Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics Discussions
Open Access

Characterizing Uncertainties in Atmospheric Inversions of Fossil Fuel CO_2 Emissions in California

Kieran Brophy[1], Heather Graven[1], Alistair J. Manning[2], Emily White[3], Tim Arnold[4,5], Marc L Fischer[6], Seongeun Jeong[6], Xinguang Cui[6], Matthew Rigby[3]

1. Department of Physics, Imperial College London, London, UK
2. Hadley Centre, Met Office, Exeter, UK
3. School of Chemistry, University of Bristol, Bristol, UK
4. National Physical Laboratory, London, UK
5. University of Edinburgh, Edinburgh, UK
6. Lawrence Berkeley National Laboratory, Berkeley, CA

Corresponding author:
Kieran Brophy
Email: kb613@ic.ac.uk

Key Words: Fossil Fuel, Carbon Dioxide, Simulation Experiment, Inversion, Transport Error

Abstract

Atmospheric inverse modelling has become an increasingly useful tool for evaluating emissions of greenhouse gases including methane, nitrous oxide and synthetic gases such as hydrofluorocarbons (HFCs). Atmospheric inversions for emissions of CO_2 from fossil fuel combustion ($ffCO_2$) are currently being developed. The aim of this paper is to investigate potential errors and uncertainties related to the spatial and temporal prior representation of emissions and modelled atmospheric transport for the inversion of $ffCO_2$ emissions in the U.S. state of California. We perform simulation experiments based on a network of ground-based observations of CO_2 concentration and radiocarbon in CO_2 (a tracer of $ffCO_2$), combining prior (bottom-up) emission models and transport models currently used in many atmospheric studies. The potential effect of errors in the spatial and temporal distribution of prior emission estimates is investigated in experiments by using perturbed versions of the emissions estimates used to create the pseudo data. The potential effect of transport error was investigated by using three different atmospheric transport models for the prior and pseudo data simulations. We find that the magnitude of biases in posterior state-total emissions arising from errors in the spatial and temporal distribution in prior emissions in these experiments are 1-15% of posterior state-total emissions, and generally smaller than the 2-σ uncertainty in posterior emissions. Transport error in these experiments introduces biases of -10% to +6% in posterior state-total emissions. Our results indicate that uncertainties in posterior state-total $ffCO_2$ estimates arising from the choice of prior emissions or atmospheric transport model are on the order of 15% or less for the ground-based network in California we consider. We highlight the need for temporal variations to be included in prior emissions, and for continuing efforts to evaluate and improve the representation of atmospheric transport for regional $ffCO_2$ inversions.

1. Introduction

The U.S. state of California currently emits roughly 100 Tg C of fossil fuel CO_2 ($ffCO_2$) each year (CARB, 2017), or approximately 1% of global emissions (Boden et al., 2016). The passing of California's "Global Warming Solutions Act" (AB-32) in 2006 requires that overall greenhouse gas emissions in California be reduced to their 1990 levels by 2020 (a 15% reduction compared to business as usual emissions) with further reductions of 40% below 1990 levels planned for 2030, and 80% below by 2050. The California Air Resources Board (CARB) is responsible for developing and maintaining a "bottom-up" inventory of greenhouse gas emissions to verify these reduction targets. However, previous studies have shown such inventories may have errors or incomplete knowledge of sources (e.g. Marland et al, 1999; Andres et al., 2012). Uncertainties in inventories of annual $ffCO_2$ emissions from most developed countries (i.e. UNFCCC Annex I and Annex II) have been estimated to be between 5-

10% (Andres et al., 2012), and uncertainties can become much larger at subnational levels (Hogue et al., 2016). In a recent study Fischer et al., (2017) found discrepancies between bottom-up gridded inventories of $ffCO_2$ emissions were 11% of California's state total emissions.

Previous research has shown that inferring $ffCO_2$ emissions from atmospheric measurements, including measurements of $ffCO_2$ tracers, could provide independent emissions estimates on urban to continental scales (e.g. Basu et al., 2016; Fischer et al., 2017; Graven et al. 2018; INFLUX ref). Such estimates are derived from observations through the use of an atmospheric chemical transport model and a suitable inverse method in a process often referred to as "inverse modelling" or an "inversion". Distinguishing enhancements of CO_2 due to anthropogenic or biogenic sources can be done using measurements of radiocarbon in CO_2 ($\Delta^{14}CO_2$), since CO_2 emitted from fossil fuel combustion is devoid of $^{14}CO_2$ due to radioactive decay (Levin et al., 2003).

Recent studies have shown that both simulated (Fischer et al. 2017) and observed (Graven et al. 2018) measurements of $\Delta^{14}CO_2$ at a network of sites could be used to estimate monthly mean California $ffCO_2$ emissions in a regional inversion with posterior uncertainties of ~5-8%, levels that are useful for the evaluation of bottom-up $ffCO_2$ emissions estimates. Furthermore, Graven et al., 2018 found their posterior emissions estimates were not significantly different from the California Air Resources Board's reported $ffCO_2$ emissions, providing tentative validation of California's reported $ffCO_2$ emissions in 2014-15. In another study using aircraft-based $\Delta^{14}CO_2$ measurements, Turnbull et al. (2011) found $ffCO_2$ emissions from Sacramento County in February 2009 were broadly consistent (mean difference of -17%, range: -43% to +133%) with the Vulcan emissions estimate (Gurney et al., 2009).

Although atmospheric inversions may provide a method for estimating emissions that is useful for evaluating emissions reduction policies, such as AB-32, systematic errors can arise from the atmospheric transport and prior emission models (e.g., Nassar et al., 2014; Liu et al., 2014; Hungershoefer et al., 2010; Chevallier et al., 2009, Gerbig et al., 2003). Comparisons of CO_2 simulated by different transport models have been conducted globally (e.g. Gurney et al., 2003, Peylin et al. 2013), and on the European continental scale (Peylin et al., 2011). The latter found that transport model error resulted in differences in modelled $ffCO_2$ concentrations that were 2-3 times larger than using the same transport model but different prior emissions, depending on the location and time of year. However, comparisons of $ffCO_2$ simulated by different high resolution models (25 km or less) at regional scales are still lacking.

The objective of this paper is to examine the sensitivity of a regional inversion for Californian $ffCO_2$ emissions to errors in the prior emissions estimate and transport model. We build on previous work by Fischer et al. (2017) that developed an Observation System Simulation Experiment to estimate the uncertainties in both California statewide $ffCO_2$ emissions and biospheric fluxes that might be obtained using an atmospheric inversion. Their inversion was driven by a combination of in situ tower measurements, satellite column measurements from OCO-2, prior flux estimates, a regional atmospheric transport modelling system, and estimated uncertainties in prior CO_2 flux models, $ffCO_2$ measurements using radiocarbon, OCO-2 measurements, and in atmospheric transport. In contrast to Fischer et al., 2017 we focus only on $ffCO_2$ emissions and use a network of flask samples without incorporating satellite measurements.

Our approach is to use simulation experiments to quantify representation and transport error using the inversion setup and the observation network from Graven et al. (2018) as a test case.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Specifically we test whether the inversion can estimate the "true" emissions that were used to produce the pseudo data, within the uncertainties, when the prior emissions estimate includes spatial and temporal representation errors within the scope of current emissions estimates (Vulcan v2.2 and EDGAR v4.2 FT2010). We further test whether the inversion can estimate "true" emissions, within the uncertainties, when the transport model used for the prior simulation is different from the transport model used to produce the pseudo data, emulating transport error.

2. Data and Methods

The analysis approach applies a Bayesian inversion developed from previous work that combines atmospheric observations, atmospheric transport modelling, prior flux models, and an uncertainty specification (Jeong et al., 2013; Fischer et al., 2017). Here, the inversion scales prior emission estimates in 15 regions (Figure 1a, Table 1) termed "air basins", classified by the California Air Resources Board for air quality control (https://www.arb.ca.gov/desig/adm/basincnty.htm).

2.1 Observation Network

As a test case to explore uncertainties in $ffCO_2$ inversions, we use the observation network of 9 tower sites in California that was used to collect flask samples for measurements of CO_2 and radiocarbon in CO_2 in 2014-15 (Figure 1a) (Graven et al. 2018). Three month-long campaigns were conducted in May 2014, October-November 2014 and January-February 2015, with flasks sampled approximately every 2-3 days at 22:30 GMT (14:30 local standard time). The time of observation was chosen as the planetary boundary layer is usually deepest in the afternoon so that errors in the modelled boundary layer concentration are considered smaller (Jeong et al., 2013), and afternoon concentrations are more representative of large regions.

The observed $ffCO_2$ concentration at a given site can be calculated by (Levin et al., 2003; Turnbull et al. 2009):

$$ffCO_2 = C_{obs}\left(\frac{\Delta_{bg}-\Delta_{obs}}{\Delta_{bg}-\Delta_{ff}}\right) + \beta \qquad (1)$$

Where C_{obs} is the total observed CO_2 concentration at a given site. Δ refers to $\Delta^{14}C$, the ratio of $^{14}C/C$ reported in part per thousand deviation from a standard ratio, including corrections for mass-dependent isotopic fractionation and sample age (Stuiver and Polach, 1977). Δ_{bg}, Δ_{obs} and Δ_{ff} are the $\Delta^{14}CO_2$ of background, observed and fossil fuel CO_2, respectively, where Δ_{ff} is -1000‰ since $ffCO_2$ is devoid of $^{14}CO_2$. The term β is a correction for the effect of other influences on $\Delta^{14}CO_2$, principally heterotrophic respiration (Turnbull et al. 2009). Following Fischer et al. (2017), total observational uncertainty for $ffCO_2$ was assumed to be 1.5 ppm (1-σ). This is consistent with Graven et al. (2018), who estimated total uncertainty in $ffCO_2$ for individual samples of 1.0 to 1.9 ppm.

