A Multi-tower Measurement Network Estimate of California's Methane Emissions

Seongeun Jeong[1], Ying-Kuang Hsu[2], Arlyn E. Andrews[3], Laura Bianco[3,4], Patrick Vaca[2], James M. Wilczak[3], Marc L. Fischer[1]

[1]Environmental Energy Technologies Division, Lawrence Berkeley National Lab, Berkeley, CA, USA

[2]California Air Resources Board, 1001 "I" Street, Sacramento, CA, USA

[3]Earth System Research Laboratory, NOAA, Boulder, CO, USA

[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

Corresponding author:

Seongeun Jeong

MS 90R2002

1 Cyclotron Rd.

Berkeley, CA 94720 USA

phone: 510-486-7216

email: sjeong@lbl.gov

Key Points:

- Multi-site observations constrain Central Valley CH_4 emissions
- California total emissions are likely 1.3 -1.8 times the state inventory
- Additional measurements will help guide CH_4 mitigation activities

Abstract

We present an analysis of methane (CH_4) emissions using atmospheric observations from five sites in California's Central Valley across different seasons (September 2010 to June 2011). CH_4 emissions for spatial regions and source sectors are estimated by comparing measured CH_4 mixing ratios with transport model (WRF-STILT) predictions based on two 0.1 degree CH_4 (seasonally varying "California-specific" (CALGEM) and a static global (EDGAR42)) prior emission models. Region-specific Bayesian analyses indicate that for California's Central Valley the CALGEM- and EDGAR42-based inversions provide consistent annual total CH_4 emissions (32.87±2.09 vs. 31.60±2.17 Tg CO_2eq yr^{-1}; 68% C.I., assuming uncorrelated errors between regions). Summing across all regions of California, optimized CH_4 emissions are only marginally consistent between CALGEM- and EDGAR42-based inversions (48.35±6.47 vs. 64.97±11.85 Tg CO_2eq), because emissions from coastal urban regions (where landfill and natural gas emissions are much higher in EDGAR than CALGEM) are not strongly constrained by the measurements. Combining our results with those from a recent study of the South Coast air basin narrows the range of estimates to 43 – 57 Tg CO_2eq yr^{-1} (1.3 - 1.8 times higher than the current state inventory). These results suggest that the combination of rural and urban measurements will be necessary to verify future changes in California's total CH_4 emissions.

Keywords: methane, greenhouse gas, emission inventory, atmospheric transport, inverse model

Index Terms: 0365, 0345, 0368

1. Introduction

Methane (CH_4) is the second highest contributor to climate change among greenhouse gases (GHGs) behind carbon dioxide (CO_2), based on its concentration changes in the atmosphere since the start of the industrial revolution, the long residence time of CH_4 and its ability to absorb infrared radiation. Atmospheric CH_4 levels have increased by about 150% since 1750 accounting for ~ 25% of the global total radiative forcing from all long-lived and globally mixed GHGs [Hofman et al., 2006; Montzka et al., 2011]. Given the significance of CH_4 as a GHG it is

important to be able to quantify changes in emissions. However, bottom-up emission inventory models are highly uncertain due to lack of driver data and incomplete understanding of emission processes. Atmospheric inverse modeling, which uses observed concentration changes in CH_4 to infer sources, potentially provides an effective tool for understanding CH_4 emissions [Houweling et al., 1999; Gimson and Uliasz, 2003; Kort et al., 2008; Zhao et al., 2009; Jeong et al., 2012a].

California currently emits approximately 500 Tg of CO_2eq GHGs, with CH_4 estimated to contribute ~6% of the total [California Air Resources Board (CARB), 2011]. Because California has committed to an ambitious plan to reduce GHG emissions to 1990 levels by 2020 through Assembly Bill 32 (AB-32), planning effective mitigation efforts and verifying future emission reductions require accurate accounting of CH_4 emissions.

This paper quantifies regional CH_4 emissions from California within a Bayesian inverse modeling framework, representing the first analysis of CH_4 emissions in California using atmospheric observations from multiple sites across different seasons (September 2010 – June 2011). The work expands on studies by Zhao et al. [2009] and Jeong et al. [2012a] that quantified CH_4 emissions from central California using a single tower near Walnut Grove, California (WGC) by combining measurements from the additional sites in the Central Valley and including published emission estimates from the Los Angeles metropolitan area to capture emissions from California's urban regions. In Section 2, we describe the methods we employed, including atmospheric measurements, *a priori* CH_4 emissions inventories, meteorology and trajectory transport modeling, and the Bayesian inverse method. Section 3 presents results, including seasonal variation in footprints, and the inferred surface emissions of CH_4 from California for different regions and sources. Section 4 further discusses the results and presents conclusions for CH_4 emissions in California.

