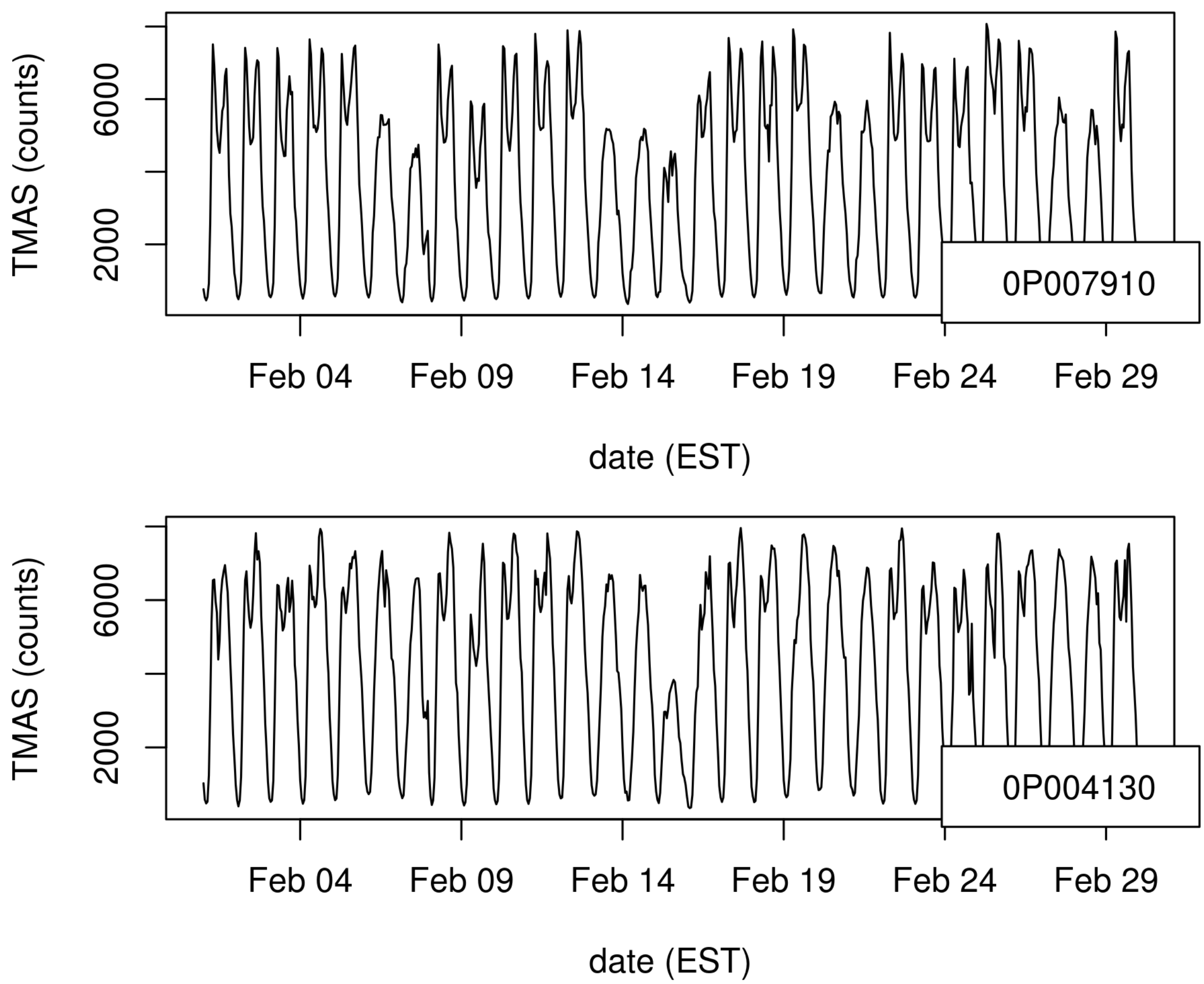


Figure S19: Hourly traffic counts for two TMAS stations placed in Washington, DC and Baltimore during the month of February 2016.

