Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The San Joaquin Valley

Yu Yan Cui, Jerome Brioude, Wayne M Angevine, Jeff Peischl, Stuart A Mckeen, Si-Wan Kim, J Andrew Neuman, Daven K Henze, Marc L Fischer, Seongeun Jeong, et al.

Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The San Joaquin Valley

Yu Yan Cui[1,2], Jerome Brioude[1,2,3], Wayne M. Angevine[1,2], Jeff Peischl[1,2], Stuart A. McKeen[1,2], Si-Wan Kim[1,2], J. Andrew Neuman[1,2], Daven K. Henze[4], Nicolas Bousserez[4], Marc L. Fischer[5], Seongeun Jeong[5], Hope A. Michelsen[6], Ray P. Bambha[6], Zhen Liu[6,7], Gregory W. Santoni[8], Bruce C. Daube[8], Eric A. Kort[9], Gregory J. Frost[2], Thomas B. Ryerson[2], Steven C. Wofsy[8], Michael Trainer[2]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA.
[2]Chemical Sciences Division, Earth System Research Laboratory, NOAA, Boulder, CO, USA.
[3]Laboratoire de l'Atmosphere et des Cyclones, UMR8105, CNRS-Meteo France-Universite La Reunion, La Reunion, France.
[4]Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA.
[5]Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
[6]Combustion Research Facility, Sandia National Laboratories, Livermore, CA, USA.
[7]Now at Ramboll Environ US Corporation, Novato, CA, USA.
[8]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA.
[9]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA.

Corresponding Author: Yu Yan Cui (Yuyan.Cui@noaa.gov)

Key Points:

- Estimate methane emissions in the San Joaquin Valley using inverse modeling and a mass-balance approach
- Methane emissions are estimated to be greater than the bottom-up inventory by a factor of 1.7
- Livestock largely account for differences between the optimized and prior methane emission estimates

Index terms:

0322 Constituent sources and sink

3260 Inverse theory

0345 Pollution: urban and regional

0365 Troposphere: composition and chemistry

Keywords:

methane; emission inventory; inverse modeling; mass-balance estimate; the San Joaquin Valley of California

Abstract

We quantify methane (CH_4) emissions in California's San Joaquin Valley (SJV) using four days of aircraft measurements from a field campaign during May-June 2010 together with a Bayesian inversion method and a mass-balance approach. For the inversion estimates, we use the FLEXible PARTicle dispersion model (FLEXPART) to establish the source-receptor relationship between sampled atmospheric concentrations and surface fluxes. Our prior CH_4 emissions estimates are from the California Greenhouse Gas Emissions Measurements (CALGEM) inventory. We use three meteorological configurations to drive FLEXPART and subsequently construct three inversions to analyze the final optimized estimates and their uncertainty (one standard deviation)**.** We conduct May and June inversions independently, and derive similar total CH_4 emissions estimates for the SJV: 135±28 Mg/hr in May and 135±19 Mg/hr in June. The inversion result is 1.7 times higher than the prior estimate from CALGEM. We also use an independent mass-balance approach to estimate CH_4 emissions in the northern SJV for one flight when meteorological conditions allowed. The mass-balance estimate provides a confirmation of our inversion results, and these two independent estimates of the total CH_4 emissions in the SJV are consistent with previous studies. In this study, we provide optimized CH_4 emissions estimates at 0.1° horizontal resolution. Using independent spatial information on major CH_4 sources, we estimate that livestock contribute 75–77% and oil/gas production contributes 15–18% of the total CH_4 emissions in the SJV. Livestock explain most of the discrepancies between the prior and the optimized emissions from our inversion.

1. Introduction

Methane (CH_4) is the second most significant greenhouse gas. It has a large global-warming potential and mediates global tropospheric chemistry. Globally, more than 60% of total CH_4 emissions are attributed to human activities [*EPA,* 2015], such as the natural gas and petroleum industries, domestic livestock operations, landfills, rice cultivation, and coal mining. Reducing CH_4 from human activity is important for reducing risks associated with climate change. As the most populous state of the US and a major CH_4 emitter, California enacted State Assembly Bill 32 (http://www.arb.ca.gov/cc/ab32/ab32.htm) in 2006 to reduce greenhouse gas emissions to 1990 emission levels by the year 2020, and to reduce greenhouse gas emissions to 40 percent below 1990 levels by year 2030. Achieving this goal requires accurate accounting of the magnitude and source attribution of CH_4 emissions.

The Central Valley covers about 14% of California's total land area and is the leading dairy-farming and most productive agricultural region in California. Twenty percent of US milk production occurs in California, mostly in the Central Valley (http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1103). The California Greenhouse Gas Emissions Measurements (CALGEM, http://calgem.lbl.gov) project found that the Central Valley is the California region with the highest CH_4 emissions [*Zhao et al.*, 2009; *Jeong et al.,* 2012; *Jeong et al.,* 2013]. The San Joaquin Valley (SJV), the southern portion of the Central Valley, contains a variety of potential CH_4 sources of anthropogenic origin, including approximately 2 million head of cattle and calves [National Agricultural Statistics Service, 2013], more than 75,000 active oil wells, and many cities.

Current bottom-up inventories of CH_4 sources in the SJV are quite uncertain. The Emission Database for Global Atmospheric Research (EDGAR) version 4.2 global emission inventory at 0.1° × 0.1° horizontal resolution (http://edgar.jrc.ec.europa.eu) reports that the CH_4 emissions from livestock in the SJV are 26.7 Mg/hr. However, a bottom-up study from CALGEM at 0.1° × 0.1° horizontal resolution calculated CH_4 emissions from livestock in the San Joaquin Valley to be 60.4 Mg/hr, more than twice that of EDGAR version 4.2 [*Jeong et al.,* 2013]. The SJV is also a significant region for petroleum and natural gas production. A new bottom-up study from *Jeong et al.* [2014] reports 3 to 7 times higher emissions from petroleum and natural gas production than the California Air Resources Board (CARB) 2013 Oil and Gas Industry Survey Results and 2014 greenhouse gas emissions inventory.

To improve emissions quantification, atmospheric measurements have increasingly been used to constrain the bottom-up emissions estimates. In the SJV, there are ongoing studies using the tower measurements to estimate CH_4 emissions [*Zhao et al.*, 2009; *Jeong et al.,* 2013, 2016]. Current satellite data have been used to constrain CH_4 in California, but CH_4 emissions estimates using satellite observations over the Central Valley remain difficult because of the scarcity of observations [*Wecht et al.,* 2014; *Bousserez et al.,* 2016].

A field campaign named the California Research at the Nexus of Air Quality and Climate Change (CalNex, *Ryerson et al.,* 2013) took place in California during May and June 2010. During CalNex, the NOAA WP-3 aircraft collected intensive measurements, including CH_4 mixing ratios, over the South Coast Air Basin and the Central Valley. To identify contributions from individual source categories, the aircraft flew close to emission sources with extensive horizontal and vertical coverage. The CalNex aircraft measurements provide a good opportunity to conduct a top-down estimate of the CH_4 emissions in these regions of California [*Peischl et al.*, 2013; *Cui et al.,* 2015]. The large spatial coverage of the aircraft enables sampling of multiple CH_4 sources distributed across the complex terrain of the SJV, providing a useful complement to ground-based and remote-sensing measurements.

This study uses a mesoscale inverse modeling technique to estimate CH_4 emissions in the SJV based on aircraft measurements from CalNex. This mesoscale inverse modeling system has already been employed to estimate CH_4 emissions in the South Coast Air Basin of California [*Cui et al.,* 2015] using measurements from the same campaign. The mass-balance approach [*White et al.,* 1976], an independent top-down method, is applied in part of the SJV to provide confirmation of the inverse modeling results. We compare our top-down CH_4 emissions estimates to three different inventories. We also compare our results with another inversion analysis of the same region using tower measurements [*Jeong et al.,* 2013, 2016].

The details of our methodology are described in Section 2. Our optimized emissions and interpretation of the results are presented in Section 3. Conclusions are given in Section 4.

2. Methods

In this section, we describe the atmospheric measurements of CH_4 mixing ratios from the National Oceanic and Atmospheric Administration (NOAA) WP-3 aircraft. We describe the prior CH_4 emission inventories, the construction of our atmospheric transport model used to build the source-receptor relationships, and the design of our Bayesian inverse modeling. The mass-balance approach, which provides an independent estimate of CH_4 emissions based on the aircraft measurements, is described.

2.1 Measurements

In CalNex, the NOAA WP-3 aircraft obtained in situ measurements over the SJV during four daytime flights (May 7, May 12, June 16, and June 18) (Figure 1). We classify the 8 counties of the SJV into two sub-regions named D1 and D2 (Figure 1 (A)). D1 is the southern SJV including Madera, Fresno, Tulare, Kings, and Kern Counties, and D2 is the northern SJV including San Joaquin, Stanislaus, and Merced Counties. D1 and D2 correspond to regions #12 and #8, respectively, of *Jeong et al.* [2013]. The May 7 and June 16 flights flew over D1, and the May 12 and June 18 flights flew over D2 (Figures 1 (C) and (D). We excluded flight portions over the ocean and during takeoff and landing from the Los Angeles area.

CH_4 mixing ratios observed by the NOAA P-3 aircraft were measured once per second using wavelength-scanned-cavity-ring-down spectroscopy (WS-CRDS; Picarro 1301 m) [*Peischl et al.,* 2012, 2013]. The precision of the 1-Hz CH_4 measurement is ± 1.4 ppbv, and accuracy is estimated at ±1.2 ppbv. We aggregate these observations into 30-s averages for use in the inversion framework, which, at a ground speed of approximately 100 m s^{-1}, correspond to segments of about 3 km horizontally (Figure 2). This aggregated dataset provides the receptor points in our backward trajectory simulations from the atmospheric transport models described in Section 2.3 and is used in an inverse-modeling analysis.

2.2 Prior emission inventory

A prior inventory provides critical information for Bayesian inversion modeling, particularly when atmospheric measurements alone cannot fully constrain the spatial distribution of the emissions sources. Inaccurate representation of the spatial distribution of emissions sources in a prior limits the performance of inverse modeling [*Xiang et al.,* 2013]. Therefore, we need to select the best available inventory for the prior input. We compared three available CH_4 inventories: a recent gridded top-down inventory based on the US EPA National Emissions Inventory (NEI 2011, https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data) [*Ahmadov et al.,* 2015], a recent gridded bottom-up inventory designed to be consistent with the US EPA Inventory of US Greenhouse Gas Emissions and Sinks (GHGI) for 2012 [*Maasakkers et al.,* 2016], and a gridded bottom-up inventory from CALGEM designed to match the CARB inventory for 2008 [*Jeong et al.,* 2012, 2013]. These three inventories provide annual average CH_4 emissions estimates.