2.2 Prior Emissions Estimates and Prior Uncertainty

The two prior emissions estimates used here are gridded products produced by EDGAR (version FT2010) (EDGAR, 2011) for the year 2008 and Vulcan (version 2.2) for 2002 (Gurney et al., 2009). EDGAR is produced at an annual resolution whilst Vulcan has an hourly resolution. The two models use different emissions data and different methods to spatially allocate emissions with annually averaged statewide emissions differing by 17.8 TgC (~19% of mean emissions), and up to 11.6 TgC for individual air basins of California (Table 1). Although our campaigns took place in 2014-2015, we use emissions estimates from Vulcan for the year 2002 and EDGAR for 2008 as emissions estimates are not available from Vulcan and EDGAR for 2014-15. The

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

difference in state total emissions between 2002, 2008 and 2014-15 is 3-6 TgC (CARB, 2017), much less than the EDGAR-Vulcan difference of 17.8 TgC.

We estimate prior uncertainty in the same way as in Fischer et al. (2017), using a comparison of four gridded emissions estimates in California (Vulcan v2.2, EDGAR FT2010, ODIAC v2013 and FFDAS v2) as well as a comparison across an ensemble of emissions estimates for one model (FFDAS v2, Asefi-Najafabady et al.,2014). The relative 1-σ standard deviation is between 8% and 100% for individual air basins (Table 1), and this is what we use to specify the 1-σ uncertainty in the prior emissions from each air basin. This estimate of prior uncertainty is referred to as "SD prior uncertainty". We also conduct tests with an alternative prior uncertainty of 70% for each air basin (referred to as "70% prior uncertainty"). This was done to test the sensitivity of our results to the choice of prior uncertainty. Emissions occurring outside California were assumed to have an uncertainty of 100% for both cases.

2.3 Atmospheric Transport Models

We simulate ffCO_2 using three different atmospheric transport models outlined in Table 2. These models are commonly used in regional atmospheric transport modelling and greenhouse gas inversion studies but to date have not been compared in California. Two of the transport models use different versions and parameterizations of the Weather Research and Forecast (WRF) model combined with the Stochastic Time-Inverted Lagrangian Transport (STILT) model. The third transport model uses meteorology from the UK Met Office's Unified Model (UM) combined with the Numerical Atmospheric dispersion Modelling Environment (NAME).

The first WRF-STILT model is run at Lawrence Berkeley National Laboratory (WS-LBL, Fischer et al. 2017; Jeong et al. 2016; Bagley et al. 2017) and uses WRF version 3.5.1 (Lin, 2003; Nehrkorn et al., 2010). Estimates for Planetary Boundary Layer Height (PBLH) are based on the Mellor–Yamada–Nakanishi–Niino version 2 (MYNN2) parameterization (Nakanishi and Niino 2004, 2006). As in Jeong et al. (2016), Fischer et al. (2017) and Bagley et al. (2017), two land surface models (LSMs) are used depending on the location of the observation site. A 5-layer thermal diffusion land surface model is used in the Central Valley for the May campaign whilst the Noah LSM (Chen et al., 2001) is used in the remaining campaigns and regions of California. We implement multiple nested domains, with the outermost domain spanning 16-59°N and 154-137°W with a 36km resolution, a second domain of 12km resolution over western North America, and a third domain of 4km resolution over California. Two urban domains of 1.3 km resolution are used in the San Francisco Bay area and the metropolitan area of Los Angeles. Footprints describing the sensitivity of an observation to surface emissions are calculated by simulating 500 model particles and tracking them backward for 7 days. The footprint of a given site and observation time is produced hourly for particles below 0.5 times the PBLH.

The second WRF-STILT model is from CarbonTracker-Lagrange (WS-CTL), an effort led at NOAA to produce standard footprints for greenhouse gas observation sites in North America (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker-lagrange). WS-CTL uses WRF version 2.1.2 and the Yonsei University (YSU) (Hong et al., 2006) PBLH scheme coupled with the Noah land surface model and the MM5 (fifth generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model, Grell et al., 1994) similarity theory-based surface layer scheme. As with WS-LBL, simulations are run for 7 days and particles below 0.5 times the PBLH are used in the calculation of the footprint. Footprints have a spatial resolution of 0.1° for the first 24 hours and 1° for the remaining 6 days. Footprints are hourly dis-aggregated for the first 24 hours and then aggregated for the remaining 6 days. The 0.1° spatial resolution domain is 31° longitude by 21° latitude with the domain centered on the release location. The 1° resolution has a domain of 170°E to 50°E longitude and 10° N to 80° N latitude. The WRF domain covers most of continental North America (Fig. 1 in Nehrkorn et al., 2010) with 30 km resolution and has a finer nest with 10 km spatial resolution over the continental United States. WS-CTL simulates footprints for 500 particles released over a 2-hour period between 21:00 and 23:00 GMT (13:00 and 15:00 PST). An exception is Sutro Tower (STR), where footprints are only available for an instantaneous release of 500 particles at 22:10 GMT. Walnut Grove (WGC) footprints are available only for a release height of 30m a.g.l, which is lower than the sampling height of 91m a.g.l. used in the observation campaign (Graven et al. 2018) and used in the other two transport models. Footprints were available for 2014 but not for 2015, so the WS-CTL model is used for simulations of the May and Oct-Nov 2014 campaigns but not for the Jan-Feb 2015 campaign.

The third model, UM-NAME, is the UK Met Office's NAME model, Version 3.6.5 (Jones et al., 2007), driven by meteorology from the Met Office's global numerical weather prediction model, the Unified Model (UM) (Cullen et al., 1993). The UM model has a horizontal resolution of ~25 km up to July 2014, covering the period of the May 2014 campaign. The horizontal resolution was then increased to ~17 km covering both the October-November 2014 and January-February 2015 campaigns. The temporal resolution of the UM meteorology is every 3 hours for all campaigns. Following a similar approach as for the WRF-STILT models, 500 particles were released instantaneously at 22:30 GMT and simulated for hourly dis-aggregated footprints for the first 24 hours and aggregated for the remaining 6 days. The footprints are calculated for the same horizontal resolution as the UM meteorology (25 or 17km), where the particles present in the

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

layer between 0 and 100 m above ground level are used to calculate the footprint. The computational domain covers 175.0°W to 75°W longitude and 6.0°N to 74°N latitude.

Simulated $ffCO_2$ signals (the enhancement of CO_2 concentration due to $ffCO_2$ emissions within the model domain) are calculated by taking the product of the footprint and an emissions estimate. Following previous work, we assume a transport model uncertainty of 0.5 times the mean monthly signal in the pseudo-observations at each site (Jeong et al., 2013; Fischer et al 2017).

Ten ensembles were run for UM-NAME to test the effect of random errors on the calculation of the footprint. The RMSE was within 10% of the mean monthly signal for most observation sites. This is similar to the findings of Jeong et al. (2012), which the transport model uncertainty is based on. Two observation sites (THD and VTR) had slightly higher RMSE, but both were within 20% of the mean monthly signal.

2.4 Inversion Method

Our inversion method is a Bayesian synthesis inversion to scale emissions in separate regions of California. We follow the same approach as Fischer et al. (2017) to solve for a vector of scaling factors, λ, for 15 air basins and a 16th region representing the area outside of California. Unlike Fischer et al. (2017), we do not split the San Joaquin Valley into two regions. The inversion uses the set of observations, c, and the matrix of predicted $ffCO_2$ signals from each air basin, K, to optimize the cost function J:

$$J_\lambda = (c - \boldsymbol{K}\lambda)^T \boldsymbol{R}^{-1} (c - \boldsymbol{K}\lambda) + (\lambda - \lambda_{prior})^T \boldsymbol{Q}_\lambda{}^{-1} (\lambda - \lambda_{prior}) \quad (2)$$

λ_{prior} is the prior estimate of the scaling factors (a vector of ones with length equal to the number of regions) and R and Q_λ are the error covariance matrices relating to observational and model transport errors, and prior emissions estimate errors respectively. The non-diagonal elements of R and Q_λ are zero, assuming uncorrelated errors in the prior emissions in each air basin and in the model and observations. This assumption for R is robust as we only generate one pseudo observation every 2-3 days. Included in R are observational errors and transport model errors, added in quadrature. Therefore if the average signal at an observation site is very small, then observational uncertainty (1.5 ppm) will dominate R. Minimizing J using the standard least squares formulation under the assumption of Gaussian distributed uncertainties gives the posterior estimate for λ following:

$$\boldsymbol{\lambda}_{post} = \left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{K} + \mathbf{Q}_\lambda^{-1}\right)^{-1}\left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{c} + \mathbf{Q}_\lambda^{-1}\boldsymbol{\lambda}_{prior}\right) \qquad (3)$$

With the posterior error covariance given as:

$$V_{post} = \left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{K} + \mathbf{Q}_\lambda^{-1}\right)^{-1} \qquad (4)$$

λ_{post} and V_{post} are computed separately for each of the three campaigns outlined in section 2.1. Posterior emissions estimates are the product of λ_{post} and the prior emissions estimate from each air basin. State total emissions are then calculated by summing the emissions in each air basin. Uncertainty in the state-wide Californian posterior flux, including error correlations, is calculated as:

$$\boldsymbol{\sigma}_{\mathbf{E}}^2 = \boldsymbol{E}_{prior}\boldsymbol{V}_{post}\boldsymbol{E}_{prior}^T \qquad (5)$$

Where E_{prior} is a vector of $ffCO_2$ emissions from each air basin.

2.5 Simulation Experiments

We conduct a series of experiments to test the performance of the inversion in estimating the true emissions when the emissions estimates or transport models used to produce pseudo-observations are different to those used to produce the prior simulations. The tests explore the effect differences in the magnitude, spatial distribution, and temporal variation of prior emissions have on posterior emissions. We also examine the effect of using different transport models to simulate pseudo observations and to simulate prior concentrations.

As part of these experiments, we evaluate the impact of outlier removal on the simulation experiments. Outlier removal is generally used in atmospheric inversions when there is an issue with the ability of the model to simulate a particular observation. We use the outlier removal method outlined in Graven et al. (2018) and compare with inversion results where no outliers are removed. In this outlier removal method, an observation (here, a pseudo-observation) is designated as an outlier if (1) the absolute difference between the $ffCO_2$ signals in the observation and the prior simulation is greater than the average of the observed and simulated $ffCO_2$, and (2) either the observed or simulated $ffCO_2$ is greater than 5 ppm.