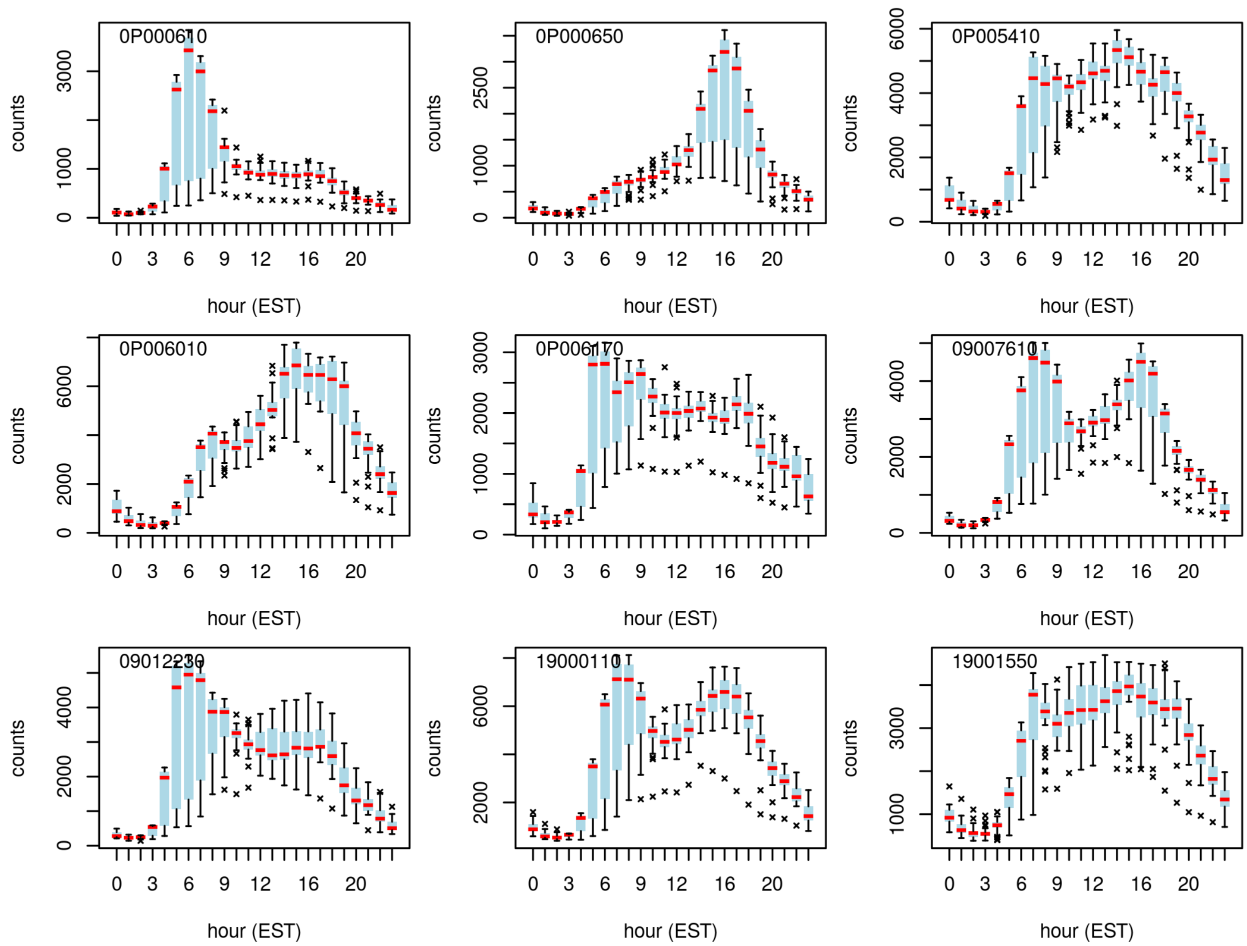


Figure S20: Daily cycle of the hourly traffic counts for nine TMAS stations placed within the accounting box for the month of February 2016.