The spatial distributions of the three inventories are shown in Figure S1, and their total CH_4 emissions for the SJV and its D1 and D2 sub-regions are listed in Table 1. The three inventories' SJV total CH_4 emissions estimates range from 68-107 Mg/hr. We find distinct variations between the three inventories' spatial distributions of CH_4 emissions from livestock and active oil and gas wells. CALGEM, developed by *Zhao et al.* [2009] and *Jeong et al.* [2012], relies on more detailed local information about source locations and activity to generate the gridded CH_4 emissions estimates, compared with the other two inventories based on EPA's NEI and GHGI. For example, CALGEM's spatial distributions for livestock and oil/gas sources are based on the California Department of Water Resources land-use survey database [*Salas et al.*, 2009] and the California Department of Conservation's Division of Oil, Gas, and Geothermal Resources database (http://www.conservation.ca.gov/dog/pubs_stats/annual_reports/Pages/annual_reports.aspx), respectively. Among the three inventories considered, CALGEM contains the most accurate spatial distributions for the major CH_4 sources in the Central Valley, and we therefore use CALGEM as the foundation of our prior inventory. We also update the oil/gas source sector of CALGEM in the SJV according to emissions from *Jeong et al.* [2014]. The CALGEM inventory is available at 0.1° × 0.1° spatial resolution, and we optimize the inventory at the same resolution.

Similar to *Cui et al.* [2015], our study adjusts the magnitude of total CH_4 emissions in each grid cell of the prior annual average inventory, without differentiating source sectors. When we calculate the contributions from different source sectors independently, we require extra spatial information. Figure 1 (B) presents the spatial information for the two dominant CH_4 sources in the SJV: dairies (an important livestock-related activity across the SJV) and active oil/gas wells [*Jeong et al.,* 2013]. Like CALGEM, we obtained the spatial information for livestock sources from *Salas et al.* [2008], and the spatial distribution of the active oil and gas wells was taken from California's Department of Conservation Division of Oil, Gas, and Geothermal Resources database (http://www.conservation.ca.gov/dog/pubs_stats/annual_reports/Pages/annual_reports.aspx). Livestock sources are highly concentrated in both the D1 and D2 sub-regions. Oil and gas production is mainly found in the southern part of D1. In the SJV, the oil and gas production sector has much larger CH_4 emissions than oil/gas processing, transmission and distribution [*Jeong et al.,* 2014].

Although livestock and oil/gas production are the two major sources in the SJV, they are rarely collocated in the same 0.1° grid cell, allowing for the estimation of total emissions from each of them. In this study, if a grid cell includes more than one sector, only the sector with the highest emission in that cell is represented (this situation occurs less than 5% of the time). We assume that the uncertainty of the total emissions estimates due to the spatial partitioning of the two major sources is smaller than the transport uncertainty, and we did not explicitly include the spatial partitioning uncertainty for the source contribution estimate in this study. The similar spatial patterns shown in Figure 1 (A) and (B) demonstrate that the prior inventory captures the spatial patterns of major sources.

2.3 Atmospheric transport modeling

Following *Cui et al.,* [2015], the FLEXPART-WRF Lagrangian model version 3.1 [*Brioude et al.,* 2013] is used to calculate source-receptor relationships, a.k.a. footprints. The surface footprints ($s\ m^2\ kg^{-1}$) represent the residence time within a surface layer (below 100 m above ground level) weighted by the atmospheric density. We conducted three atmospheric transport simulations using FLEXPART driven by three different meteorology configurations from the Weather Research Forecasting Model (WRF) (Table 2). The three WRF meteorological fields have a 4 x 4 km horizontal grid spacing. The first and second meteorology configurations (WRF1 and WRF2) are from *Angevine et al.* [2012]. The third WRF configuration (WRF3) is from *Kim et al.* [2016]. Using measurements from the same field campaign, WRF1 and WRF2 have been used to estimate nitrous oxide emissions in the Central Valley [*Xiang et al.,* 2013], and WRF3 has been used to estimate ozone in the Los Angeles region [*Kim et al.,* 2016]. Detailed information on evaluations of planetary boundary layer height (PBLH), wind speed, and wind direction from the three transport models can be found in *Angevine et al.* [2012] and *Kim et al.,* [2016]. Here we show model evaluations using observations from the four flights in Figures S2-S4 and Table S1.

Correlations between any of the three CH_4 simulations with differing meteorological configurations are no larger than the correlations between any model simulation and the observations. Therefore, the three model simulations can be treated as independent representations of the meteorology. Each model is used in our inverse modeling system to derive the posterior emissions estimates, and the final optimized emissions estimates are based on the mean value from the three estimates. Three meteorological models can only represent part of the phase space of model uncertainties. A complete estimate of transport model uncertainty would require a larger ensemble and more comprehensive characterization [*Angevine et al.,* 2014].

Ten thousand FLEXPART-WRF back trajectories were initiated at each receptor point along the flight track and run for three days backward in time. We derive our surface footprint from FLEXPART-WRF at the same spatial resolution (0.1° x 0.1°) as the prior. The surface footprints for the May and June inversions from each of the transport models are shown in Figure 3.

Figure 4 presents the mean vertical profiles of CH_4 mixing ratios in 100-m vertical intervals over the SJV from the aircraft measurements and from the three transport models using the CH_4 prior inventory. The error bars represent the standard deviations among the three different transport models. There is no obvious bias in the simulated vertical mixing. There is a small bias in simulating CH_4 in the upper part of the mean profile, but the bias is statistically insignificant as it is smaller than the uncertainty range of the CH_4 background determination (see next section). There is a systematic low bias in the modeled CH_4 concentrations below 1600-1800 m above sea level (ASL), which is attributed to a bias in the prior emissions estimates as shown below.

2.4 Bayesian inverse modeling

We perform a 4-dimensional (three spatial dimensions in the model plus time) inversion using a Bayesian framework by minimizing a cost function assuming lognormal distributions for the observed enhancements and surface fluxes [*Brioude et al., 2011*]. The cost function used in the inversion framework is

$$J = \frac{1}{2}(\ln(y_0) - \ln(Hx))^T R^{-1}(\ln(y_0) - \ln(Hx)) + \frac{1}{2}\alpha(\ln(x) - \ln(x_b))^T B^{-1}(\ln(x) - \ln(x_b)),$$

where $\boldsymbol{y_o}$ is the measured time series of CH_4 mixing ratio enhancement above defined background, $\boldsymbol{H}$ is the source-receptor relationship matrix calculated by FLEXPART-WRF, $\boldsymbol{R}$ and $\boldsymbol{B}$ are the error covariance matrices of the model-observation mismatch and the prior information, respectively, $\boldsymbol{x_b}$ is the prior emission inventory, and $\boldsymbol{x}$ is the posterior emission inventory to be determined. The parameter $\boldsymbol{\alpha}$ [*Henze et al., 2009*] balances the errors of both covariance matrices in the minimization of the cost function to calculate the best estimates of emissions.

The surface emissions optimization applied in this study is based on the inverse modeling framework applied in *Cui et al.* [2015]. Most CH_4 mixing ratio enhancements were measured below 2.0 km altitude ASL during the four flights. To reduce the potential uncertainty in the transport models' ability to distinguish between the PBL and the free troposphere, we focus on the measurements (i.e., receptor points) below 1.5 km ASL (Figure 2). Choosing a threshold of 2.0 km or 1.0 km ASL does not significantly affect our results.

For each flight, we plot the histogram of the observed CH_4 mixing ratios below 1.5 km ASL on the upwind side of the domain. We choose the mode of this distribution as the background value. Based on the width of this distribution, we estimate a 10 ppbv uncertainty in the background mixing ratio for each flight.

The NOAA P-3 flights over the SJV flew close to surface sources, so that the measurements were obtained within hours from the time of emission. Therefore, it is reasonable to assume that photochemical loss of CH_4 can be neglected. Hence, CH_4 is treated as a passive tracer in our mesoscale inverse system.

We conduct a cluster aggregation process for the spatial grid cells as described by *Cui et al.* [2015]. Surface grid cells in the domain are clustered using a neighbor method based on the information from the Fisher information matrix [*Bocquet et al.,* 2011]. We use this method to obtain inversion solutions efficiently and to reduce cross correlations between surface fluxes during the inverse modeling. In this study, 4544 (64 x 71) grid cells resulted in 2024 clusters in our inverse modeling system.

The $\boldsymbol{R}$ and $\boldsymbol{B}$ covariance matrices are assumed to be diagonal matrices. $\boldsymbol{R}$ is calculated by the addition in quadrature of the 30-second aggregation uncertainty (i.e., the standard deviation of a 30-s interval, 10 ppbv for the mean value), the background uncertainty (10 ppbv), and the uncertainty of each transport model (50% [*Angevine et al.,* 2014], 50 ppbv for the mean value) in simulating CH_4 enhancements above background. The largest uncertainty in $\boldsymbol{R}$ is that of the transport models. We assume a larger uncertainty in the models in this study than in the Los Angeles basin [*Cui et al.,* 2015] because of the inherent difficulty in modeling the transport within the complex terrain of the Central Valley.

Jeong et al. [2013] classified the state of California into 13 sub-regions to conduct their inverse modeling and assumed 70% uncertainties in each sub-region for their prior inventory (CALGEM). We assume a 100% relative uncertainty for each cluster in our prior, since one sub-region from *Jeong et al.* [2013] is comprised of multiple clusters of our grid cells and because we updated the magnitude and spatial locations of oil and natural gas production in the CALGEM inventory. We test the sensitivity of our results to the 70% assumption of the prior's uncertainty (compare Table S2 to Table 3). Using a prior uncertainty of 70% instead of 100% for each cluster does not significantly affect our optimized emission estimates.

To carry out inverse modeling in the lognormal framework, we define all uncertainties as the arithmetic standard deviation (***SD[X]***) for a variable (**X**), including the measurements, the background determination, the transport model, the prior inventory, and the posterior estimates of each inversion. We define the covariance error matrixes ($\boldsymbol{R}$ and $\boldsymbol{B}$) as the squared scale parameter ($\boldsymbol{\sigma^2}$) of the variable (**X**). ***SD[X]*** and $\boldsymbol{\sigma^2}$ have the following relationship: $\sigma^2 = ln\left(1 + \frac{(SD[X])^2}{(E[X])^2}\right)$, where ***E[X]*** is the arithmetic mean.

For each sub-region, the total emissions estimate is calculated by summing the emissions estimates of the clusters in the region. The total uncertainty estimate for each sub-region is calculated as the square root of the sum of the variances along the diagonal in the posterior error covariance matrix. We do not include the off-diagonal elements of this matrix because some are negative (indicating anti-correlation between two grid cells), and including them would result in a slightly smaller uncertainty estimate. Instead we report the larger, more conservative uncertainty based on the diagonal elements only. A similar uncertainty estimate was also used in *Jeong et al.* [2013]. The optimized emissions estimates from each of the transport models are shown in Table 3. The final optimized estimates and the associated uncertainties are built by a resampling method shown in Table 3 from the three inversions based on the three transport models.