2.5.1 Difference in magnitude of emissions

First we test how well the inversion estimates the true emissions if the prior emissions have a systematic error in magnitude, but no error in the spatial or temporal distribution of emissions and no error in atmospheric transport. In this experiment, the prior emissions estimate is given by EDGAR and the true $ffCO_2$ signals were generated by scaling the EDGAR emissions in each air basin to match the annually averaged Vulcan emissions in that air basin. These differences range from 0.1 TgC in San Diego to 11.6 TgC in the San Joaquin Valley (Table 1). The EDGAR state total emissions are 12% higher than Vulcan, so the bias in the prior estimate in the state total $ffCO_2$ emissions is +12%. The experiment is run for all the transport models with no temporal variation in emissions. This experiment assesses the performance of the inversion and the

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

strength of the data constraint provided by the observation network in the simplest case where there are regional errors in the magnitude of prior emissions.

2.5.2 Difference in spatial distribution of emissions

To investigate the bias in the posterior emissions estimate that could result from errors in the spatial distribution of prior emissions within each air basin, we now use annually averaged Vulcan emissions as the true emissions and EDGAR emissions scaled in each air basin to match the annually averaged Vulcan emissions in that region as the prior estimate of emissions. In this experiment, the prior estimate of the total emissions in each air basin is unbiased, and we assess how differences in the spatial distribution of emissions between Vulcan and EDGAR in each air basin may lead to a bias in the posterior emissions estimate. As shown in Figure 1c, the most significant discrepancies in spatial distribution are in the major urban areas of Los Angeles and the San Francisco Bay. This experiment is also run for all the transport models using the same transport model for both the true and prior simulation and including no temporal variation in emissions.

2.5.3 Difference in temporal variation of emissions

To assess the impact of temporally-varying emissions on the inversion result, we generated true $ffCO_2$ signals with temporally-invariant annually-averaged Vulcan emissions and prior $ffCO_2$ signals with temporally-varying Vulcan emissions. We scaled the temporally-varying Vulcan emissions in each air basin so that the total $ffCO_2$ emissions were the same magnitude as the total $ffCO_2$ emissions in the annually averaged Vulcan emissions for each field campaign. As shown in Figure 1d, scaling was less than 10% of annual mean emissions with campaigns occurring during maxima and minima of the annual emissions cycle. Here the prior estimate is again

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics Discussions
Open Access

unbiased, and we assess how differences in the diurnal variation of emissions (see Fig 1b) may lead to a bias in the posterior emissions estimate. This experiment is also run for all the transport models using the same transport model for both the true and prior simulation.

2.5.4 Difference in Atmospheric Transport

To test the effect of differences in the simulated atmospheric transport of emissions, the same emissions estimate (annually-averaged Vulcan) is coupled with two different transport models to generate prior and true ffCO_2 signals. This experiment investigates potential effects of transport errors, within the variations in transport across the three models we use. WS-LBL is considered the "true" atmospheric transport while UM-NAME and WS-CTL are used for the prior simulation in individual experiments. Here the prior estimate is again unbiased, and we assess how differences in the modeled atmospheric transport may lead to a bias in the posterior emissions estimate.

3 Results

3.1 Simulated ffCO_2 Observations

Before presenting the results of the inversion experiments, we first examine simulated ffCO_2 contributions from different regions at each of the 9 observation sites. This allows us to quantify which air basins have the largest influence on simulated concentrations at observation sites and better interpret the results of the experiments. Figure 2 shows simulated concentrations at observation sites resulting from emissions in the 6 highest-emitting air basins in California, and from outside California. The highest signals (> 10 ppm) are simulated at urban sites (e.g. CIT and SBC) for emissions from urban air basins (e.g., South Coast, 14.SC). The 9 air basins not shown in Fig. 2 contributed signals below 0.1 ppm due to the small size or low emissions of the air basin (e.g. Lake County and Lake Tahoe), or distance from the observation network (e.g. Northeast Plateau, Great Basin Valleys and Salton Sea). In general, the northern sites (THD to SLT in Fig 2) are sensitive to northern air basins (Sacramento and San Joaquin Valleys and SF Bay), and the southern sites (VTR to SIO) are sensitive to emissions from southern air basins (Mojave Desert, South Coast and San Diego). All transport models show the observation sites are sensitive to more air basins in the Oct-Nov and Jan-Feb campaigns, compared to the May campaign (Fig. 2). Signals simulated by WS-CTL come from fewer air basins than UM-NAME or WS-LBL, particularly in May.

In our simulation experiments, signals from outside California are generally small compared to the total signal for most sites (<10% on average), although they can average 20-50% for STB, STR, SLT and SIO for individual campaigns. For THD, located near the northern border of the state, a larger influence from outside California is found, 10-90%, due to a combination of relatively low local emissions and northerly winds transporting emissions from the northwestern United States and Canada .

3.2.1 Difference in magnitude of emissions

Figure 3 (a) shows the statewide inversion result for the experiment testing the effect of a bias in magnitude in regional emissions in the prior simulation. In this figure, and similar figures that follow for the other experiments, prior estimates are represented by black markers and posterior estimates are represented by colored markers, with the 2-σ uncertainty on the x-axis and the bias relative to the truth on the y-axis. The diagonal lines show 1:1 and 1:-1 lines, so that a marker lying to the right of these lines indicate the posterior bias is smaller than the posterior uncertainty, whereas a marker to the left of these lines indicate the posterior bias is larger than

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

the posterior uncertainty. Filled markers show results using SD prior uncertainty and empty markers show results using 70% prior uncertainty. Prior and posterior uncertainties are expressed as 2-σ.

For all transport models and campaigns, the inversion is able to reduce prior bias and scale posterior emissions towards the truth. The +12% bias in the statewide emissions in the prior was reduced to a posterior bias of between 0 and +9% (mean bias = +5%) for SD prior uncertainty. Using 70% prior uncertainty reduced prior bias to between -3 and +6 (mean = +1%). Statewide posterior uncertainty was 10-14% (mean 12%) and 14-32% (mean = 21%) for SD and 70% prior uncertainty respectively, where uncertainty is expressed as 2-σ, lower than the statewide prior uncertainties of 16% for SD and 69% for 70% prior uncertainty. There were no outliers identified in this experiment.

To determine what is driving the statewide results, we examine the individual air basin inversion results. Figure 3 (b) shows the inversion results for the six main emission regions of California, with San Joaquin Valley (8.SJV) and South Coast (14.SC) having the largest prior biases. For the San Joaquin Valley (8.SJV) and South Coast (14.SC) regions with the largest prior bias, the biases are reduced in most cases, however, only the posterior estimates from the 70% prior uncertainty experiment overlap the true emissions. The posterior estimates for SD prior uncertainty do not overlap with the truth, indicating that the 2-σ prior uncertainty of 24% in South Coast (14.SC), for example, restricts the inversion from eliminating biases of 30% in these regions (Table 1), given the observations available. The 9 air basins omitted from Fig. 3(b) are generally not being scaled by the inversion due to a lack of constraint from the observation network, low emissions, or small prior uncertainty (Figure S1).

The bias in the posterior estimate of statewide emissions is larger in May than in Oct-Nov and Jan-Feb (Fig 3a, triangles). This poorer performance of the inversion in May can be largely attributed to the San Joaquin Valley (8.SJV), where the posterior emissions are largely unchanged from the prior in May. There is no observation site in the San Joaquin Valley, and as shown in Fig. 2, emissions in the San Joaquin Valley do not reach observation sites in neighboring air basins in May, but they do reach these sites in Oct-Nov and Jan-Feb. In contrast, the South Coast (14.SC) influences the two observation sites, CIT and SBC, located in the region as well as several other sites (Fig. 2). Both CIT and SBC show prior signals are too high compared to true signals for all campaigns and models (Fig. 3c), reflecting the positive bias in prior emissions in the South Coast region, which is reduced in the posterior.

3.2.2 Difference in spatial distribution of emissions

The statewide inversion results for the experiment including errors in the spatial distribution of emissions are shown in Figure 4 (a). In this case the magnitude of prior emissions in each air basin is equal to true emissions and we aim to quantify how errors in the spatial distribution of emissions (EDGAR as prior and Vulcan as true distribution) lead to bias in posterior emissions estimates. Posterior emissions are negatively biased, apart from WS-LBL in January-February. Posterior bias was between -10% and +1% (mean -4%) for SD prior uncertainty and between -10% and +4% (mean = -4%) for 70% prior uncertainty across transport models and campaigns. As might be expected from the experimental setup with an unbiased prior, posterior emissions estimates generated using SD prior uncertainty have a smaller mean bias and smaller range of posterior estimates compared to those generated using 70% prior uncertainty. Statewide uncertainty was reduced from 16% to 10-14% (mean = 12%) for SD prior uncertainty and from 58% to 14-21% (mean = 18%) for 70% air basin prior uncertainty. Biases induced are smaller

Atmospheric Chemistry and Physics Discussions
Open Access

than the 2-σ posterior uncertainty across all transport models, campaigns and choice of prior uncertainty.

Posterior emissions results in the two largest emitting air basins (the San Francisco Bay and South Coast) are also negatively biased in most cases (Fig 4b). In several cases, posterior biases are larger than the associated posterior uncertainties, for example in the South Coast for WS-LBL in all cases. Considering Figure 4 (c), prior $ffCO_2$ signals are being overestimated more often than underestimated, particularly for the relatively more urban sites CIT and SLT. Sacramento Valley (3.SV) and the San Joaquin Valley (8.SJV) have higher posterior emissions in WS-LBL in most cases, possibly due to the inversion compensating for reduced posterior emissions in the San Francisco Bay (13.SFB) and South Coast (14.SC).

Since the prior emissions from EDGAR have been scaled to have the same total as Vulcan (the true emissions) in each region, the pattern of more negative posterior emissions is only caused by the sub-regional spatial distribution of emissions. Comparing Vulcan and EDGAR native grid cell emissions in Figures 1c and S2, EDGAR tends to have greater emissions in high-emission grid cells. In other words, the emissions are more concentrated in EDGAR and more dispersed in Vulcan. This pattern explains the negative bias in posterior emissions for the urban South Coast air basin. The opposite effect does not appear to hold for rural observation sites and regions, perhaps because rural emissions are already rather dispersed and have less of an influence on the observations.