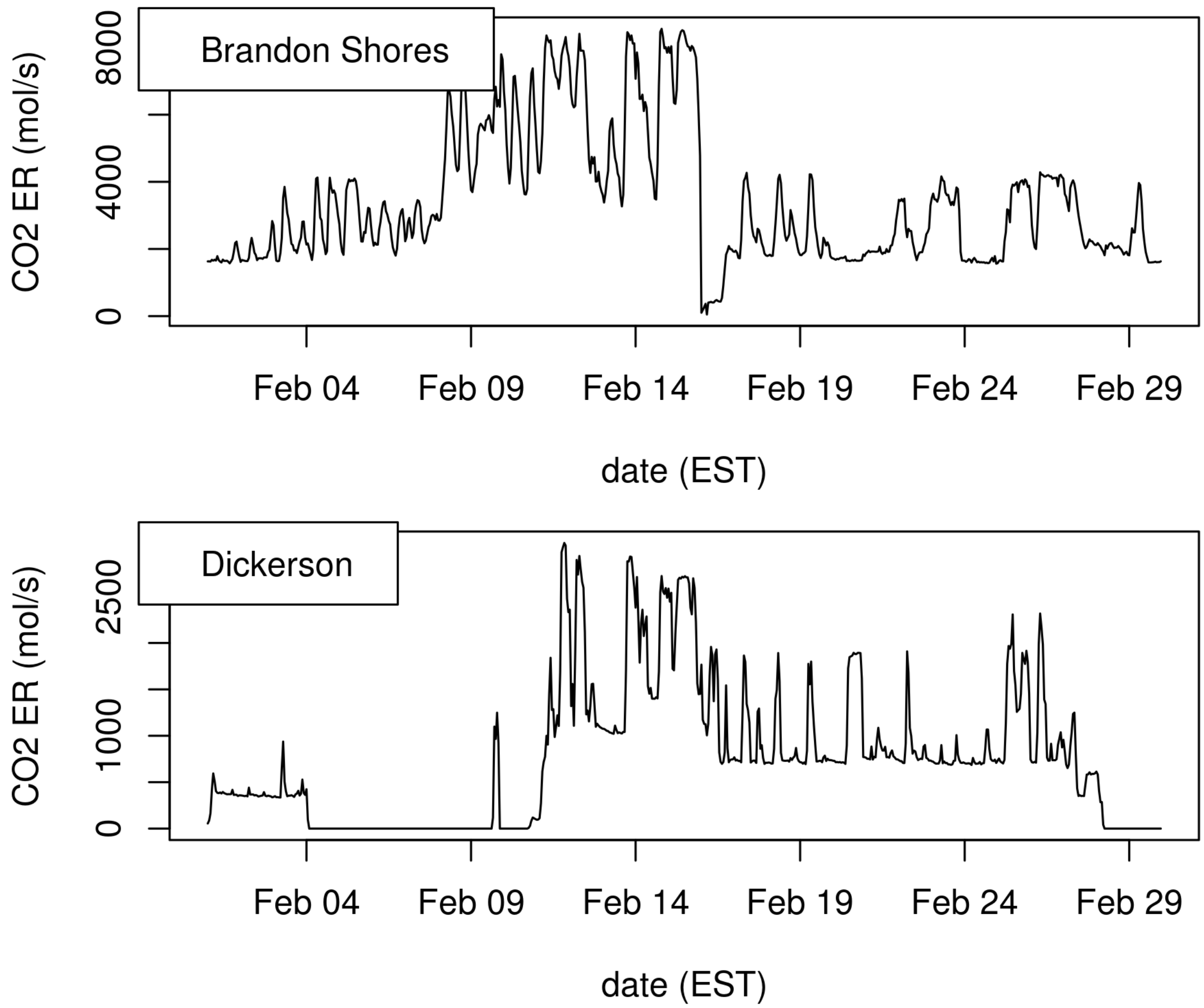


Figure S21: Hourly CO_2 emission rate for two Power Plants in the area during the month of February 2016.