2.5 Mass-balance approach

CH_4 emission fluxes were determined using the mass-balance approach [*White et al.,* 1976] for comparison with the inversions. In this study, we use this approach to quantify CH_4 emissions using measurements made both upwind and downwind of the emission sources. We estimate the total CH_4 emissions from the D2 sub-region of the SJV when favorable meteorological conditions were observed, including steady horizontal winds, and a well-developed PBL that was well mixed vertically. The uncertainties associated with the assumptions of the technique are included. The details of the mass-balance approach are described in *Peischl et al.,* [2015].

3. Results and discussion

3.1 San Joaquin Valley CH_4 emissions estimates from the inversions

We optimize the spatially resolved CH_4 emissions estimates in the SJV using the mesoscale inverse modeling system with the CalNex airborne measurements (Figure 5). The optimized estimates are from two independent inversions using observations in the May and June 2010 flights. The May and June inversions derive similar total CH_4 emissions estimates for the SJV (Table 1). We estimate the total CH_4 emissions from the SJV to be 135±28 Mg/hr in May 2010 and 135±19 Mg/hr in June 2010. The difference in total emissions between May and June is statistically insignificant. In general, the spatial patterns of the CH_4 prior inventory are consistent with those of the optimized emissions estimates (Figure 5). However, the optimized emissions in May and June both indicate that the magnitudes of the prior emissions in the SJV are much lower than the optimized estimates (Figure 5 (B) and (D)). The highest emission rates (and the largest adjustments to the prior) are seen in the region from Hanford to Visalia in the southern sub-region (D1) and from Merced to Stanislaus in the northern sub-region (D2) of the SJV. Our optimized estimates on average in the SJV are higher by a factor of 1.7 than the prior estimates based on the CALGEM inventory.

The optimized total CH_4 emissions estimates from each transport model are shown in Table 3. The transport model evaluations shown in Table S1 indicate that WRF3 has a large (57%) bias in simulating PBLH in D2 for the May inversion case. Therefore, in Table 3 we also list the overall estimates based only on WRF1 and WRF2 simulations. In May, using only these two simulations results in only a 10% difference in estimated SJV CH_4 emissions compared with the results based on three WRF simulations; differences in June are much smaller. We therefore base our main conclusions on results from the three WRF simulations for both May and June.

To evaluate the optimized emissions, we compare the measured CH_4

enhancements above background and those simulated by FLEXPART-WRF using the optimized emissions estimates and the prior estimates (Figure 6 and Figure 7, Table 1). The FLEXPART-WRF simulation using the optimized emissions captures the observations with a coefficient of determination (r^2) of 0.76 and 0.71 for the May and June inversions, respectively. These correlations are higher than for the simulations using the prior estimates (r^2 = 0.49 and 0.47, respectively). Moreover, there is a large decrease in the mean bias using the optimized emissions. The mean biases between the observed and simulated CH_4 enhancements using the prior inventory in the May and June inversions are -55.2 and -31.8 ppbv, respectively. In contrast, the observed-simulated biases using the optimized emissions are only -9.1 and -5.5 ppbv, respectively, an 83% decrease for both inversions compared to the corresponding results based on the prior inventory. Additionally, the vertical profiles of CH_4 mixing ratios are well captured by the models when we use the optimized CH_4 emissions estimates (Figure 4).

We compare optimized emissions estimates in the present study to the top-down estimate from *Jeong et al.* [2013, 2016] (Table 1). The total emissions estimates for the SJV in this study are similar to estimates from *Jeong et al.* [2016] (98-170 Mg CH_4/hr). In this study, we use many more grid clusters than the number of grid cells in *Jeong et al.,* [2013] to invert for the surface fluxes in the SJV. The total emissions estimates are similar, while the partitioning of CH_4 emissions between sub-regions D1 and D2 differ between our study and *Jeong et al.* [2013]. We estimate total CH_4 emissions from D1 to be 80±17 Mg/hr in May and 79±17 Mg/hr in June (Table 1), and the total CH_4 emissions from D2 to be 55±18 Mg/hr in May and 56±13 Mg/hr in June. The differences between May and June are statistically insignificant. The estimated emissions for D1 are lower than those of *Jeong et al.* [2013], while those for D2 are higher on average than those of *Jeong et al.* [2013]. *Jeong et al.* [2013] only used two grid cells to represent the domain of the SJV in their inversions, while we substantially improved the spatial resolution by aggregating 4544 grid cells (0.1° × 0.1°) into 2024 clusters. The difference in spatial resolution between the two studies results in different transport and emissions estimates.

3.2 San Joaquin Valley CH_4 emissions estimates from the mass-balance approach

We use the same CalNex aircraft measurements and an independent mass-balance approach to derive CH_4 emissions from the SJV. We determined emissions in the northern SJV sub-region (D2) using measurements from the May 12 flight, the only day with favorable meteorological conditions in the Central Valley during CalNex.

On the May 12 flight, the upwind transect in San Joaquin County (Figure 1 (C)) resulted in a CH_4 flux of 28±19 Mg/hr (1-sigma uncertainty) coming mainly from the nearby Sacramento Valley. The downwind transect in Merced County resulted in a flux of 97±45 Mg/hr. The difference between the upwind and downwind transects, 69±47 Mg CH_4/hr, represents the estimated emissions from sub-region D2, assuming the upwind sources were constant while the wind traveled from the upwind transect to the downwind transect. Details of the mass-balance calculation are given in Table 4. Within the stated uncertainties, the mass-balance emissions estimate agrees with our inversion in D2 (55±18 Mg CH_4/hr in May). Therefore, an independent method purely based on the measurements confirms our optimized inversion results.

We did not conduct a mass-balance analysis for the southern SJV region (D1) in this study because CH_4 surface emissions from D2 strongly influenced CH_4 in D1 (Figure

8). In addition, the nighttime Fresno eddy [*Bao et al.,* 2008] complicates the application of a mass-balance approach to the flights over D1, such as leading to a build-up of CH_4 enhancements in the entire domain the following day and violating the steady wind assumption. Therefore, favorable conditions for mass-balance estimates in D1 are difficult to obtain during CalNex. Similarly, winds over the D1 and D2 regions during the June flights had a westerly component that transported emissions through the eastern edge of the San Joaquin Valley and beyond the extent of the downwind flight legs, so we could not carry out mass-balance estimates using the June flights. These limitations to using the mass-balance approach in the SJV show the value of inverse modeling estimates for the region.

3.3 Major source contributions in the San Joaquin Valley

Livestock sources (including dairies and animal feeding operations) are the largest source of CH_4 emissions in both sub-regions of the San Joaquin Valley. Livestock and oil/gas production sources are rarely collocated in the same 0.1° grid cell. In the few cases where a grid cell contains more than one CH_4 source, the source type of the cell is determined by the dominant source. Combining our optimized 0.1° resolution CH_4 emissions estimates (Figure 5) and the locations of two major sources (Figure 1 (B)), we estimate the CH_4 emissions from livestock sources in the SJV to be 103±29 Mg/hr and 105±25 Mg/hr for May and June, respectively (Table 5), which are higher than the prior CH_4 emissions by a factor of 1.8. Livestock emissions contribute 75–77 % of the total CH_4 emissions in the SJV according to our optimized results on average. Our estimates are consistent with the analysis of *Jeong et al.* [2016], who estimate SJV CH_4 emissions from the livestock source sector are 81-177 Mg/hr. Moreover, our finding for livestock sources is consistent with the analysis of *Johnson et al.* [2016], who estimate a factor of 2 higher emissions from a top-down approach compared with the CALGEM inventory.

Active oil/gas wells are mainly located in the southern SJV (Figure 1 (B)). We estimate CH_4 emissions in the SJV from the active oil/gas wells to be 24±11 Mg/hr in May and 21±7 Mg/hr in June (Table 4), which are higher than the prior CH_4 emissions by a factor of 1.6. On average, the wells emissions contribute 15–18% of the total CH_4 emissions in the SJV according to our optimized results. Our results are in agreement with the *Jeong et al.* [2014, 2016] estimates of 19 Mg/hr from oil and natural gas production in the SJV.

We also calculate the fractional adjustment in each of the two sources relative to the fractional change between the prior and optimized estimates of the SJV total CH_4 emissions. On average, livestock sources explain 82-86% of the discrepancy between our prior and optimized estimates, while oil/gas production explains 13-18% of the discrepancy.

4. Conclusions

Using airborne measurements collected during the CalNex 2010 study, we apply a mesoscale inverse model to perform a top-down estimate of CH_4 emissions in the San Joaquin Valley of California. Our optimized estimates of total CH_4 emissions in the San Joaquin Valley in May 2010 (June 2010) are 135±28 (135±19) Mg CH_4/hr. Our optimized CH_4 emissions estimates are higher by a factor of 1.7 than the prior estimates based on CALGEM.

We compare our inversions based on CalNex four days of aircraft measurements with inversions conducted using tall tower measurements [*Jeong et al.,* 2013, 2016]. The total SJV CH_4 emissions derived from these complementary inversion approaches agree within the uncertainties, while our inversions provide SJV emissions estimates at a finer spatial distribution than these previous studies. The optimized spatial emissions information that we derive helps to refine source attributions. We also compare our inversions with the annual average SJV CH_4 emissions (107 Mg CH4/hr) from a recent national bottom-up CH_4 inventory [Maasakkers et al., 2016], and within the uncertainties our optimized estimates agree with these bottom-up estimates.

Our optimized estimates, based on only four days of aircraft measurements in the summer of 2010, do not capture episodic or seasonal variations in SJV emissions. Therefore, we cannot carry out fully quantitative comparisons with the annual average emissions of the CALGEM prior and Maasakkers et al. [2016] inventories, nor with the longer analysis periods of the inversions performed by Jeong et al. [2013, 2016] in different years than 2010.

Compared with the prior CALGEM inventory, our optimized estimates for CH_4 emissions from livestock sources are higher by a factor of 1.8, while our optimized CH_4 emissions from oil/gas production are higher by a factor of 1.6. Livestock are the most important source of CH_4 emissions in the SJV, and we find that livestock sources explain most of the discrepancies between the prior and our optimized CH_4 emissions estimates. Our use of high-frequency aircraft observations and a model with high spatial resolution allow us to distinguish signals from livestock and oil/gas sources and to provide a quantitative top-down constraint on the emissions from these sectors.

To validate our optimized emissions estimates, we also conduct a mass-balance estimate for one flight and one sub-region as an independent approach. Our optimized estimates are in agreement with the mass-balance estimate within the combined uncertainty of the two approaches. The mass-balance method using aircraft observations can be used to estimate emissions from a region under favorable meteorological conditions, but such conditions do not always occur. For instance, no mass-balance estimates could be performed for the southern SJV in this study. Mesoscale inverse modeling therefore offers a reliable, complementary technique for quantifying emissions from multiple CH_4 sources over a large area.