In these experiments, 0-3% of observations were identified as outliers, but excluding outliers did not change the statewide result significantly (<1% change in mean bias).

3.2.3 Difference in temporal variation of emissions

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Figure 5 (a) shows the statewide inversion result for the experiment where the emissions are Vulcan temporally-varying in the prior simulation (see Fig. 1b) but Vulcan temporally-invariant in the true simulation. Posterior bias was between -13 and +5% (mean = -3%) for SD uncertainty and between -15% and +6% (mean = -3%) for 70% prior uncertainty. Posterior uncertainty was 11-15% (mean = 12%) for SD prior uncertainty and 15-24% (mean = 18%) in posterior emissions for SD (70%) prior uncertainty. Outlier removal resulted in 0-1% (mean = 0%) of data points being removed, which did not change the statewide results.

The posterior estimate for WS-LBL in May with SD prior uncertainty has a significant negative bias of -13%, approximately the same magnitude as the associated 2-σ posterior uncertainty. As can be seen by the air basin results of Figure 5 (b), the statewide bias for WS-LBL in May is being driven by a large regional bias in the South Coast, but also in the San Francisco Bay and San Diego air basins. These regional biases are larger than their associated posterior uncertainties. Figure 5 (c) shows the prior ffCO_2 signals at CIT average ~7ppm too high in May for WS-LBL. In contrast, prior ffCO_2 signals at CIT and SBC are too low in Oct-Nov for WS-CTL, leading to a high bias in posterior emissions from the South Coast. San Diego also exhibited both high and low biases in the posterior emissions. Overall, temporal variations in emissions led to posterior biases generally within ±6%, but as large as 15%; however, a consistent pattern in the posterior bias due to the temporal representation in emissions was not found.

3.2.4 Difference in Atmospheric Transport

The statewide inversion results for the experiment where the atmospheric transport in the prior simulation uses WS-CTL or UM-NAME but the atmospheric transport in the true simulation

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics Discussions

uses WS-LBL are shown in Figure 6 (a). Outliers were identified in these experiments and we present results for inversions including all data and for inversions where outliers were removed.

When all data are included, differences in atmospheric transport model introduces a bias in statewide posterior emissions of between -42% and -3% (mean = -12%) for SD prior uncertainty and between -32% and 0% (mean = -15%) for 70% prior uncertainty. For one case, using WS-CTL to generate prior signals in October-November, the bias in the posterior emissions estimate was larger than the 2-σ uncertainty for both SD and 70% prior uncertainty.

Removing outliers significantly improved the inversion results (Figure 6 b): the mean bias was between -10% and 0% (mean = -3%) for SD prior uncertainty and between -9% and +6% (mean = -5%) for 70% prior uncertainty when outliers were removed. Posterior uncertainty was 9-15% (mean = 12%) and 15-24% (mean = 18%) for SD and 70% prior uncertainty respectively, with all posterior estimates within 2-σ of the true statewide emissions. The reduction in posterior bias when outliers are removed is mostly due to the removal of a few large positive outliers in prior simulated signals by WS-CTL (Figure 7). Figure 7 illustrates the time series of simulated $ffCO_2$ in each model with outliers shown as an x. Outliers removed were between 6.9% and 20.6% of all observations (mean = 10.5%). This is similar to the fraction of outliers identified in Graven et al. 2018 using the same method with real data (~8%). It is also similar to that of Jeong et al., 2012a and b (0-27%) for monthly inversions of CH_4 in California using a different method of identifying outliers where model-data residuals are larger than 3-σ of model-data uncertainty.

While the statewide posterior emissions estimate is significantly biased in only one case (WS-CTL in Oct-Nov) when outliers are not removed, the posterior emissions estimates for the main emissions regions are significantly biased in several cases (Fig 6c). The largest bias is in the

South Coast region where posterior estimates are biased by more than -75% (with 1% posterior uncertainty) in Oct-Nov when using WS-CTL to generate prior signals. This large posterior emissions bias in the South Coast and the statewide total can be attributed to overestimates in prior $ffCO_2$ signal of more than 6ppm on average at CIT and SBC and more than 2ppm at WGC and STR (Fig. 6e) due to some high outliers in the WS-CTL simulations (Fig. 7). Posterior estimates for San Francisco Bay, South Coast and San Diego were also significantly biased in some other cases, particularly for 70% prior uncertainty but also for SD prior uncertainty. This indicates that regional biases caused by differences in atmospheric transport appear to compensate over the statewide scale, and that results for individual regions are less robust than aggregate results for the statewide network. It also suggests that a dense observation network is beneficial to reducing the impact of uncertainty in atmospheric transport.

To investigate the differences in simulated $ffCO_2$ and assess whether these could be attributed to specific aspects of modelled meteorology, we compared PBLH and wind speed in WS-LBL and the UM for 5 of the 9 observation sites where PBLH output was available. PBLH was not available for WS-CTL. Estimates for PBLH in WS-LBL are based on the Mellor–Yamada–Nakanishi–Niino version 2 (MYNN2) parameterization scheme that estimates PBLH using localized turbulence kinetic energy closure parameterization (Nakanishi and Niino 2004, 2006). Estimates of PBLH are calculated internally within the UM. PBLH and wind speed were averaged over 6 hours from 12 to 6pm Pacific Standard Time to compare the afternoon means (Seibert et al., 2000). We found no consistent correlation between differences in PBLH or wind speed and differences in simulated $ffCO_2$ between models across sites and campaigns (Figure S3). Absolute values of wind direction and $ffCO_2$ did not show consistent correlations either. The lack of correlation suggests we cannot attribute differences in simulated $ffCO_2$ to any single meteorological variable estimated at any individual station in the transport models.

We also examined if differences in simulated $ffCO_2$ signals across transport models could be explained by the differences in spatial resolution of the models. WS-CTL footprints were re-gridded from a 0.1° native grid to the coarser UM-NAME grid of 17 or 25km and then used to simulate $ffCO_2$. For this comparison, we simulated $ffCO_2$ every day over the campaign period. We found no consistent reduction in mean $ffCO_2$ bias between sites over the 2 campaigns, however there is a reduction in spread of bias at 4 sites for both campaigns (WGC, SLT, SBC and SIO), suggesting model resolution could potentially have an impact for these sites. In general however, we cannot say that transport model resolution error in atmospheric transport is a key driver of $ffCO_2$ signal bias across observation sites (Figure S4).

4 Discussion

Our results show that atmospheric inversions can reduce a hypothetical bias in the magnitude of prior $ffCO_2$ emissions estimates for the U.S. state of California using the ground-based observation network from Graven et al. (2018), under the idealized assumptions of perfect atmospheric transport and perfect spatio-temporal distribution of emissions in the prior estimate. By exploring differences in model transport and spatio-temporal distribution of prior emissions, we found that biases of magnitude 1-15% in monthly posterior estimates of statewide emissions can result from differences in the temporal variation, spatial distribution and modelled transport of the prior simulation. However, these biases were less than the 2-σ posterior uncertainty in state-total emissions, when outliers were removed. In some cases, the biases in posterior emissions for individual air basins were significant, compared to the posterior uncertainties,

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics Discussions

suggesting that estimates for individual regions are less reliable than the aggregate estimates of the state-wide total.

The largest bias in statewide posterior estimates was found to be caused by errors in the temporal variation in emissions. This highlights the necessity for temporally-varying emissions to be estimated and included in prior emissions estimates, particularly for urban regions. Similar results have been found in other regions including Indianapolis (Turnbull et al. 2015) and Europe (Peylin et al. 2011), and more generally, for high-emission regions around the globe (Zhang et al. 2016). Although the afternoon sampling is near to the diurnal maximum in emissions in California (Fig. 1c, Gurney et al. 2009), which might be expected to lead to higher simulated $ffCO_2$ in temporally-varying vs temporally-invariant emissions, we did not find consistently positive biases in $ffCO_2$ but rather both positive and negative biases. This suggests the overall impact of temporally-varying emissions depends on the model transport and the characteristics of the observation site. Furthermore, uncertainties in the temporal distribution of emissions at an hourly resolution have not yet been fully quantified (Nassar et al., 2013).

Errors in model transport, as represented in our experiments by using different transport models, were shown to bias posterior $ffCO_2$ emissions by 10% or less, when outliers were removed. These biases related to transport error are somewhat lower than estimated by similar simulation experiments for $ffCO_2$ emissions estimates for the U.S. by Basu et al. (2016) using different transport models (>10%), although their spatial scale was larger and the alternate model they used was intentionally biased. In contrast, the three models we use are all actively applied in regional greenhouse gas inversions. Our results are comparable to the estimate of ±15% uncertainty in atmospheric transport in WS-LBL using comparisons with atmospheric observations of CO in California (Bagley et al. 2017).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

The fraction of pseudo-observations we identified as outliers in these transport error experiments (10.5%, range 6.9-20.6%), was similar to Graven et al., 2018, where 8% of all observations were removed as outliers using the same method. The outliers in our experiments were primarily high $ffCO_2$ signals simulated by WS-CTL in Oct-Nov. When included in the inversion, these did lead to significant biases in the posterior estimates for the experiment on model transport. This highlights the need for careful examination of simulated $ffCO_2$ and consideration of outliers in atmospheric $ffCO_2$ inversions.

Attributing differences in simulated $ffCO_2$ between the different transport models to specific meteorological variables proved inconclusive, and model resolution error did not appear to explain the differences in simulated signals, although we were not able to investigate aggregation error in comparison to the high-resolution WS-LBL model. Wang et al. (2017) found aggregation error to be only a minor contributor to errors in simulated $ffCO_2$ in Europe, while Feng et al., (2016) found that high-resolution gridded emissions estimates could be more important than high resolution transport models for simulations of greenhouse gases in Greater Los Angeles. We found that differences in the spatial representation of prior emissions in EDGAR compared to Vulcan led to consistently lower, although not significantly different, posterior state-wide estimates due to the emissions in EDGAR being more concentrated in urban regions. The spatial allocation of emissions between urban and rural regions in gridded emissions estimates have much larger uncertainties than national totals (Hogue et al. 2016), suggesting that several different gridded emissions estimates should be used in regional $ffCO_2$ inversions to capture this source of uncertainty.