3 Tables

Table S1: Transport model configurations summary with the labels used to identify them throughout the text.

Label	**Model**	**IC/BC**	**HYSPLIT vertical mixing**
HR	HRRR	RAP	Kantha / Clayson
YU	WRF-YSU	NARR	Kantha / Clayson
MY	WRF-MYNN	NARR	TKE
MY2	WRF-MYNN	NARR	Experimental (K_z)
QS	WRF-QNSE	HRRR	TKE
BL	WRF-BouLac+ UCM	HRRR	TKE

Table S2: Summary of the emissions inventories used as priors along with the labels used to identify them throughout the text.

Tracer	**Label**	**Name**	**Period**	**Total* (mol s^{-1})**
CO_2	VU	VULCAN	Feb – 2002	63 10^3
	OD	ODIAC	Feb – 2015	49 10^3
	DA	DARTE + LandUse	2012	42 10^3
	FF	FFDAS	2010	42 10^3
	AC	ACES	2011	59 10^3
	EB	ENSEMBLE	---	51 10^3
	AC2	ACES2	Feb – 2013 & 2014 (Afternoon hours)	94 10^3
	FL	FLAT	---	14 10^3
	SP	SIMPLE	---	42 10^3
CH_4	EP	EPA	2012	153
	EG	EDGAR	2012	237
	EB	ENSEMBLE	---	195
	FL	FLAT	---	14
CO	AC	ACES**	2011	362
	EG	EDGAR	2012	436
	NI	NEI	2011	932
	FL	FLAT	---	14

*Washington DC / Baltimore area accounting box. **Scaled using the mean observed $\Delta CO:\Delta CO_2$ ratio.

Table S3: Summary of the sensitivity analysis cases along with the labels used to identify them throughout the text.

Case	**Background**	**Prior Covariance**	**Transport Covariance**
OBC05	Optimized lagrangian background	50% prior emissions	Full covariance
C05	Non-Optimized lagrangian background	50% prior emissions	Full covariance
OBC1	Optimized lagrangian background	100% prior emissions	Full covariance
OBC1*	Optimized lagrangian background	100% prior emissions	Full covariance
C1	Non-Optimized lagrangian background	100% prior emissions	Full covariance
SBC1	Scaled lagrangian background	100% prior emissions	Full covariance
OBC2	Optimized lagrangian background	200% prior emissions	Full covariance
C2	Non-Optimized lagrangian background	200% prior emissions	Full covariance
C1P01	Constant background (P1%)	100% prior emissions	Full covariance
C1P05	Constant background (P5%)	100% prior emissions	Full covariance
C1P10	Constant background (P10%)	100% prior emissions	Full covariance
EDC1	Optimized lagrangian background	100% prior emissions	Diagonal covariance

*Uncertainty due to the potential mis-location of the reference value (σ_{mis}) excluded.

References

(1) Salmon, O. E.; Shepson, P. B.; Ren, X.; Marquardt Collow, A. B.; Miller, M. A.; Carlton, A. G.; Cambaliza, M. O. L.; Heimburger, A.; Morgan, K. L.; Fuentes, J. D.; Stirm, B. H.; Grundman II, R.; Dickerson, R. R. Urban emissions of water vapor in winter. *Journal of Geophysical Research: Atmospheres* **2017**, *122*, 9467–9484.

(2) Salmon, O. E.; Shepson, P. B.; Ren, X.; He, H.; Hall, D. L.; Dickerson, R. R.; Stirm, B. H.; Brown, S. S.; Fibiger, D. L.; McDuffie, E. E.; Campos, T. L.; Gurney, K. R.; Thornton, J. A. Top-Down Estimates of NOx and CO Emissions From Washington, D.C.-Baltimore During the WINTER Campaign. *Journal of Geophysical Research: Atmospheres* **2018**, *123*, 7705–7724.