Our inversions based on high quality aircraft measurements provide estimates of CH_4 emissions in the San Joaquin Valley that agree with previous inversion calculations based on tall tower observations. These independent top-down estimates confirm that major CH_4 sources in the Valley are underestimated by the CALGEM prior inventory. This study shows that applying an inverse model to tower and aircraft measurements to assess and improve emissions estimates can inform bottom-up inventories and could ultimately be useful in evaluating emissions reduction strategies.

Acknowledgments:

FLEXPART-WRF model is available at the official FLEXPART website (http://flexpart.eu). NOAA P-3 data are available and can be downloaded at http://www.esrl.noaa.gov/csd/projects/calnex. The optimized emission inventory is available online as supporting information in NetCDF format. The lognormal Bayesian inverse software was developed at NOAA/ESRL/CSD and CIRES. The WRF initial/boundary data were provided by ERA-Interim and NOAA/NCEP. U.S. EPA NEI 2011 provided information that was compared to our prior inventory. We thank NOAA's High Performance Computing Program for their support in running FLEXPART-WRF. This work was supported in part by NOAA's Atmospheric Chemistry, Carbon Cycle, and Climate Program. J.B., D.K.H., N.B., and M.T. acknowledge support from the NOAA Climate Program Office (CPO) (NA14OAR4310136). M.L.F. and S.J. acknowledge support from the California Energy Commission Public Interest Environmental Research Program to LBNL under contract no. DE-AC02-05CH11231. Z.L., R.P.B., and H.A.M are supported under the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DEAC04-94AL85000.

References:

Ahmadov, R., S. McKeen, M. Trainer, R. Banta, A. Brewer, S. Brown, P.M. Edwards, J.A. de Gouw, G.J. Frost, J. Gilman, D. Helmig, B. Johnson, A. Karion, A. Koss, A. Langford, B. Lerner, J. Olson, S. Oltmans, J. Peischl, G. Pétron, Y. Pichugina, J.M. Roberts, T. Ryerson, R. Schnell, C. Senff, C. Sweeney, C. Thompson, P. Veres, C. Warneke, R. Wild, E.J. Williams, B. Yuan, and R. Zamora, Understanding high wintertime ozone pollution events in an oil and natural gas producing region of the western US, *Atmospheric Chemistry and Physics*, doi:10.5194/acp-15-411-2015, 2015.

Angevine, W. M., L. Eddington, K. Durkee, C. Fairall, L. Bianco, and J. Brioude (2012), Meteorological model evaluation for CalNex 2010, Mon. Weather Rev., 140, 3885–3906, doi:10.1175/MWR-D-12-00042.1.

Angevine, W. M., J. Brioude, S. McKeen, and J. S. Holloway (2014), Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty at mesoscale, Geosci. Model Dev., 7, 2817–2829, doi:10.5194/gmd-7-2817-2014.

Bao, J-W, S. A. Michelson, P. O. G. Persson, I. V. Djalalova, and J. M. Wilczak (2008), Observed and WRF-Simulated Low-Level Winds in a High-Ozone Episode during the Central California Ozone Study. J. Appl. Meteor. Climatol., 47, 2372–2394, DOI: http://dx.doi.org/10.1175/2008JAMC1822.1.

Bocquet, M., L. Wu, and F. Chevallier (2011), Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism, Q. J. R. Meteorol. Soc., 137, 1340–1356.

Bousserez, N., D. K. Henze, B. Rooney, A. Perkins, K. J. Wecht, A. J. Turner, V. Natraj, J. R. Worden (2016), Constraints on methane emissions in North America from future geostationary remote sensing measurements, Atmos. Chem. Phys., 16, 6175–6190, doi:10.5194/acp-16-6175-2016.

Brioude, J., et al. (2011), Top-down estimate of anthropogenic emission inventories and their interannual variability in Houston using a mesoscale inverse modeling technique, J. Geophys. Res., 116, D20305, doi:10.1029/2011JD016215.

Brioude, J., Arnold, D., Stohl, A., Cassiani, M., Morton, D., Seibert, P., Angevine, W., Evan, S., Dingwell, A., Fast, J. D., Easter, R. C., Pisso, I., Burkhart, J., and Wotawa, G.: The Lagrangian particle dispersion model FLEXPART-WRF version 3.1, Geosci. Model Dev., 6, 1889-1904, doi:10.5194/gmd-6-1889-2013, 2013.

Chen, F., and J. Dudhia, (2001), Coupling an advanced land surface– hydrology model with the Penn State–NCAR MM5 modeling system Part I: Model implementation and sensitivity. Mon. Wea. Rev., 129, 569–585, doi: http://dx.doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Cui, Y. Y., J. Brioude, S. A. McKeen, W. M. Angevine, S.-W. Kim, G. J. Frost, R. Ahmadov, J. Peischl, N. Bousserez, Z. Liu, T. B. Ryerson, S. C. Wofsy, G. W. Santoni, E. A. Kort, M. L. Fischer, and M. Trainer (2015), Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin. J. Geophys. Res. Atmos., 120, 6698–6711. doi: 10.1002/2014JD023002.

Dudhia, Jimy, (1996), A multi-layer soil temperature model for MM5 the Sixth PSU/NCAR Mesoscale Model Users' Workshop.

EPA (2015), Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013, http://epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2015-Main-Text.pdf

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with explicit treatment of entrainment processes. Mon. Wea. Rev., 134, 2318–2341.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012), Seasonal variation of CH_4 emissions from central California, J. Geophys. Res., 117, D11306, doi:10.1029/2011JD016896.

Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak and M. L. Fischer (2013), A Multi-tower Measurement Network Estimate of California's Methane Emissions. J. Geophys. Res., doi: 10.1002/jgrd.50854.

Jeong, S., D. Millstein and M. L. Fischer (2014), Spatially Explicit Methane Emissions from Petroleum Production and the Natural Gas System in California, Env. Sci. Technol., doi: 10.1021/es4046692.

Jeong, S., et al. (2016), Estimating methane emissions in California's urban and rural regions using multi-tower observations, J. Geophys. Res. Atmos., 121, doi:10.1002/2016JD025404.

Johnson et al., (2016), Investigating seasonal methane emissions in northern California using airborne measurements and inverse modeling, J. Geophys. Res. Atmos., doi: 10.1002/2016JD025157.

Kalnay, E., M. Kanamitsu, and W. E. Baker (1990), Global numerical weather prediction at the National Meteorological Center, Bull. Amer. Meteor. Soc., 71, 1410–1428, doi: 10.1175/1520-0477(1990)071<1410:GNWPAT>2.0.CO;2.

Kim, S.-W., B. C. McDonald, S. Baidar, S. S. Brown, B. Dube, R. A. Ferrare, G. J. Frost, R. A. Harley, J. S. Holloway, H.-J. Lee, et al. (2016), Modeling the weekly cycle of NOx and CO emissions and their impacts on O3 in the Los Angeles-South

Coast Air Basin during the CalNex 2010 field campaign, J. Geophys. Res. Atmos., 121, 1340–1360, doi:10.1002/2015JD024292.

Maasakkers, J. D., et al. (2016), Gridded national inventory of U.S. methane emissions, Environ. Sci. Technol., 50, 13,123–13,133, doi:10.1021/ acs.est.6b02878.

Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems, Rev. Geophys. 20(4), 851–875, doi:10.1029/RG020i004p00851.

National Agricultural Statistics Service (2013), USDA, http://www.nass.usda.gov/Statistics_by_State/California/Publications/California_Ag_Statistics/2013cas-all.pdf

Peischl, J., T. B. Ryerson, J. S. Holloway,M. Trainer, A. E. Andrews, E. L. Atlas, D. R. Blake, B. C. Daube, E. J. Dlugokencky, M. L. Fischer, A. H. Goldstein, A. Guha, T. Karl, J. Kofler, E. Kosciuch,P. K. Misztal, A. E. Perring, I. B. Pollack, G. W. Santoni, J. P. Schwarz, J. R. Spackman, S. C. Wofsy, and D. D. Parrish (2012), Airborne observations of methane emissions from rice cultivation in the Sacramento Valley of California, J. Geophys. Res. Atmos., 117, D00V25, doi:10.1029/2012jd017994.

Peischl, J., T. B. Ryerson, J. Brioude, K. C. Aikin, A. E. Andrews, E. Atlas, D. Blake, B. C. Daube, J. A. de Gouw, E. Dlugokencky, G. J. Frost, D. R. Gentner, J. B. Gilman, A. H. Goldstein, R. A. Harley, J. S. Holloway, J. Kofler, W. C. Kuster, P. M. Lang, P. C. Novelli, G. W. Santoni, M. Trainer, S. C. Wofsy, D. D. Parrish (2013), Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118, 4974–4990, doi:10.1002/jgrd.50413.

Peischl, J., Ryerson, T. B., Aikin, K. C., deGouw, J. A., Gilman, J. B., Holloway, J. S., Lerner, B. M., Nadkarni, R., Neuman, J. A., Nowak, J. B., Trainer, M., Warneke, C. and Parrish, D. D. (2015), Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. J. Geophys. Res. Atmos., 120: 2119–2139, doi: 10.1002/2014JD022697.

Ryerson, T. B., A. E. Andrews, W. M. Angevine, T. S. Bates, C. A. Brock, B. Cairns, R. C. Cohen, O. R. Cooper, J. A. de Gouw, F. C. Fehsenfeld, R. A. Ferrare, M. L. Fischer, R. C. Flagan, A. H. Goldstein, J. W. Hair, R. M. Hardesty, C. A. Hostetler, J. L. Jimenez, A. O. Langford, E. McCauley, S. A. McKeen, L. T. Molina, A. Nenes, S. J. Oltmans, D. D. Parrish, J. R. Pederson, R. B. Pierce, K. Prather, P. K. Quinn, J. H. Seinfeld, C. J. Senff, A. Sorooshian, J. Stutz, J. D. Surratt, M. Trainer, R. Volkamer, E. J. Williams, S. C. Wofsy (2013), The 2010 California Research at the Nexus of Air Quality and Climate Change (CalNex) field study, J. Geophys. Res. Atmos., 118, 5830– 5866, doi:10.1002/jgrd.50331.

Salas, W. A., et al. (2008), Developing and applying process-based models for estimating greenhouse gas and air emission from California dairies, California Energy

Commission, PIER Energy-Related Environmental Research, CEC-500-2008-093, http://www.energy.ca.gov/2008publications/CEC-500-2008-093/CEC-500-2008-093.PDF.

Salas, W., C. Li, F. Mitloehner, and J. Pisano (2009), Developing and applying process-based models for estimating greenhouse gas and air emissions from California dairies, Rep. CEC-500-2008-093, Public Interest Energy Res. Program, Calif. Energy Comm., Sacramento, Calif.