The results of these experiments suggest that the choice of prior emissions estimate and transport model (among those considered here and currently used in the community) used in our $ffCO_2$

inversion would result in differences of 15% or less in posterior state-wide $ffCO_2$ emissions in California, using the observation network from Graven et al. (2018). These differences are generally not significant, compared to the posterior 2-σ uncertainties of 10 to 15%. In comparison, Graven et al. (2018) found that posterior state-wide $ffCO_2$ emissions were not statistically different when using temporally-varying emissions from Vulcan, as compared to annual mean emissions from Vulcan or EDGAR, with posterior uncertainties of ±15 to ±17%. Our results may be specific to the California region, observation network and inversion setup we consider here, but we expect that similar differences of 1-15% are likely to be found elsewhere in similar inversions at comparable regional scales.

In our results, emissions from many small or rural air basins did not have a significant contribution to the local enhancement of $ffCO_2$ at the observation sites and were not adjusted by the inversion in most cases (Figure 2, Figure S1). Combined with our experimental setup specifying the magnitude of prior emissions to be equal to true emissions, it might be expected that our results could underestimate the predicted biases in posterior emissions. However, these experiments were designed specifically to quantify representation and transport error using the inversion setup and the observation network from Graven et al. (2018) as a test case. Fischer et al. (2017), showed in individual simulation experiments that using either EDGAR or a spatially uniform flux of 1 μ mol m^{-2} s^{-1} as a biased prior produced posterior emissions that are substantially closer to true emissions, but only if the prior uncertainties are set large enough to encompass biases in prior emissions. Further experiments using a different experimental setup such as choice of mismatch error or specification of inversion regions (e.g. to change the inversion region size based on proximity to the observation network, Manning et al., 2011), would help to characterize uncertainties in regional $ffCO_2$ inversions and the robustness of posterior estimates to the choices made in the inversion setup.

Conclusion

We have shown that atmospheric inversions for the U.S. state of California can reduce a hypothetical bias in the magnitude of prior emissions estimates of $ffCO_2$ in California using the ground-based observation network from Graven et al. (2018). Experiments to characterize the effect of differences in the spatial and temporal distribution in prior emissions resulted in biases in posterior state-total emissions with magnitudes of 1-15%, similar to monthly posterior estimates of Basu et al., 2016 for the western United States. Our results highlight the need for (1) temporal variation to be included in prior emissions, (2) different estimates of the spatial distribution of emissions between urban and rural regions to be considered, and (3) representation of atmospheric transport in regional $ffCO_2$ inversions to be further evaluated.

Acknowledgments

This project was funded by the Natural Environment Research Council (NERC, UK), the UK Met Office, NASA Carbon Monitoring System (NNX13AP33G and NNH13AW56I), and the European Commission through a Marie Curie Career Integration Grant. The authors thank A. Andrews and the CarbonTracker-Lagrange team for providing footprints. Support for CarbonTracker-Lagrange has been provided by the NOAA Climate Program Office's Atmospheric Chemistry, Carbon Cycle, & Climate (AC4) Program and the NASA Carbon Monitoring System.

References

R. J. Andres, T. A. Boden, F. M. Bréon, P. Ciais, S. Davis, D. Erickson, J. S. Gregg, A. Jacobson, G. Marland, J. Miller, T. Oda, J. G J Olivier, M. R. Raupach, P. Rayner, and K. Treanton. A synthesis of carbon dioxide emissions from fossil-fuel combustion. Biogeosciences, 9(5):1845–1871, 2012. ISSN 17264170. doi: 10.5194/bg-9-1845-2012.

S. Asefi-Najafabady, P. J. Rayner, K. R. Gurney, A. McRobert, Y. Song, K. Coltin, J. Huang, C. Elvidge, and K. Baugh. A multiyear, global gridded fossil fuel CO 2 emission data product: Evaluation and analysis of results. Journal of Geophysical Research: Atmospheres, 119(17):10,213–10,231, 2014.

Sourish Basu, John Bharat Miller, and Scott Lehman. Separation of biospheric and fossil fuel fluxes of CO2 by atmospheric inversion of CO2 and 14CO2 measurements: Observation System Simulations. Atmospheric Chemistry and Physics, 16(9):5665–5683, 2016. ISSN 16807324. doi: 10.5194/acp-16-5665-2016.

Tom Boden, Bob Andres, and Gregg Marland. Ranking of the world's countries by 2013 total CO2 emissions from fossil fuel burning, cement production, and gas flaring. Technical report, 2016. URL: cdiac.ornl.gov/trends/emis/top2013.tot.

J. Brioude, W. M. Angevine, R. Ahmadov, S. W. Kim, S. Evan, S. a. McKeen, E. Y. Hsie, G. J. Frost, J. a. Neuman, I. B. Pollack, J. Peischl, T. B. Ryerson, J. Holloway, S. S. Brown, J. B. Nowak, J. M. Roberts, S. C. Wofsy, G. W. Santoni, T. Oda, and M. Trainer. Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: Assessing anthropogenic emissions of CO, NOx and CO2 and their impacts. Atmospheric Chemistry and Physics, 13(7):3661–3677, 2013.

CARB, 2017. California Greenhouse Gas Emission Inventory - 2017 Edition, California Air Resources Board. Available at http://www.arb.ca.gov/cc/inventory/data/data.htm

EDGAR. EDGAR Greenhouse Gas Emissions Inventory v4.2 FT2010, 2011. URL http://edgar.jrc.ec.europa.eu/index.php.

Marc L. Fischer, Nicholas Parazoo, Kieran Brophy, Xinguang Cui, Seongeun Jeong, Junjie Liu, Ralph Keeling, Thomas E. Taylor, Kevin Gurney, Tomohiro Oda, and Heather Graven. Simulating estimation of California fossil fuel and biosphere carbon dioxide exchanges combining in situ tower and satellite column observations. Journal of Geophysical Research: Atmospheres, 122(6):3653–3671, 2017. ISSN 21698996. doi: 10.1002/2016JD025617.

Mathias Göckede, Anna M. Michalak, Dean Vickers, David P. Turner, and Beverly E. Law. Atmospheric inverse modeling to constrain regional-scale CO2 budgets at high spatial and temporal resolution. Journal of Geophysical Research: Atmospheres, 115(15), 2010.

Graven, H, ML Fischer, T Lueker, S Jeong, TP Guilderson, RF Keeling, R Bambha, K Brophy, W Callahan, X Cui, C Frankenberg, KR Gurney, BW LaFranchi, SJ Lehman, H Michelsen, JB Miller, S Newman, W Paplawsky, NC Parazoo, C Sloop, SJ Walker, 2018. Assessment of Fossil Fuel CO_2 Emissions in California Using Atmospheric Observations and Models, *Environmental Research Letters*, in review.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Heather D Graven, Thomas P Guilderson, Ralph, and F Keeling. METHODS FOR HIGHPRECISION 14 C AMS MEASUREMENT OF ATMOSPHERIC CO 2 AT LLNL. RADIOCARBON, 49(2):349–356, 2007.

Heather D. Graven, Thomas P. Guilderson, and Ralph F. Keeling. Observations of radiocarbon in CO2 at seven global sampling sites in the Scripps flask network: Analysis of spatial gradients and seasonal cycles. Journal of Geophysical Research Atmospheres, 117(2):n/a–n/a, January 2012. ISSN 01480227. doi: 10.1029/2011JD016535. URL http://doi.wiley.com/10.1029/2011JD016535.

Kevin R. Gurney, Daniel L. Mendoza, Yuyu Zhou, Marc L. Fischer, Chris C. Miller, Sarath Geethakumar, and Stephane De La Rue Du Can. High resolution fossil fuel combustion CO2 emission fluxes for the United States. Environmental Science and Technology, 43(14):5535–5541, 2009. 14

Kevin Robert Gurney, Rachel M. Law, a. Scott Denning, Peter J. Rayner, David Baker, Philippe Bousquet, Lori Bruhwiler, Yu Han Chen, Philippe Ciais, Songmiao Fan, Inez Y. Fung, Manuel Gloor, Martin Heimann, Kaz Higuchi, Jasmin John, Eva Kowalczyk, Takashi Maki, Shamil Maksyutov, Philippe Peylin, Michael Prather, Bernard C. Pak, Jorge Sarmiento, Shoichi Taguchi, Taro Takahashi, and Chiu Wai Yuen. TransCom 3 CO2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. Tellus, Series B: Chemical and Physical Meteorology, 55(2): 555–579, 2003.

Susannah Hogue, Eric Marland, Robert J Andres, Gregg Marland, and Dawn Woodard. Uncertainty in gridded CO2 emissions estimates. Earth's Future, pages 1–15, 2016. ISSN 23284277. doi: 10.1002/2015EF000343. URL http://doi.wiley.com/10.1002/2015EF000343.

IPCC. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Technical report, 2014.

Daniel J Jacob. LECTURES ON INVERSE MODELING. 2007.

Seongeun Jeong, Ying Kuang Hsu, Arlyn E. Andrews, Laura Bianco, Patrick Vaca, James M. Wilczak, and Marc L. Fischer. A multitower measurement network estimate of California's methane emissions. Journal of Geophysical Research Atmospheres, 118(19):11339–11351, October 2013. ISSN 21698996. doi: 10.1002/jgrd.50854. URL http://doi.wiley.com/10.1002/jgrd.50854.