(3) Ren, X.; Salmon, O. E.; Hansford, J. R.; Ahn, D.; Hall, D.; Benish, S. E.; Stratton, P. R.; He, H.; Sahu, S.; Grimes, C.; Heimburger, A. M. F.; Martin, C. R.; Cohen, M. D.; Stunder, B.; Salawitch, R. J.; Ehrman, S. H.; Shepson, P. B.; Dickerson, R. R. Methane Emissions From the Baltimore-Washington Area Based on Airborne Observations: Comparison to Emissions Inventories. *Journal of Geophysical Research: Atmospheres* **2018**, *123*, 8869–8882.

(4) Enting, I. *Inverse Problems in Atmospheric Constituent Transport*; Cambridge Atmospheric and Space Science Series; Cambridge University Press, 2002.

(5) Tarantola, A.; for Industrial, S.; Mathematics, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Other titles in applied mathematics; Society for Industrial and Applied Mathematics, 2005.

(6) Lopez-Coto, I.; Ghosh, S.; Prasad, K.; Whetstone, J. Tower-based greenhouse gas measurement network design—The National Institute of Standards and Technology North East Corridor Testbed. *Advances in Atmospheric Sciences* **2017**, *34*, 1095–1105.

(7) Stohl, A.; Seibert, P.; Arduini, J.; Eckhardt, S.; Fraser, P.; Greally, B. R.; Lunder, C.; Maione, M.; Mühle, J.; O'Doherty, S.; Prinn, R. G.; Reimann, S.; Saito, T.; Schmidbauer, N.; Simmonds, P. G.; Vollmer, M. K.; Weiss, R. F.; Yokouchi, Y. An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons. *Atmospheric Chemistry and Physics* **2009**, *9*, 1597–1620.

(8) Lauvaux, T.; Schuh, A. E.; Uliasz, M.; Richardson, S.; Miles, N.; Andrews, A. E.; Sweeney, C.; Diaz, L. I.; Martins, D.; Shepson, P. B.; 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* **2012**, *12*, 337–354.

(9) Henne, S.; Brunner, D.; Oney, B.; Leuenberger, M.; Eugster, W.; Bamberger, I.; Meinhardt, F.; Steinbacher, M.; Emmenegger, L. Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling. *Atmospheric Chemistry and Physics* **2016**, *16*, 3683–3710.

(10) Stein, A. F.; Draxler, R. R.; Rolph, G. D.; Stunder, B. J. B.; Cohen, M. D.; Ngan, F. NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society* **2015**, *96*, 2059–2077.

(11) Lin, J. C.; Gerbig, C.; Wofsy, S. C.; Andrews, A. E.; Daube, B. C.; Davis, K. J.; Grainger, C. A. A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research: Atmospheres* **2003**, *108*.

(12) Benjamin, S. G.; Weygandt, S. S.; Brown, J. M.; Hu, M.; Alexander, C. R.; Smirnova, T. G.; Olson, J. B.; James, E. P.; Dowell, D. C.; Grell, G. A.; Lin, H.; Peckham, S. E.; Smith, T. L.; Moninger, W. R.; Kenyon, J. S.; Manikin, G. S. A North

American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review* **2016**, *144*, 1669–1694.

(13) Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D.; Duda, M. G.; yu Huang, X.; Wang, W.; Powers, J. G. *A Description of the Advanced Research WRF Version 3*; 2008.

(14) Nakanishi, M.; Niino, H. An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorology* **2006**, *119*, 397–407.

(15) Hong, S.-Y. A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. *Quarterly Journal of the Royal Meteorological Society* **2010**, *136*, 1481–1496.

(16) Mesinger, F.; DiMego, G.; Kalnay, E.; Mitchell, K.; Shafran, P. C.; Ebisuzaki, W.; JoviÄĞ, D.; Woollen, J.; Rogers, E.; Berbery, E. H.; Ek, M. B.; Fan, Y.; Grumbine, R.; Higgins, W.; Li, H.; Lin, Y.; Manikin, G.; Parrish, D.; Shi, W. North American Regional Reanalysis. *Bulletin of the American Meteorological Society* **2006**, *87*, 343–360.