Wecht, K. J., D. J. Jacob, M. P. Sulprizio, G. W. Santoni, S. C. Wofsy, R. Parker, H. Bösch, and J. Worden (2014), Spatially resolving methane emissions in California: constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI, geostationary) satellite observations, Atmos. Chem. Phys., 14, 8173-8184, doi:10.5194/acp-14-8173-2014.

White, W., J. Anderson, D. Blumenthal, R. Husar, N. Gillani, J. Husar, and W. Wilson (1976), Formation and transport of secondary air pollutants: Ozone and aerosols in the St. Louis urban plume, Science, 194, 187–189, doi: 10.1126/science.959846.

Xiang, B., S. M. Miller, E. A. Kort, G. W. Santoni, B. C. Daube, Bruce C. R. Commane, W. M. Angevine, T. B. Ryerson, M. K. Trainer, A. E. Andrews, T. Nehrkorn, H. Tian, and S. C. Wofsy, (2013), Nitrous oxide (N2O) emissions from California based on 2010 CalNex airborne measurements, J. Geophys. Res. Atmos., 118, 2809–2820, doi:10.1002/jgrd.50189.

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, J. Geophys. Res., 114, D16302, doi:10.1029/2008JD011671.

Table 1. Comparison of Total CH_4 Emission Estimates in the San Joaquin Valley.

	SJV (Mg/hr)	D1 (Mg/hr)	D2 (Mg/hr)	r^2	Slope	Mean Bias (Post-Prior) (ppbv)
"May case" Optimized (This study, top-down)	135±28	80±17	55±18	0.76	0.63	-9.1
"June case" Optimized (This study, top-down)	135±19	79±17	56±13	0.71	0.61	-5.5
"May case" Prior (Based on CALGEM, bottom-up)	80	52	28	0.49	0.25	-55.2
"June case" Prior (Based on CALGEM, bottom-up)	80	52	28	0.47	0.24	-31.8
Jeong et al., [2013] (Tall tower network, top-down)	-	120±16	33±5	-	-	-
Jeong et al., [2016] (Tall tower network, top-down)	98-170	-	-	-	-	-
CH_4 annual average Inventory (based on NEI 2011, Ahmadov et al.)	68	46	22	-	-	-
CH_4 annual average Inventory (based on EPA-GHGI 2012, Maasakkers et al. [2016])	107	75	32	-	-	-
Mass-balance approach (This study, top-down)		-	69±47	-	-	-

Table 2. Names and Primary Configurations of Three WRF Runs used in This Study

Name	Version	Initialization	PBL Scheme	Grid Spacing (km)	Vertical Levels	LSM, data	Wind field
WRF1[a]	WRF 3.3	ERA-Interim	MYJ	4	60	Noah, UCM, MODIS	Time-averaged winds
WRF2[b]	WRF 3.3	NCEP-GFS	MYJ	4	40	Slab, USGS	Time-averaged winds
WRF3[c]	WRF-Chem3.4	NCEP-GFS	YSU	4	60	Noah, USGS	Time-averaged winds

[a,b] Angevine et al. [2012], [C] Kim et al. [2016]. WRF1 is initialized by the European Centre for Medium-Range Weather Forecasts' Re-Analysis-Interim (ERA-Interim). WRF1 is coupled to the Noah Land Surface Model with MODIS land products and a single-layer Urban Canopy Model (UCM) [Chen and Dudhia, 2001]. The Mellor-Yamada-Janjic (MYJ) scheme [Mellor and Yamada, 1982] is used to simulate planetary boundary layer (PBL). WRF2 is initialized by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS)[Kalnay et al., 1990]. The land surface model in WRF2 is a five-layer thermal diffusion land surface scheme ("Slab") [Dudhia, 1996] with USGS land products. WRF3 is initialized with data from the NCEP-GFS, and the PBL is simulated using the Yonsei University (YSU) boundary layer model [Hong et al., 2006].

Table 3. Optimized CH_4 Emissions in May and June from Each of Three Transport Models and the Overall Results

	May							June						
	SJV (Mg/hr)	D1 (post) (Mg/hr)	D2 (post) (Mg/hr)	r2 (prior)	r2 (post)	bias (prior) (ppbv)	bias (post) (ppbv)	SJV (Mg/hr)	D1 (post) (Mg/hr)	D2 (post) (Mg/hr)	r2 (prior)	r2 (post)	bias (prior) (ppbv)	bias (post) (ppbv)
WRF1	142±20	81±15	61±13	0.38	0.76	-60.6	-10.0	143±19	93±15	50±12	0.47	0.70	-35.7	-3.8
WRF2	156±22	88±17	68±14	0.38	0.69	-62.7	-10.4	129±18	70±12	59±13	0.33	0.60	-31.7	-4.5
WRF3	108±16	71±14	37±8	0.42	0.75	-49.9	-7.8	134±17	75±13	59±12	0.37	0.77	-31.8	-3.4
Overall[a]	135±28	80±17	55±18					135±19	79±17	56±13				
Overall[a,b]	149±22	84±17	65±14					136±20	81±18	55±13				

a. For each inversion ($X_i \pm \sigma_i$), we randomly select 10,000 values from the data range of $X \sim \mathcal{N}(X_i, \sigma_i)$. The overall estimate is the mean of all 30,000 (20,000) selected values from the three (or two) inversions and the associated uncertainty is the standard deviation of these values.

b. Including WRF1 and WRF2 simulations only, because WRF3 had a large bias in simulating PBLH in D2 in the May inversion case (see Table S1).

Table 4. Mass-Balance Inputs for the Northern San Joaquin Valley

Northern SJV Transect(s)	Terrain Ht. (m ASL)	Adjusted Mixing Ht. (m ASL)	Wind Direction (degrees)	Wind Speed (m/s)	Estimated CH_4 background (ppb)	CH_4 flux (10^{26} molec./s)	CH_4 flux (Mg/hr)
Upwind average	41 ± 41	1194 ± 243	299 ± 18	4.6 ± 2.0	1900 ± 5	2.9 ± 1.4	28 ± 19
downwind	89 ± 89	1361 ± 271	330 ± 21	6.1 ± 2.5	1900 ± 7	10.1 ± 4.7	97 ± 45

Table 5. Prior and Optimized CH_4 Emissions from two Major Source Sectors and Their Contributions to the San Joaquin Valley.

	Livestock						Oil/Gas			
	Prior (Mg/hr)	Inversion (Mg/hr)		Contribution		Prior (Mg/hr)	Inversion (Mg/hr)		Contribution	
		May	June	May	June		May	June	May	June
This study[a]	57	103±29	105±25	75%	77%	14	24±11	21±7	18%	15%
This study[a,b]		114±28	106±26	83%	77%		26±12	21±7	19%	15%
Jeong et al., [2016]		81-177		86%		Jeong et al., [2016]	19		11-19%	

a. The calculations of the final estimates are the same as Table 3.
b. Including WRF1 and WRF2 simulations only, because WRF3 had a large bias in simulating PBLH in D2 in the May inversion case (see Table S1).

Figure 1. (A) The San Joaquin Valley (SJV) and two sub-regions, the Southern SJV (D1) and the Northern SJV (D2). The background map is the prior inventory of CH_4 emissions used in this study based on CALGEM, showing the annual average emissions rate (unit: μg s^{-1} m^{-2}). (B) The spatial distribution of the two major CH_4 sources in the SJV: livestock and active oil/gas wells. (C) Two NOAA P-3 flight tracks over the SJV in May 2010. The black rectangles highlight the locations of the upwind transect in San Joaquin County and the downwind transect in Merced County used in the mass-balance estimate. (D) Two NOAA P-3 flight tracks over the SJV in June 2010.

Figure 2. Airborne measurements of CH_4 mixing ratios (averaged over 30 s) in the San Joaquin Valley, at 0-1500 m ASL and excluding measurements taken over the ocean and during takeoff and landing from the Los Angeles area. Each data point represents a receptor for the inverse modeling.

Figure 3. Surface footprints calculated by FLEXPART for the previous 72 hrs with 3 different WRF configurations and averaged for the two May flights (top row) and for the two June flights (bottom row). The surface footprints (unit: s m^2 kg^{-1}) represent the sensitivity of the airborne measurements (Figure 2) to surface emissions. Different scales are used for the footprints in the May and June cases to improve visualization.

Figure 4. Vertical profiles of 100-m averaged measurements of CH_4 enhancement mixing ratios, ΔCH_4, (measured mixing ratios in Figure 2 above a background derived for each flight; see text for details), simulations of Δch4 from FLEXPART-WRF using the prior and optimized emission estimates in the San Joaquin Valley for May (left) and June (right) 2010. The error bars represent the standard deviations (1-sigma) of simulations from the three different transport models.

Figure 5. Two-dimensional maps of CH_4 emissions estimates in the San Joaquin Valley from this study. (A) and (C) are average optimized emissions using the airborne measurements from two May flights and two June flights, respectively. (B) and (D) are the corresponding differences between the optimized emissions estimates and the prior emission inventory in Figure 1(A).