Charles D Keeling, Stephen C Piper, Robert B Bacastow, Martin Wahlen, Timothy P Whorf, Martin Heimann, and Harro A Meijer. Atmospheric CO2 and 13CO2 Exchange with the Terrestrial Biosphere and Oceans from 1978 to 2000: Observations and Carbon Cycle Implications. A History of Atmospheric CO2 and Its Effects on Plants, Animals, and Ecosystems, 177:83–113, 2005. doi: 10.1007/b138533. URL http://www.springerlink.com/index/10.1007/b138533.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Ingeborg Levin. A novel approach for independent budgeting of fossil fuel CO2 over Europe by 14CO2 observations. Geophysical Research Letters, 30(23):49–60, 2003. J. C. Lin. A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. Journal of Geophysical Research, 108 (D16):ACH 2–1–ACH 2–17, 2003.

A. J. Manning, S. O'Doherty, A. R. Jones, P. G. Simmonds, and R. G. Derwent. Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach. Journal of Geophysical Research: Atmospheres, 116(2):1–19, 2011.

Thomas Nehrkorn, Janusz Eluszkiewicz, Steven C. Wofsy, John C. Lin, Christoph Gerbig Marcos Longo, and Saulo Freitas. Coupled weather research and forecasting-stochastic time-inverted lagrangian transport (WRF-STILT) model. Meteorology and Atmospheric Physics, 107(1):51–64, 2010. ISSN 01777971. doi: 10.1007/s00703-010-0068-x.

S. Newman, S. Jeong, M. L. Fischer, X. Xu, C. L. Haman, B. Lefer, S. Alvarez, B. Rappenglueck, E. A. Kort, A. E. Andrews, J. Peischl, K. R. Gurney, C. E. Miller, and Y. L. Yung. Diurnal tracking of anthropogenic CO2 emissions in the Los Angeles basin megacity during spring 2010. Atmospheric Chemistry and Physics, 13(8):4359–4372, 2013. ISSN 16807324. doi: 10.5194/acp-13-4359-2013.

T Oda and S Maksyutov. A very high-resolution (1km??1 km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of nighttime lights. Atmospheric Chemistry and Physics, 11(2):543–556, 2011

P. Peylin, S. Houweling, M. C. Krol, U. Karstens, C. Rödenbeck, C. Geels, A. Vermeulen, B. Badawy, C. Aulagnier, T. Pregger, F. Delage, G. Pieterse, P. Ciais, and M. Heimann. Importance of fossil fuel emission uncertainties over Europe for CO2 modeling: Model intercomparison. Atmospheric Chemistry and Physics, 11(13):6607–6622, 2011.

Peylin P, Law RM, Gurney KR, Chevallier F, Jacobson AR, Maki T, Niwa Y, Patra PK, Peters W, Rayner PJ, Rödenbeck C. Global atmospheric carbon budget: results from an ensemble of atmospheric CO2 inversions. Biogeosciences. 2013 Oct 24;10:6699-720.

Clive D Rodgers. Inverse Methods for Atmospheric Sounding. WORLD SCIENTIFIC, July 2000. ISBN 978-981-02-2740-1. doi: 10.1142/3171. URL http://www.worldscientific.com/worldscibooks/10.1142/3171.

J. C. Turnbull, A. Karion, M. L. Fischer, I. Faloona, T. Guilderson, S. J. Lehman, B. R. Miller, J. B. Miller, S. Montzka, T. Sherwood, S. Saripalli, C. Sweeney, and P. P. Tans. Assessment of fossil fuel carbon dioxide and other anthropogenic trace gas emissions from airborne measurements over Sacramento, California in spring 2009. Atmospheric Chemistry and Physics, 11(2):705–721, 2011.

C. E. Yver, H. D. Graven, D. D. Lucas, P. J. Cameron-Smith, R. F. Keeling, and R. F. Weiss. Evaluating transport in the WRF model along the California coast. Atmospheric Chemistry and Physics, 13(4): 1837–1852, February 2013.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics Discussions

Nassar, R., Sioris, C. E., Jones, D. B. A., and McConnell, J. C.: Satellite observations of CO2 from a highly elliptical orbit for studies of the Arctic and boreal carbon cycle, J. Geophys. Res.- Atmos., 119, 2654–2673, doi:10.1002/2013JD020337, 2014.

Nassar R, Napier-Linton L, Gurney KR, Andres RJ, Oda T, Vogel FR, Deng F. Improving the temporal and spatial distribution of CO2 emissions from global fossil fuel emission data sets. Journal of Geophysical Research: Atmospheres. 2013 Jan 16;118(2):917-33.

Hungershoefer, K., Breon, F.-M., Peylin, P., Chevallier, F., Rayner, P., Klonecki, A., Houweling, S., and Marshall, J.: Evaluation of various observing systems for the global monitoring of CO2 surface fluxes, Atmos. Chem. Phys., 10, 10503–10520, doi:10.5194/acp-10-10503-2010, 2010.

Chevallier, F., Maksyutov, S., Bousquet, P., Bréon, F.-M., Saito, R., Yoshida, Y., and Yokota, T.: On the accuracy of the CO2 surface fluxes to be estimated from the GOSAT observations, Geophys. Res. Lett., 36, L19807, doi:10.1029/2009GL040108, 2009.

Gerbig, C., Lin, J.C., Wofsy, S.C., Daube, B.C., Andrews, A.E., Stephens, B.B., Bakwin, P.S. and Grainger, C.A., 2003. 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, 108(D24).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics
Open Access
Discussions

Marland G, Brenkert A, Olivier J. CO2 from fossil fuel burning: a comparison of ORNL and EDGAR estimates of national emissions. Environmental Science & Policy. 1999 Jun 1;2(3):265-73.

Grell GA, Dudhia J, Stauffer DR. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5).

Nakanishi M, Niino H. An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. Boundary-Layer Meteorology. 2004 Jul 1;112(1):1-31.

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 May 1;119(2):397-407.

Hong SY, Noh Y, Dudhia J. A new vertical diffusion package with an explicit treatment of entrainment processes. Monthly weather review. 2006 Sep;134(9):2318-41.

Feng S, Lauvaux T, Newman S, Rao P, Ahmadov R, Deng A, Díaz-Isaac LI, Duren RM, Fischer ML, Gerbig C, Gurney KR. Los Angeles megacity: a high-resolution land–atmosphere modelling system for urban CO 2 emissions. Atmospheric Chemistry and Physics. 2016 Jul 22;16(14):9019-45.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Turnbull JC, Sweeney C, Karion A, Newberger T, Lehman SJ, Tans PP, Davis KJ, Lauvaux T, Miles NL, Richardson SJ, Cambaliza MO. Toward quantification and source sector identification of fossil fuel CO2 emissions from an urban area: Results from the INFLUX experiment. Journal of Geophysical Research: Atmospheres. 2015 Jan 16;120(1):292-312.

Seibert P, Beyrich F, Gryning SE, Joffre S, Rasmussen A, Tercier P. Review and intercomparison of operational methods for the determination of the mixing height. Atmospheric environment. 2000 Jan 1;34(7):1001-27.

Jones A., Thomson D., Hort M., Devenish B. (2007) The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III. In: Borrego C., Norman AL. (eds) Air Pollution Modeling and Its Application XVII. Springer, Boston, MA

Chen, F., and J. Dudhia (2001), Coupling an advanced land surface hydrology model with the Penn State NCAR MM5 modeling system. Part 1:Model implementation and sensitivity, Mon. Weather Rev., 129, 569–585.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics
Discussions

Air Basin	Name	Code	Vulcan (TgC/yr)	EDGAR (TgC/yr)	SD Prior Unc. 1-σ (%)	Vulcan - EDGAR (TgC/yr)
1	North Coast	1.NC	1.0	1.6	59	-0.6
2	Northeast Plateau	2.NP	0.4	1.3	96	-1.0
3	Sacramento Valley	3.SV	6.8	7.4	8	-0.7
4	Mountain Counties	4.MC	2.2	2.0	51	0.1
5	Lake County	5.LC	0.1	0.2	65	-0.2
6	Lake Tahoe	6.LT	0.1	0.1	42	0
7	Great Basin Valleys	7.GBV	0.2	0.6	100	-0.4
8	San Joaquin Valley	8.SJV	8.6	20.2	35	-11.6
9	North Central Coast	9.NCC	6.0	2.2	71	3.8
10	Mojave Desert	10.MD	6.1	4.3	17	1.8
11	South Central Coast	11.SCC	4.4	3.4	21	1.0
12	Salton Sea	12.SS	1.4	1.7	55	-0.3
13	San Francisco Bay	13.SFB	16.4	17.5	22	-1.2
14	South Coast	14.SC	26.9	35.5	12	-8.6
15	San Diego	15.SD	6.6	6.5	10	0.1
Total California			**89.6**	**104.7**	**8**	**-17.8**

Table 1: The 15 air basins of California with respective emissions as estimated by Vulcan and EDGAR. Also shown are the SD prior uncertainty estimate (Fischer et al., 2017), and difference in magnitude between Vulcan and EDGAR for each air basin. Air basin numbers correspond to those marked in Figure 1.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics
Discussions