(17) Sukoriansky, S.; Galperin, B.; Perov, V. 'Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice'. *Boundary-Layer Meteorology* **2005**, *117*, 231–257.

(18) Bougeault, P.; Lacarrere, P. Parameterization of Orography-Induced Turbulence in a Mesobeta–Scale Model. *Monthly Weather Review* **1989**, *117*, 1872–1890.

(19) Martilli, A.; Clappier, A.; Rotach, M. W. An urban surface exchange parameterisation for mesoscale models. *Boundary-layer meteorology* **2002**, *104*, 261–304.

(20) Blaylock, B. K.; Horel, J. D.; Crosman, E. T. Impact of Lake Breezes on Summer Ozone

Concentrations in the Salt Lake Valley. *Journal of Applied Meteorology and Climatology* **2017**, *56*, 353–370.

(21) Mlawer, E. J.; Taubman, S. J.; Brown, P. D.; Iacono, M. J.; Clough, S. A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres* **1997**, *102*, 16663–16682.

(22) Thompson, G.; Rasmussen, R. M.; Manning, K. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Monthly Weather Review* **2004**, *132*, 519–542.

(23) Thompson, G.; Field, P. R.; Rasmussen, R. M.; Hall, W. D. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review* **2008**, *136*, 5095–5115.

(24) Chen, F.; Dudhia, J. Coupling an Advanced Land SurfaceâĂŞHydrology Model with the Penn StateâĂŞNCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Monthly Weather Review* **2001**, *129*, 569–585.

(25) Stull, R. B., Ed. *An Introduction to Boundary Layer Meteorology*; Springer Netherlands, 1988.

(26) Shin, H. H.; Hong, S.-Y. Intercomparison of Planetary Boundary-Layer Parametrizations in the WRF Model for a Single Day from CASES-99. *Boundary-Layer Meteorology* **2011**, *139*, 261–281.

(27) Karion, A.; Lauvaux, T.; Coto, I. L.; Sweeney, C.; Mueller, K.; Gourdji, S.; Angevine, W.; Barkley, Z.; Deng, A.; Andrews, A.; Stein, A.; Whetstone, J. Intercomparison of atmospheric trace gas dispersion models: Barnett Shale case study. *Atmospheric Chemistry and Physics* **2019**, *19*, 2561–2576.

(28) Everitt, B., Ed. *Cluster Analysis*; London: Heinemann Educational Books, 1974.

(29) Hartigan, J. A. *Clustering Algorithms*; John Wiley & Sons, Inc.: New York, NY, USA, 1975.

(30) Fovell, R. G.; Fovell, M.-Y. C. Climate Zones of the Conterminous United States Defined Using Cluster Analysis. *Journal of Climate* **1993**, *6*, 2103–2135.

(31) Gurney, K. R.; Mendoza, D. L.; Zhou, Y.; Fischer, M. L.; Miller, C. C.; Geethakumar, S.; de la Rue du Can, S. High Resolution Fossil Fuel Combustion CO2 Emission Fluxes for the United States. *Environmental Science & Technology* **2009**, *43*, 5535–5541, PMID: 19708393.

(32) Oda, T.; Maksyutov, S. A very high-resolution (1 km x 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmospheric Chemistry and Physics* **2011**, *11*, 543–556.

(33) Rayner, P. J.; Raupach, M. R.; Paget, M.; Peylin, P.; Koffi, E. A new global gridded data set of CO2 emissions from fossil fuel combustion: Methodology and evaluation. *Journal of Geophysical Research: Atmospheres* **2010**, *115*, D19306.