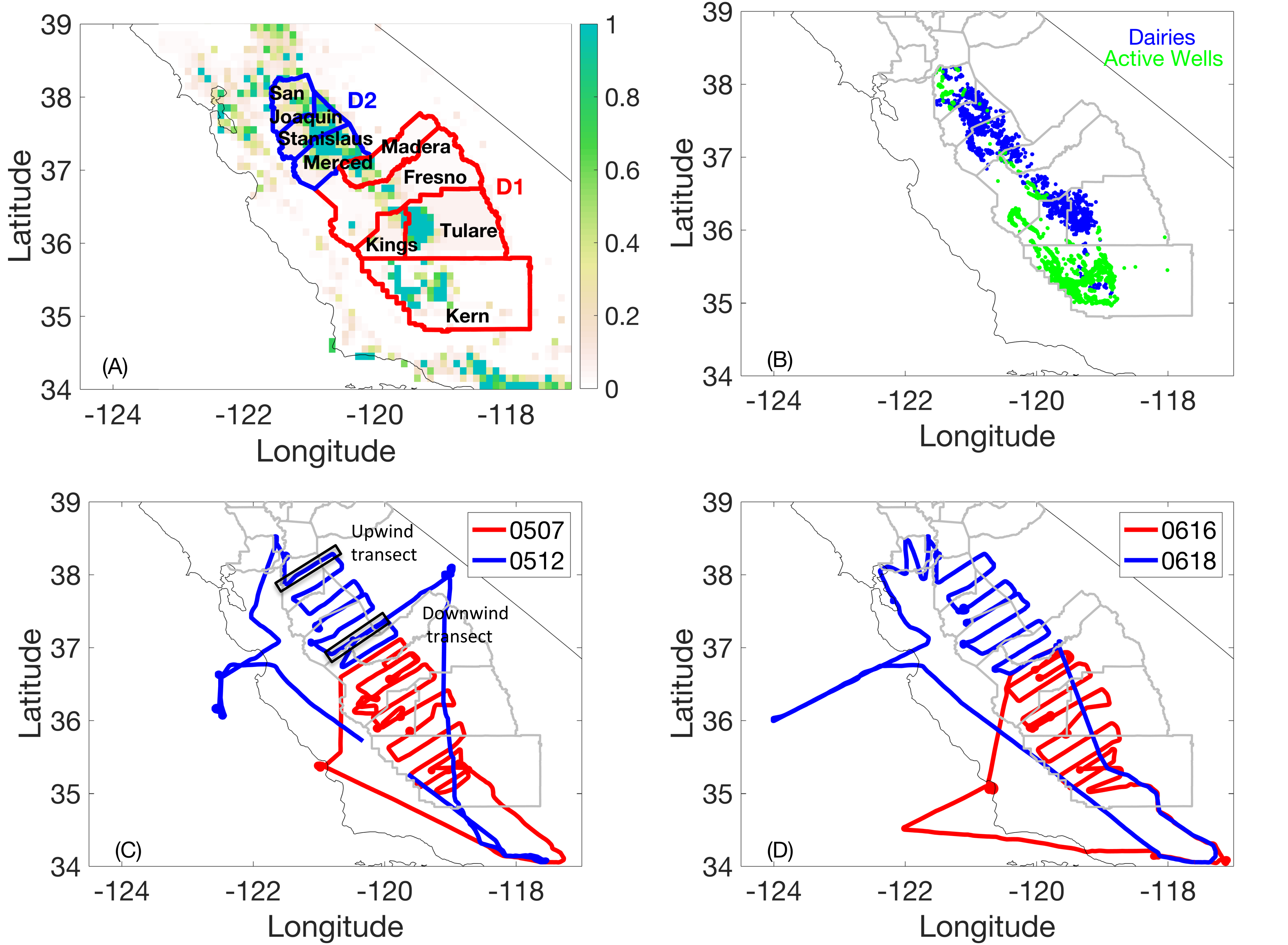
Figure 6. Airborne measurements of CH_4 enhancement mixing ratios, ΔCH_4, (measured mixing ratios in Figure 2 above a background derived for each flight; see text for details) (black line), simulations of ΔCH_4 from FLEXPART-WRF based on the prior inventory (blue lines), and simulations from FLEXPART-WRF based on the optimized emissions (red lines). Solid lines are average values based on the three transport models, and shading represents the standard deviation (1-sigma) of three transport models.

Figure 7. The relationship between observed and simulated CH_4 enhancement mixing ratios for the May (left) and June (right) flights. The simulated data points

are average values based on three transport models (the solid lines in Figure 6). The lines indicate the least squares fits to the data. We show correlations between observations and simulations with either the optimized emissions (red) or the prior inventory (blue). All correlations are significant with $P < 0.05$.

Figure 8. CH_4 enhancement mixing ratios simulated by the FLEXPART-WRF model based on the optimized CH_4 emissions from the whole domain (All, green lines) and due to CH_4 emissions from only one specific sub-region (either D1 or D2). Flights 0507 and 0616 mainly flew over D1, but were impacted by air masses from D2. Flights 0512 and 0618 mainly flew over D2 and were rarely impacted by air masses from D1. The percentages shown in the titles represent the contributions of emissions from this other sub-region (D1 or D2) to the overall airborne measurements of CH_4 mixing ratios in each flight.

Figure 1.



(A)
San Joaquin
Stanislaus
Merced
Madera
Fresno
Kings
Tulare
Kern
D1
D2
Latitude
Longitude
(B)
Dairies
Active Wells
(C)
0507
0512
Upwind transect
Downwind transect
(D)
0616
0618

Figure 2.

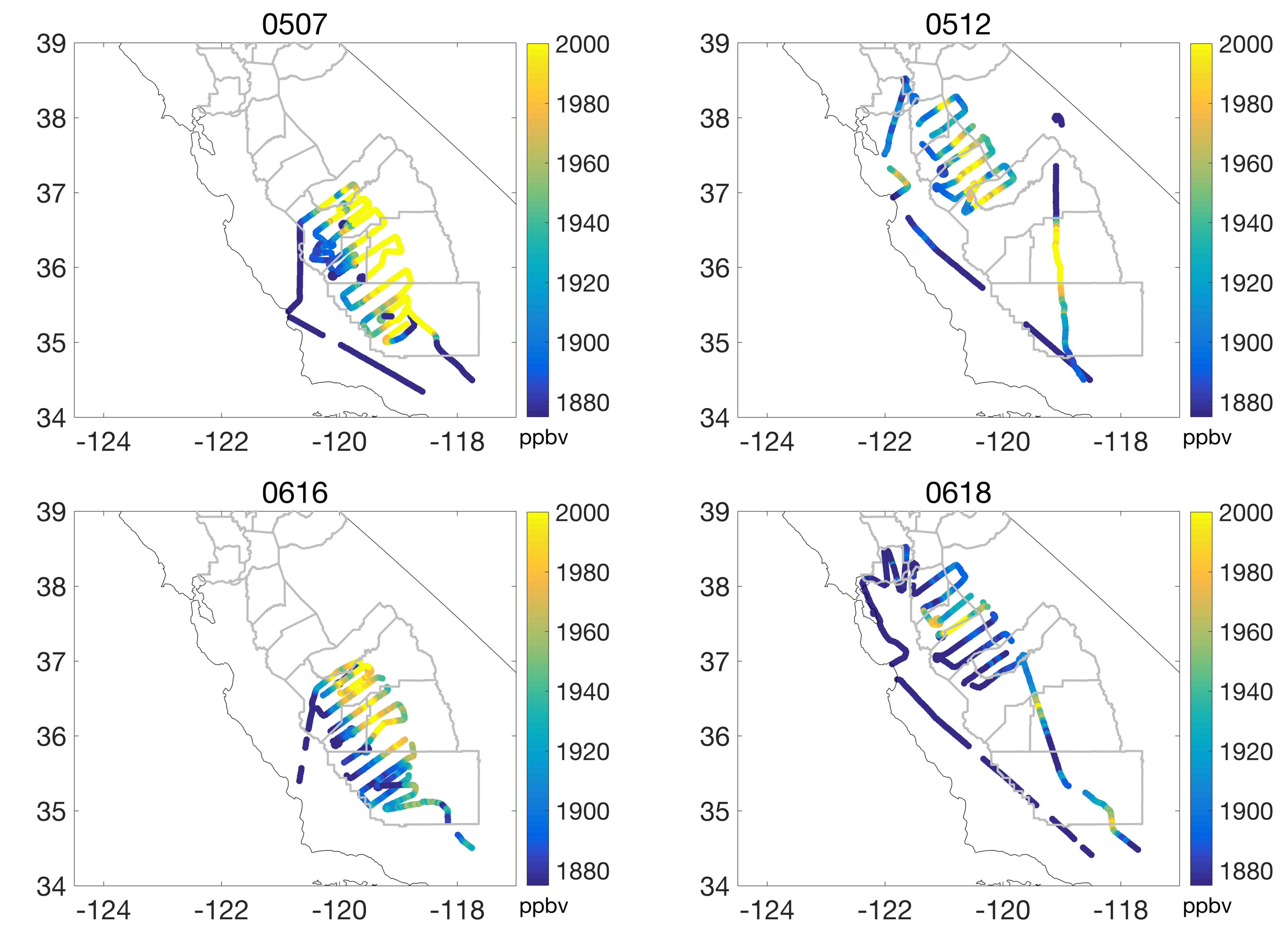

0507
0512
0616
0618
39
38
37
36
35
34
-124
-122
-120
-118
2000
1980
1960
1940
1920
1900
1880
ppbv

Figure 3.

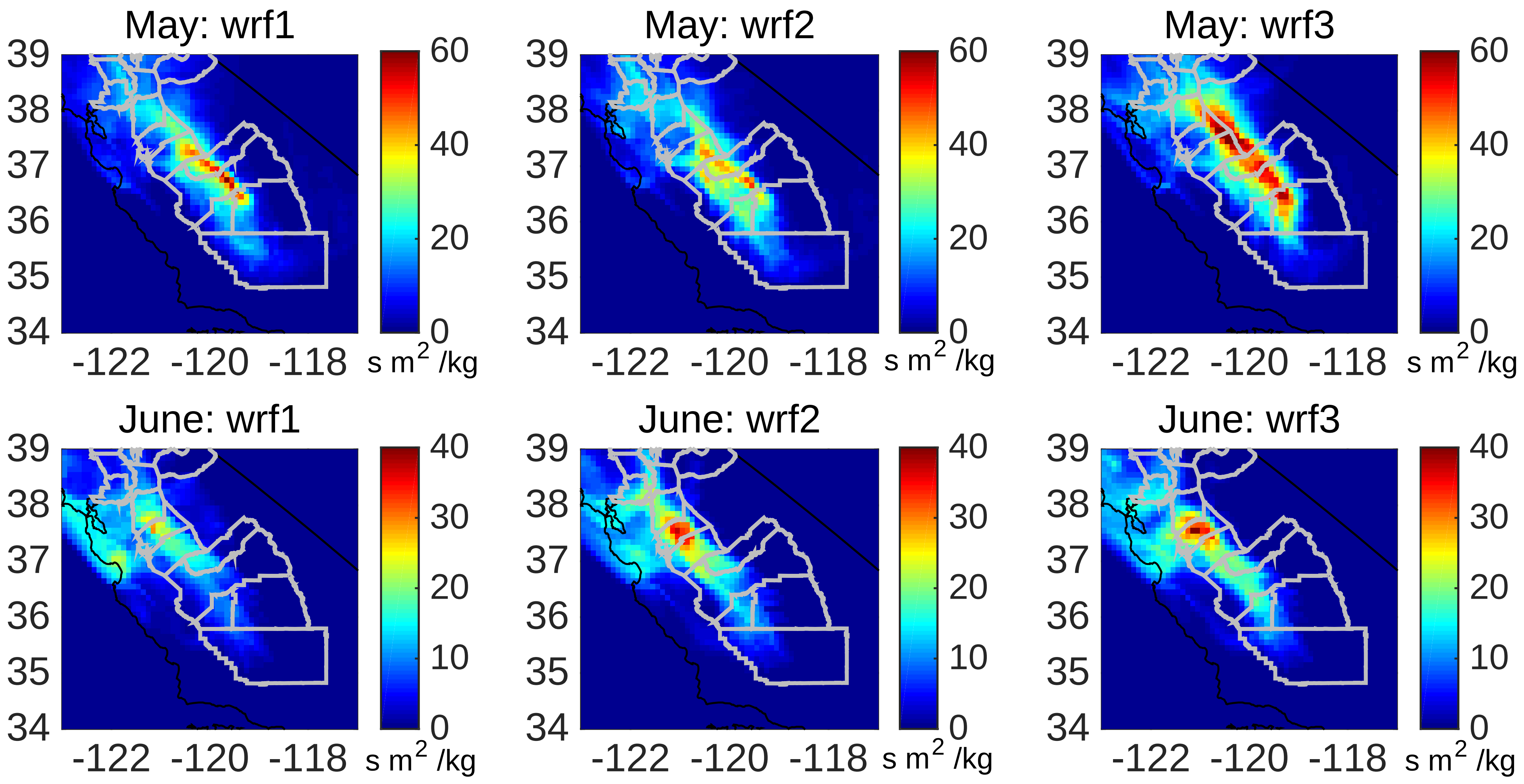

May: wrf1
May: wrf2
May: wrf3
June: wrf1
June: wrf2
June: wrf3
39
38
37
36
35
34
-122
-120
-118
60
40
20
0
30
10
s m2 /kg

Figure 4.