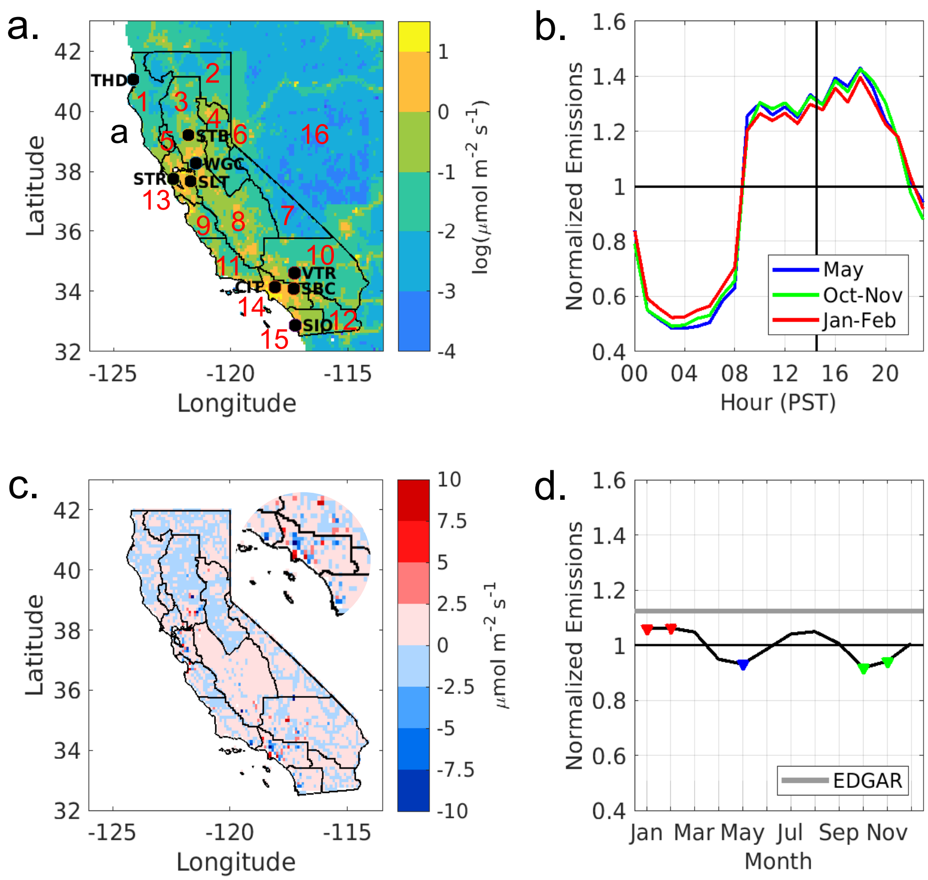

Figure 1: **a.** The location of the 9 tower sites in the observation network (marked with black circles): Trinidad Head (THD), Sutter Buttes (STB), Walnut Grove (WGC), Sutro (STR), Sandia-Livermore (LVR), Victorville (VTR), San Bernardino (SBC), Caltech (CIT) and Scripps Institute of Oceanography (SIO). The 15 air basins are marked out with black lines with region 16 representing emission from outside California within the model domain. Underlayed is a map of annual mean $ffCO_2$ emissions from the Vulcan v2.2 emission map within the United States and EDGAR v4.2 (FT2010) for emission from outside the U.S. **b.** Vulcan diurnal emissions normalized to campaign averaged emissions for the 3 campaigns, **c.** Scaled EDGAR subtracted from Vulcan emissions map, where EDGAR has been scaled to have the same air basin total emissions. The inset shows an enlarged view of southwestern California. **d.** Average monthly emissions normalized to Vulcan annual emissions. Shown in both **b** and **d** is EDGAR annual invariant emissions (grey).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Transport Model	Meteorology	Domain	Model Resolution			References
			Horizontal	Vertical (nLevels / Max Height)	Temporal	
WS-LBL	WRF (v3.5.1)	North America	1km, 4km, 12 km ,36km	50 / 16 km	1 hour	(Lin, 2003; Nehrkorn et al., 2010);
WS-CTL	WRF (v2.1.2)	North America	0.1°, 1°	29 / 25 km	1 hour	(Carbon Tracker, 2017)
UM-NAME	Unified Model	Global	17km, 25km	59 / 29 km	3 hours	(Ryall et al., 1998)

Table 2: Comparison of the three atmospheric transport models used in this study.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

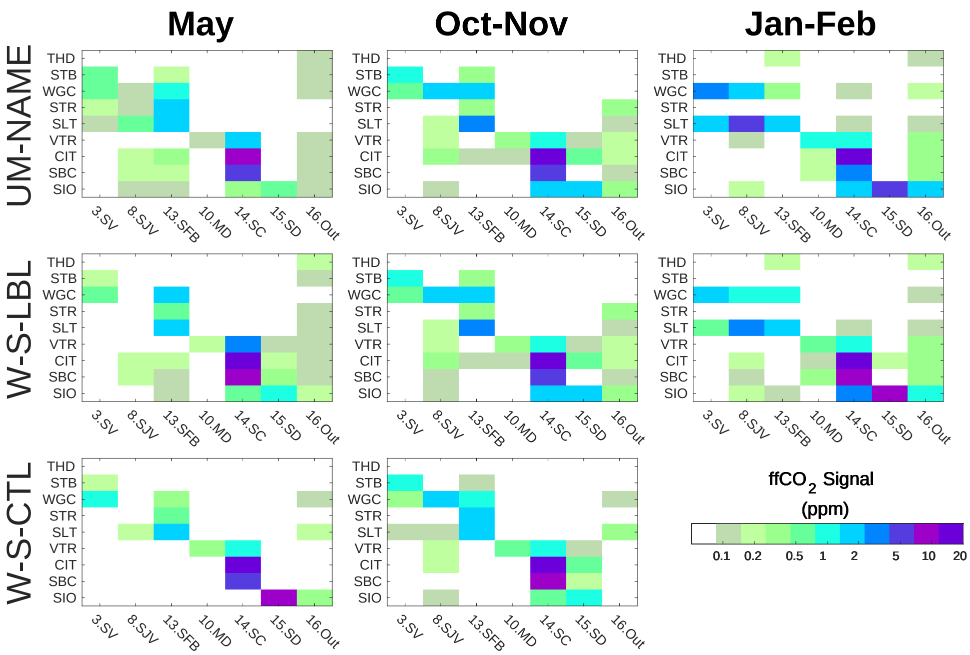

Figure 2: The average $ffCO_2$ signal (ppm) simulated by each atmospheric transport model as a result of emissions from the 6 largest emitting air basins and one outside California region at each observation site over the three measurement campaigns. Signals were simulated based on the EDGAR emission map.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics
Open Access
Discussions

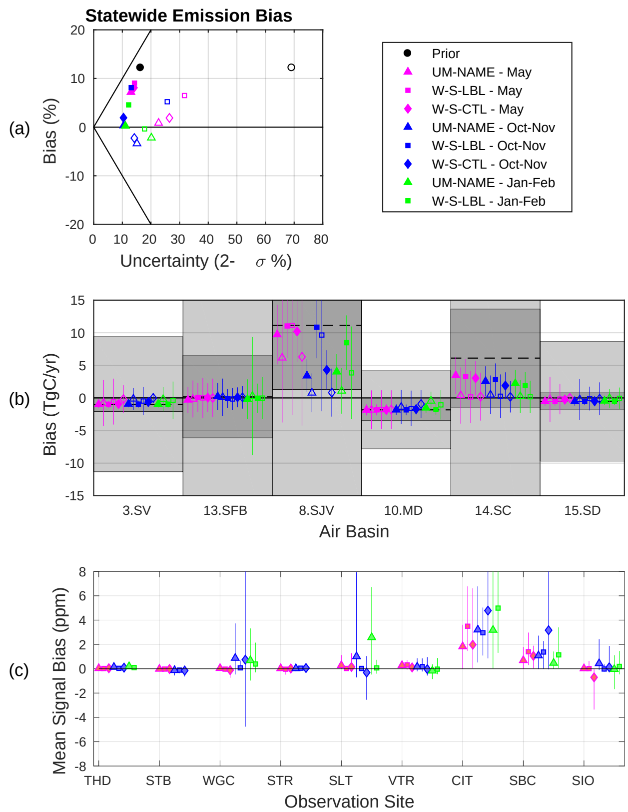

Figure 3: (a) Statewide and (b) individual air basin inversion results for an error in the magnitude of prior emissions. Prior emissions are given by EDGAR and true emissions are given by EDGAR scaled to Vulcan total in each air basin. Air basin results are shown for Sacramento Valley (3.SV), San Francisco Bay (13.SFB), San Joaquin Valley (8.SJV), Mojave Desert (10.MD), South Coast (14.SC) and San Diego (15.SD). Prior results are presented by black markers and posterior results are represented by colored markers. Filled markers show results using SD prior uncertainty and empty markers show results using 70% prior uncertainty. The prior bias in each air basin is given by the dashed lines in (b) with SD prior uncertainty (dark grey) and 70% prior uncertainty (light grey). Prior and posterior uncertainties are expressed as 2-σ. The bottom plot (c) shows the mean signal error in simulated average ffCO_2 concentration. Mean signal error is calculated by subtracting the average true signal from the average prior signal. Error lines are drawn between the maximum and minimum signal bias per campaign.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

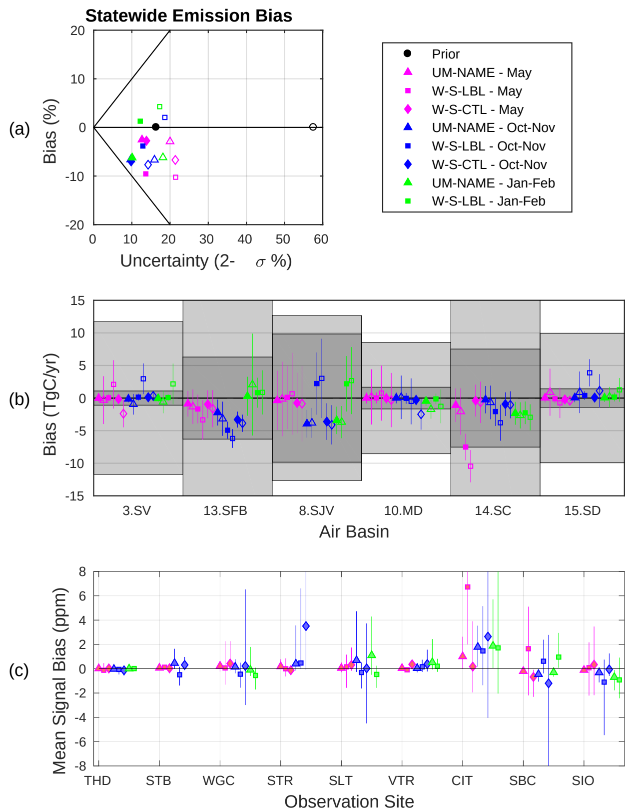

Figure 4: (a) Statewide and (b) individual air basin inversion results for an error in the spatial distribution of prior emissions. Prior emissions are given by EDGAR scaled to Vulcan emissions totals in each air basin and true emissions are given by Vulcan. The bottom plot (c) shows the mean signal error in simulated average $ffCO_2$ concentration.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