(34) Gately, C. K.; Hutyra, L. R.; Sue Wing, I. Cities, traffic, and CO2: A multidecadal assessment of trends, drivers, and scaling relationships. *Proceedings of the National Academy of Sciences* **2015**, *112*, 4999–5004.

(35) Gately, C. K.; Hutyra, L. R. Large Uncertainties in Urban-Scale Carbon Emissions. *Journal of Geophysical Research: Atmospheres* **2017**, *122*, 11,242–11,260.

(36) Maasakkers, J. D.; Jacob, D. J.; Sulprizio, M. P.; Turner, A. J.; Weitz, M.; Wirth, T.; Hight, C.; DeFigueiredo, M.; Desai, M.; Schmeltz, R.; Hockstad, L.; Bloom, A. A.; Bowman, K. W.; Jeong, S.; Fischer, M. L. Gridded National Inventory of U.S. Methane Emissions. *Environmental Science & Technology* **2016**, *50*, 13123–13133, PMID: 27934278.

(37) Janssens-Maenhout, G.; Crippa, M.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Dentener, F.; Bergamaschi, P.; Pagliari, V.; Olivier, J. G. J.; Peters, J. A. H. W.; van Aardenne, J. A.; Monni, S.; Doering, U.; Petrescu, A. M. R. EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012. *Earth System Science Data Discussions* **2017**, *2017*, 1–55.

(38) Crippa, M.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Dentener, F.; van Aardenne, J. A.; Monni, S.; Doering, U.; Olivier, J. G. J.; Pagliari, V.; Janssens-Maenhout, G. Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. *Earth System Science Data* **2018**, *10*, 1987–2013.

(39) EPA National Emissions Inventory. `https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data`, 2011.

(40) Cambaliza, M. O. L.; Shepson, P. B.; Caulton, D. R.; Stirm, B.; Samarov, D.; Gurney, K. R.; Turnbull, J.; Davis, K. J.; Possolo, A.; Karion, A.; Sweeney, C.; Moser, B.; Hendricks, A.; Lauvaux, T.; Mays, K.; Whetstone, J.; Huang, J.; Razlivanov, I.; Miles, N. L.; Richardson, S. J. Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions. *Atmospheric Chemistry and Physics* **2014**, *14*, 9029–9050.

(41) Karion, A.; Sweeney, C.; Kort, E. A.; Shepson, P. B.; Brewer, A.; Cambaliza, M.; Conley, S. A.; Davis, K.; Deng, A.; Hardesty, M.; Herndon, S. C.; Lauvaux, T.; Lavoie, T.; Lyon, D.; Newberger, T.; Petron, G.; Rella, C.; Smith, M.; Wolter, S.; Yacovitch, T. I.; Tans, P. Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. *Environmental Science & Technology* **2015**, *49*, 8124–8131, PMID: 26148550.

(42) Mueller, K. L.; Gourdji, S. M.; Michalak, A. M. Global monthly averaged CO2 fluxes

recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements. *Journal of Geophysical Research: Atmospheres* **2008**, *113*.

(43) Shiga, Y. P.; Michalak, A. M.; Randolph Kawa, S.; Engelen, R. J. In-situ CO2 monitoring network evaluation and design: A criterion based on atmospheric CO2 variability. *Journal of Geophysical Research: Atmospheres* **2013**, *118*, 2007–2018.

(44) 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.; 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* **2016**, *121*, 5213–5236.

(45) Oda, T.; Maksyutov, S.; Andres, R. J. The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data* **2018**, *10*, 87–107.

(46) Andres, R. J.; Boden, T. A.; Higdon, D. M. Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example. *Atmospheric Chemistry and Physics* **2016**, *16*, 14979–14995.

(47) Engelen, R. J.; Denning, A. S.; Gurney, K. R. On error estimation in atmospheric CO2 inversions. *Journal of Geophysical Research: Atmospheres* **2002**, *107*, ACL 10–1–ACL 10–13.

(48) Desroziers, G.; Berre, L.; Chapnik, B.; Poli, P. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society* **2005**, *131*, 3385–3396.