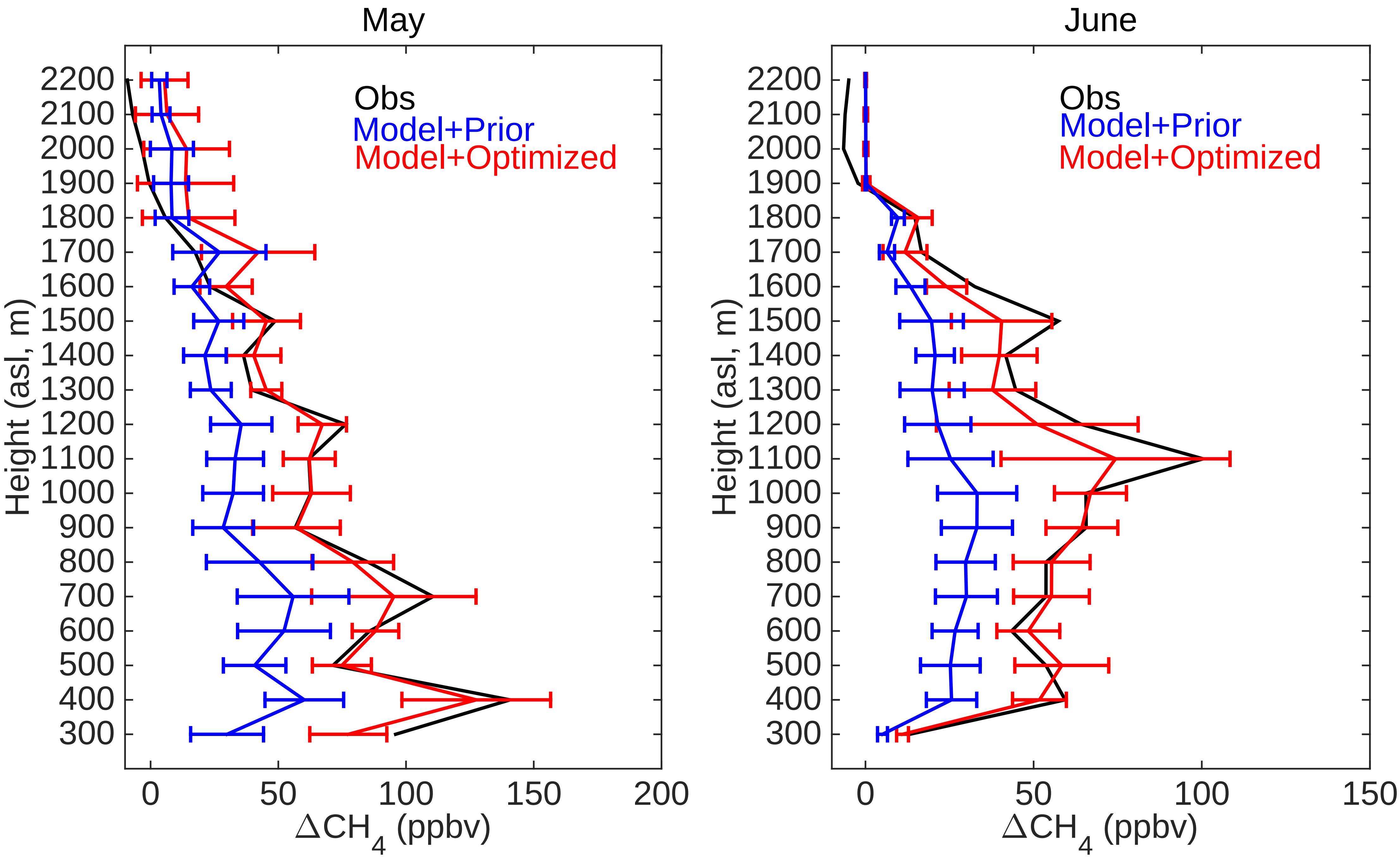
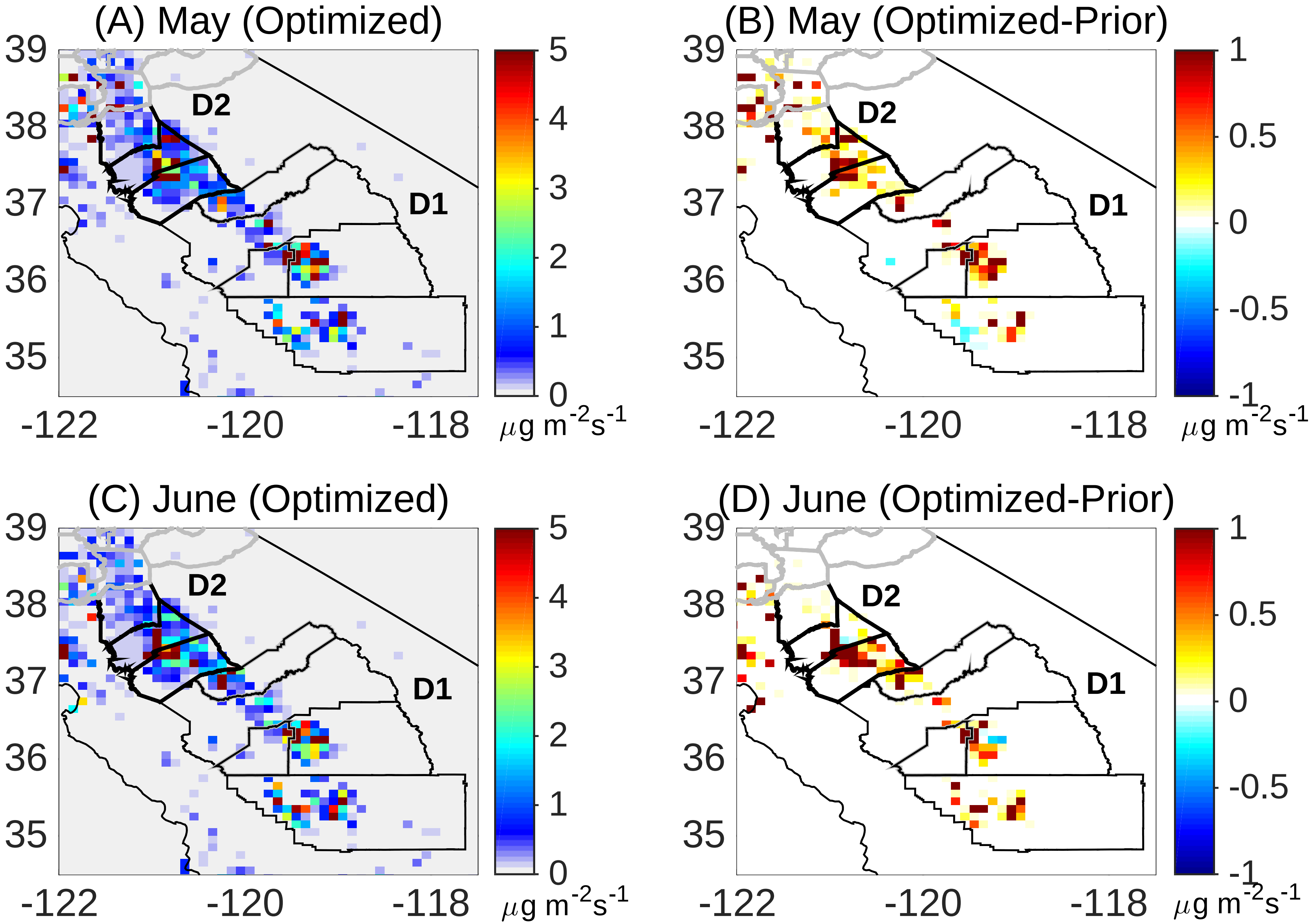

May
June
Obs
Model+Prior
Model+Optimized
Height (asl, m)
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
0
50
100
150
200
ΔCH4 (ppbv)

Figure 5.



(A) May (Optimized)
(B) May (Optimized-Prior)
(C) June (Optimized)
(D) June (Optimized-Prior)
D1
D2
39
38
37
36
35
-122
-120
-118
5
4
3
2
1
0
0.5
-0.5
-1
μg m-2s-1

Figure 6.