Atmospheric Chemistry and Physics
Open Access
Discussions

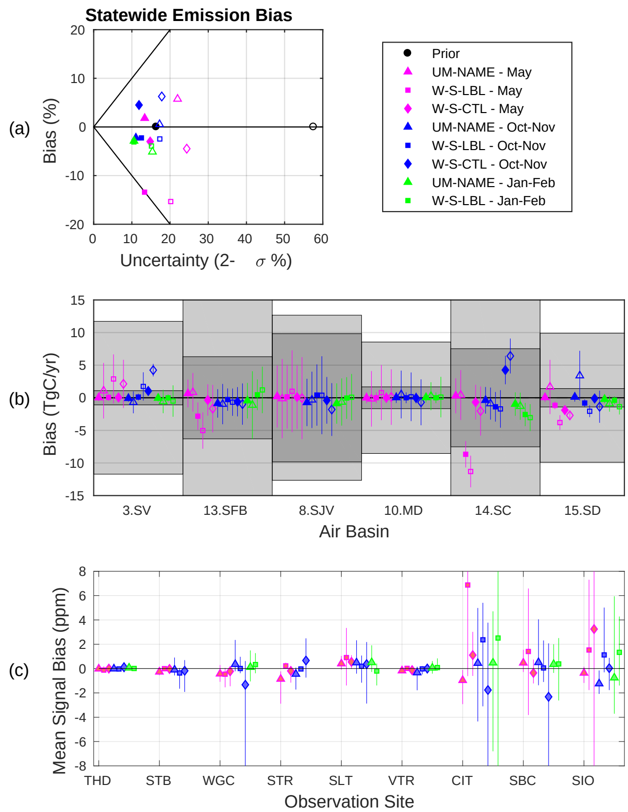

Figure 5: (a) Statewide and (b) individual air basin inversion results for an error in the temporal distribution of prior emissions. Prior emissions are given by temporally varying Vulcan and true emissions are given by annually averaged Vulcan. Prior emissions were scaled to be the equal in magnitude to annually averaged Vulcan emissions. The bottom plot (c) shows the mean signal error in simulated average ffCO$_2$ concentration.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-473
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 23 May 2018

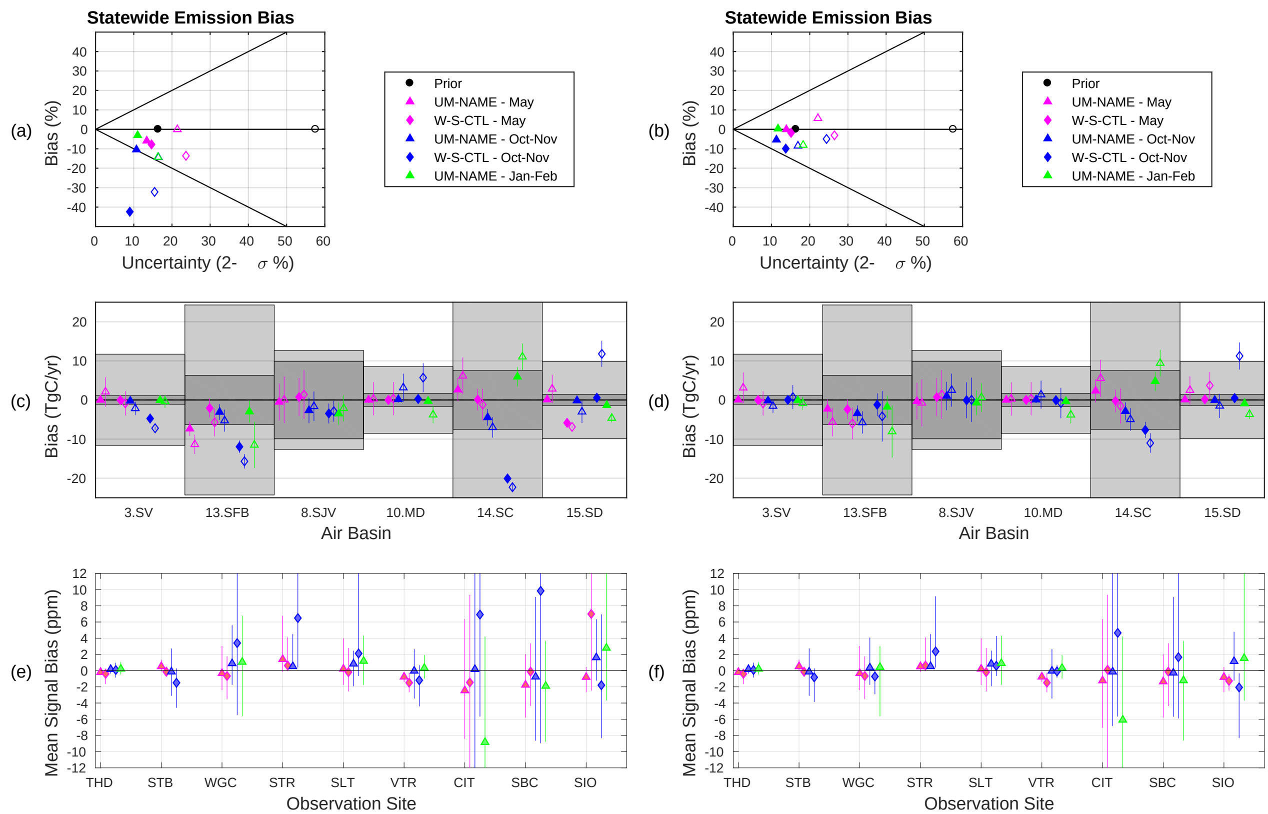

Figure 6: Inversion results for the experiment where the atmospheric transport in the prior simulation uses WS-CTL or UM-NAME but the atmospheric transport in the true simulation uses WS-LBL. Posterior statewide emissions (a, b), individual air basin emissions (c, d), and percentage error in simulated average ffCO_2 concentration (e, f) are shown with no outlier removal (first column) and outliers removed (second column). Prior and true emissions are given by annually averaged Vulcan.

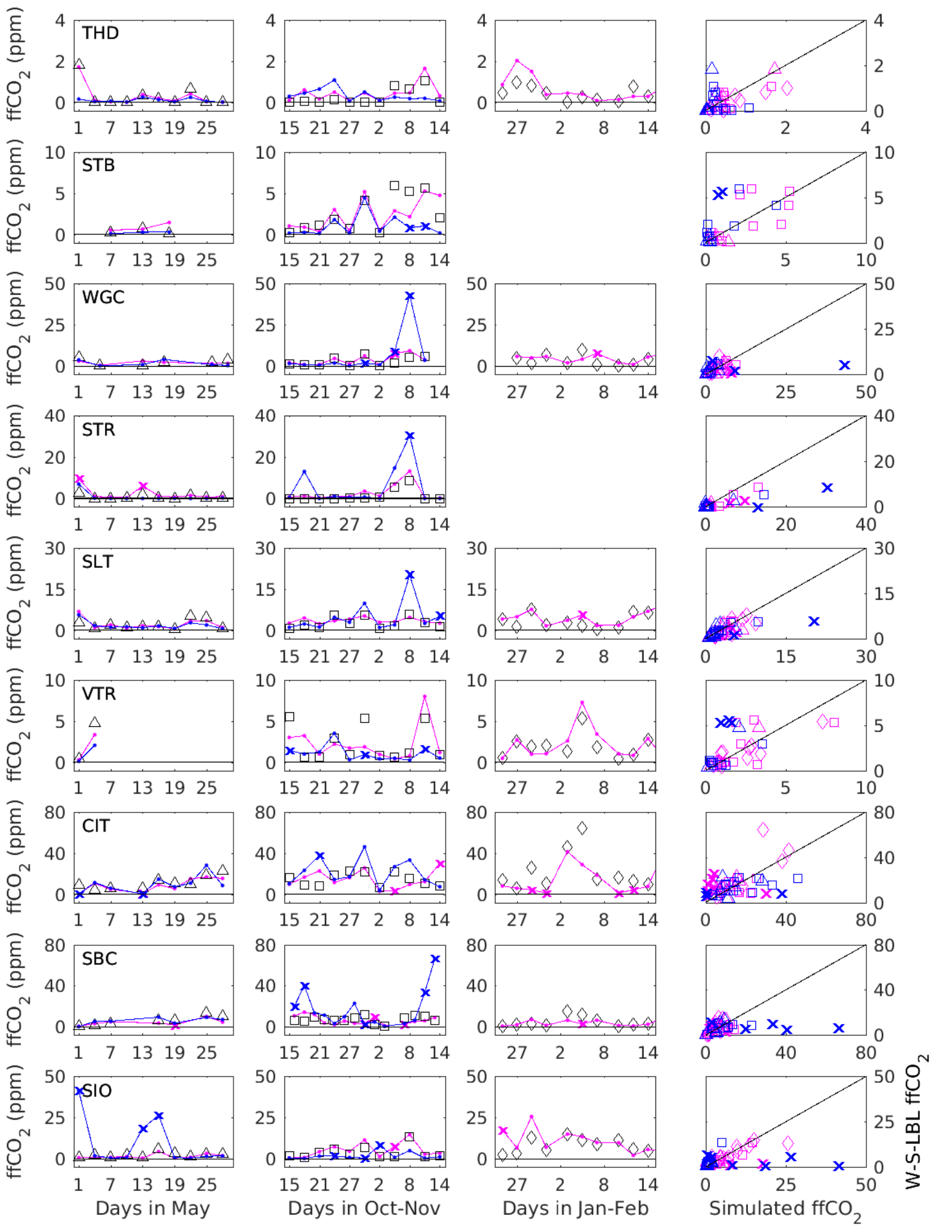

Figure 7: All simulated ffCO_2 from May (first column), October-November (second column), and January-February (third column). Simulated ffCO_2 using W-S-LBL are shown in black markers (triangles for May, squares for Oct-Nov and diamonds for Jan-Feb) whilst prior W-S-CTL signals are shown in blue and UM-NAME signals in magenta. All simulated signals are generated using the Vulcan gridded emissions map. The fourth column shows true vs prior ffCO_2 signals, with colors corresponding to models and markers corresponding to campaigns. Outliers omitted from the standard inversion are shown by an x.