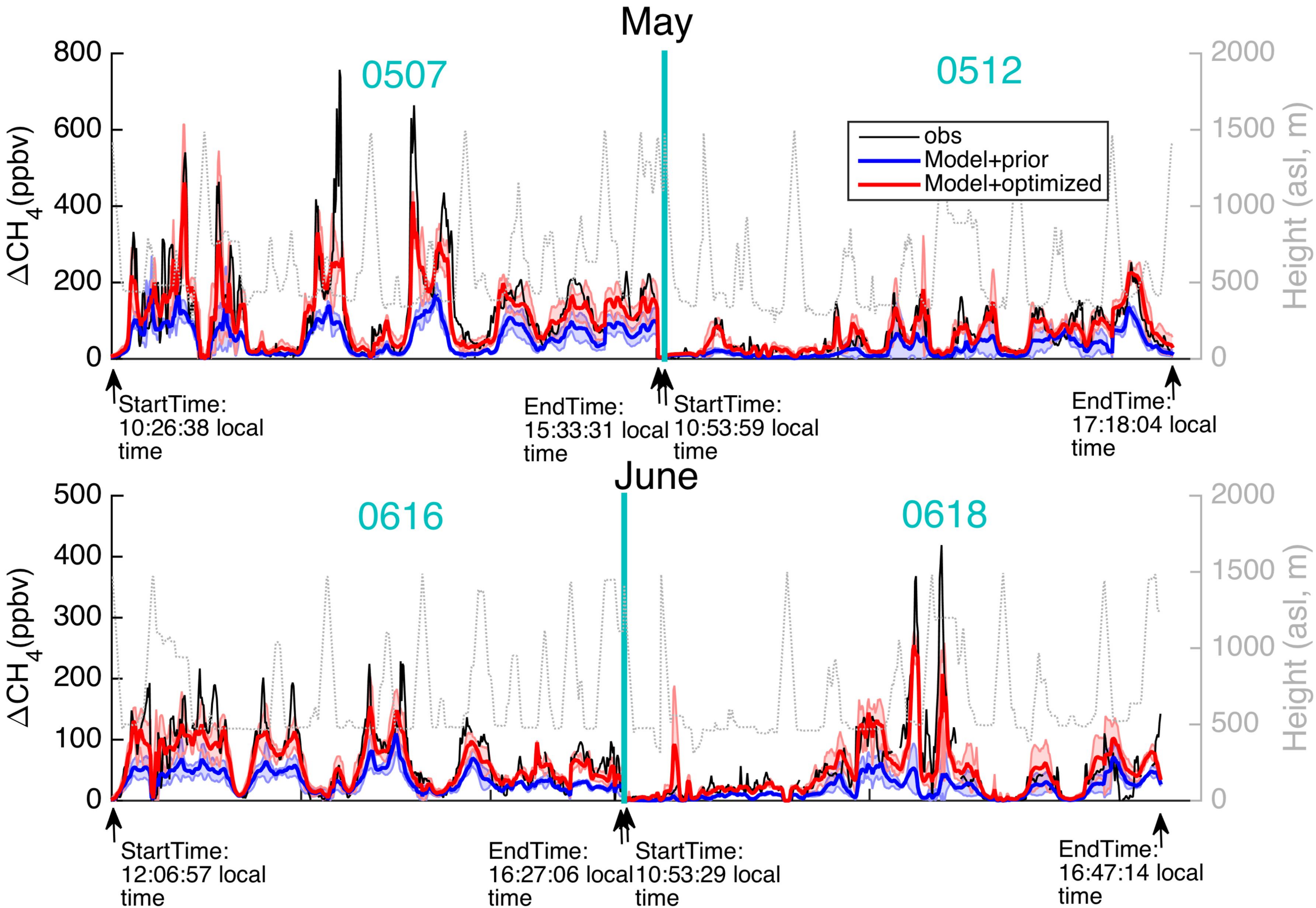
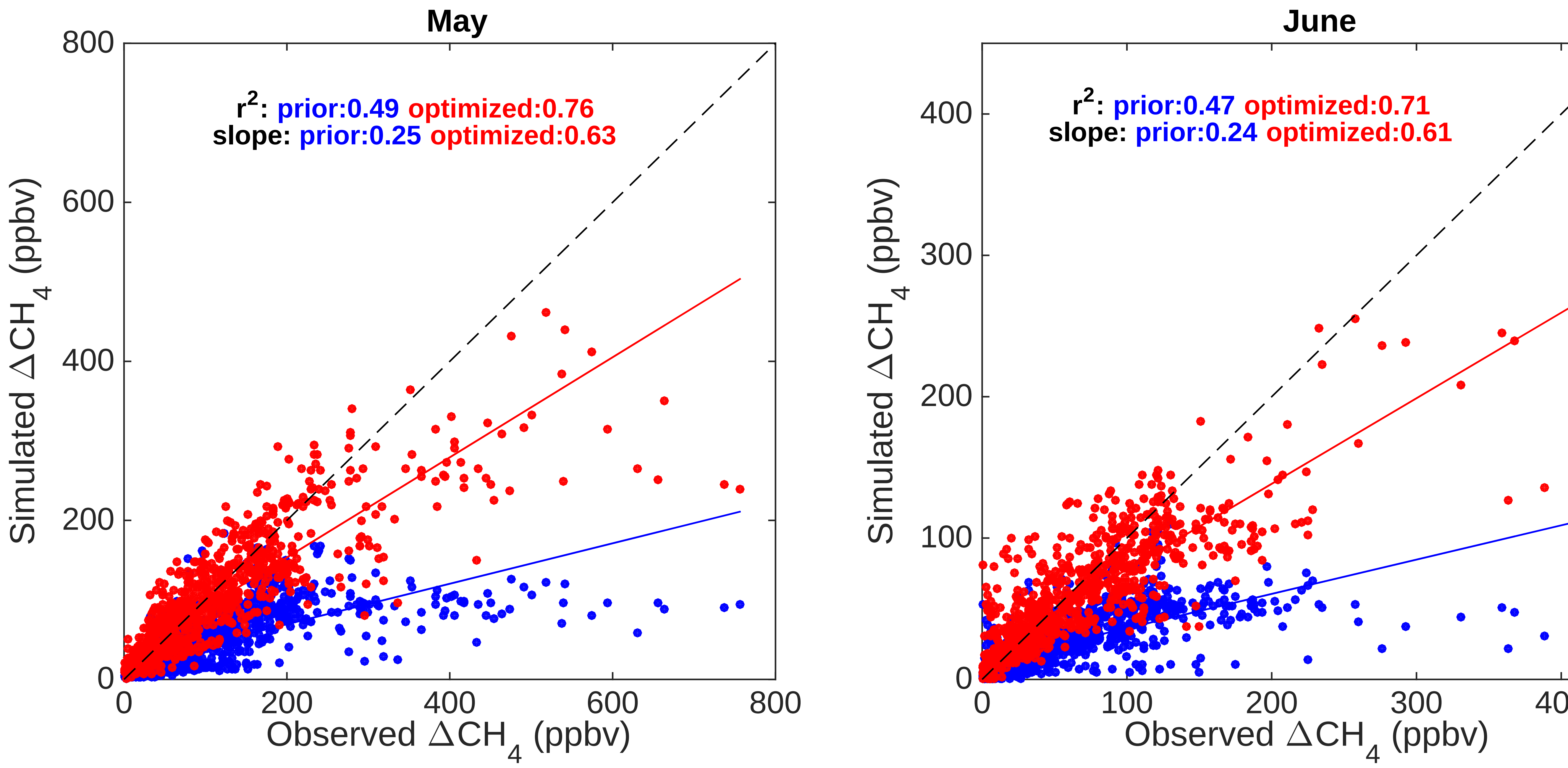
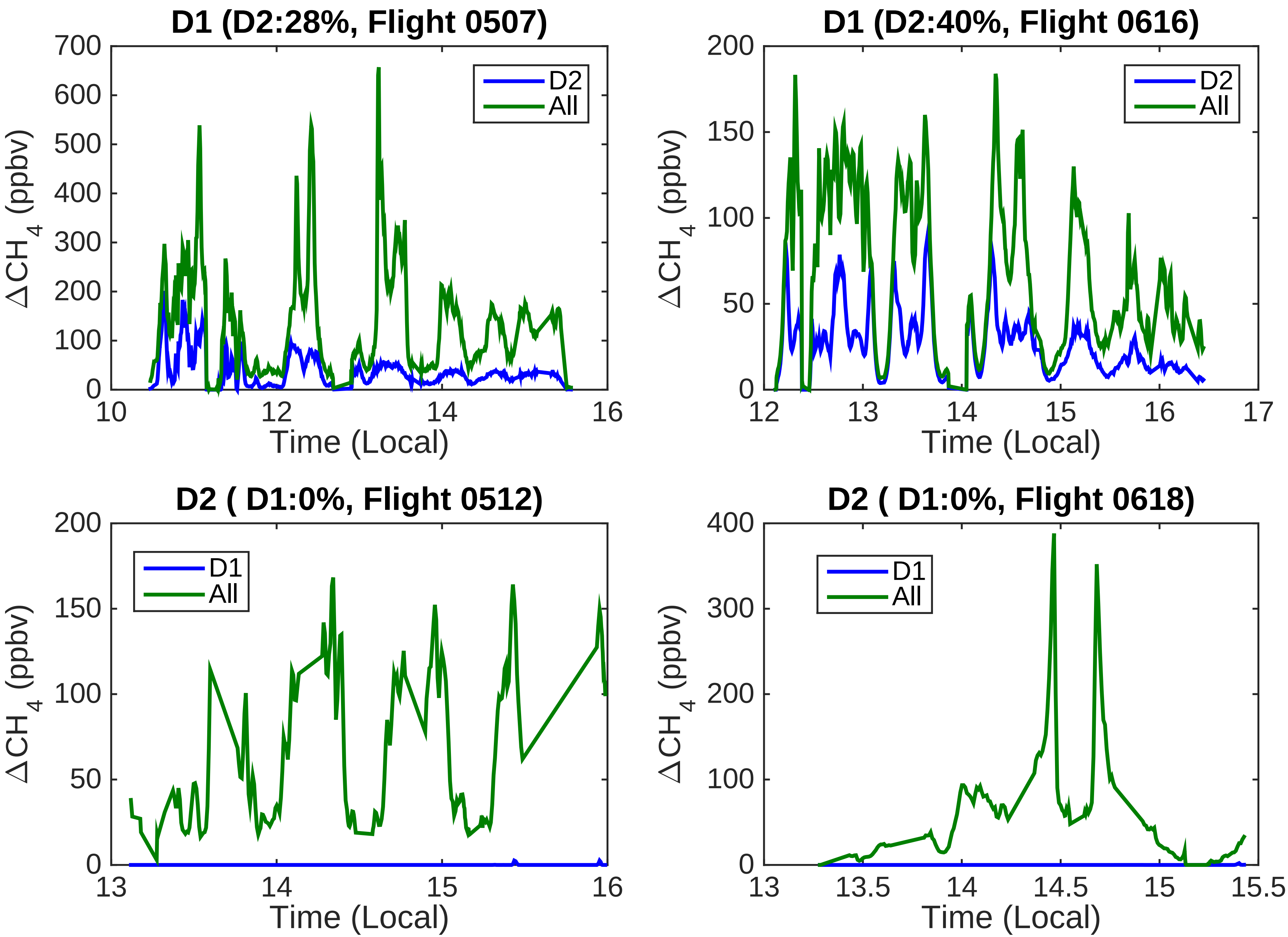

May
0507
0512
obs
Model+prior
Model+optimized
ΔCH4(ppbv)
Height (asl, m)
StartTime:
10:26:38 local
time
EndTime:
15:33:31 local
time
StartTime:
10:53:59 local
time
EndTime:
17:18:04 local
time
June
0616
0618
StartTime:
12:06:57 local
time
EndTime:
16:27:06 local
time
StartTime:
10:53:29 local
time
EndTime:
16:47:14 local
time

Figure 7.



May
r^2: prior:0.49 optimized:0.76
slope: prior:0.25 optimized:0.63
Simulated ΔCH_4 (ppbv)
Observed ΔCH_4 (ppbv)
0
200
400
600
800
June
r^2: prior:0.47 optimized:0.71
slope: prior:0.24 optimized:0.61
Simulated ΔCH_4 (ppbv)
Observed ΔCH_4 (ppbv)
0
100
200
300
400

Figure 8.



D1 (D2:28%, Flight 0507)
D2
All
ΔCH4 (ppbv)
Time (Local)
D1 (D2:40%, Flight 0616)
D2
All
ΔCH4 (ppbv)
Time (Local)
D2 (D1:0%, Flight 0512)
D1
All
ΔCH4 (ppbv)
Time (Local)
D2 (D1:0%, Flight 0618)
D1
All
ΔCH4 (ppbv)
Time (Local)