2. Data and Models

2.1. CH_4 Measurements and Boundary Conditions

CH_4 measurements were made at the collaborative five-site GHG network in California's Central Valley during September 2010 – June 2011: Arvin (35.24°N, 118.79°W; ARV), Madera (36.87°N, 120.01°W; MAD), Tranquility (36.63°N, 120.38°W; TRA), Sutter Buttes (39.21°N, 121.82°W; STB), and WGC (38.27°N, 121.49°W) (STB measurements are available only for May – June 2011). CH_4 measurements at WGC were made at 91 and 483 m above ground level on a tall tower, beginning in September 2007 [Andrews et al., 2013]. CH_4 measurements at 91 m were used for inverse modeling, and additional information about these measurements is provided by Zhao et al. [2009] and Jeong et al. [2012a]. All other stations measured CH_4 at ~10 m above the ground using Picarro model 2301 analyzers that were calibrated with standard gases from NOAA every six months and programmed to measure a standard gas every 11 hours in order to check the precision. After examining precision checks and removing special events (e.g., changing filters), raw data collected every few seconds are averaged into 3-hourly measurements for inverse modeling. We apply data filtering based on vertical mixing to data from WGC where vertical CH_4 profiles are available. As in Jeong et al. [2012a], data were selected such that the CH_4 mixing ratio difference ($C_{91} - C_{483}$) between 91 and 483 m fell within the range -1 sd $< (C_{91} - C_{483}) < 3$ sd, where sd is the standard deviation of the difference of the mean diurnal cycle between 1200 and 1700 local standard time (LST). For other sites, we use afternoon data (1200 – 1700 LST) when boundary layers are reasonably well developed in the Central Valley [Bianco et al., 2011]. As will be described in Section 3.1, for winter we use data during 1100 – 1600 LST due to earlier collapse of the boundary layer in simulations than in measurements. In Section 3.1, we also report results of a sensitivity test to periods with potentially low simulated boundary

layers that suggests our posterior emissions estimates are not significantly affected by inadequate mixing.

CH_4 boundary values were estimated using data from the NOAA Earth System Research Laboratory's Global Monitoring Division (ESRL/GMD) using an approach similar to the one used in Jeong et al. [2012b]. Marine boundary layer data from the Cooperative Air Sampling Network (http://www.esrl.noaa.gov/gmd/ccgg/flask.html) and vertical profile data from aircraft (http://www.esrl.noaa.gov/gmd/ccgg/aircraft/) were used to create a smoothed three-dimensional (3-D) curtain representing the Pacific boundary and varying with latitude, height and time. The NOAA aircraft data are primarily collected over North America and along the Pacific Coast. Since data along the coasts are sparse and the impact of surface fluxes on free tropospheric data is small, we have used all available aircraft data in our estimate. We ran back trajectories for all aircraft observations using the National Centers for Environmental Prediction (NCEP) reanalysis wind fields (global) and removed any observations for which the trajectories drop below 3 km above ground level. We defined a domain following the coast of North America and identified the latitude and altitude when the trajectory exits the domain. We also used aircraft data from Hawaii, which is outside the North American domain, and in that case the actual latitude and altitude of the observation were used. Data are binned according to the latitude and altitude where they exit the domain (10° latitude resolution over 20° - 70°, 1000 m vertical resolution, 3000 - 7000 m.a.s.l.). We then fit smooth curves to the binned data using the method of Thoning et al. [1989] (see also http://www.esrl.noaa.gov/gmd/ccgg/mbl/crvfit/crvfit.html). For altitudes below 1000 m.a.s.l., we use a Pacific version of the NOAA Greenhouse Gas Marine Boundary Layer Reference (http://www.esrl.noaa.gov/gmd/ccgg/mbl/) that is based on surface observations from the Cooperative Air Sampling Network and which varies with latitude and time. In the

range 1000 - 3000 m.a.s.l., values are interpolated between the Pacific Marine Boundary Layer Reference and the free-tropospheric curtain derived from the aircraft data. Time-varying uncertainty in the boundary curtain is estimated using the seasonal cycle of the root mean square of the residuals from the smoothed-curves. Average background values are computed for each footprint simulation by sampling the curtain at each of the 500 particle trajectory endpoints (near the domain boundary at 130°W) and calculating the average values. Uncertainty in the estimated background values is discussed in Section 2.5.

2.2. *A priori* CH_4 Emission Models

This work adopts the California Greenhouse Gas Emission Measurements (CALGEM) project *a priori* CH_4 emission model (henceforth CALGEM model) described by Jeong et al. [2012a], which is provided at a high spatial resolution (0.1º × 0.1º) for California and has seasonal components for wetlands and crop agriculture [CALGEM, 2013]. Table 1 provides CALGEM emissions used in this study by source and region, which include emissions from rice agriculture and wetlands (see Figure 1 for regions). Here the high-resolution emissions were scaled to match the CARB inventory for 2008 by sector (summing to a total of 28 Tg CO_2eq for California) [CARB, 2010]. The EDGAR42 (European Commission Joint Research Centre (JRC) and Netherlands Environmental Assessment Agency, Emission Database for Global Atmospheric Research (EDGAR), release version 4.2, 2011, http://edgar.jrc.ec.europa.eu) CH_4 emission model (annual total = 38 Tg CO_2eq or 1.4 times CALGEM total) also provides high-resolution (0.1º × 0.1º) emission maps. Table 2 shows all 16 emission source sectors from the EDGAR42 prior emission model by region, which can be compared with the CALGEM model shown in

Table 1. Bayesian inversions adjust region sums (region analysis) or source sums (source analysis) shown in Tables 1 and 2 to yield optimized (posterior) emissions.

Figure 1 shows the annual total emission maps for the CALGEM and EDGAR42 prior models along with the sub-region classification for inverse modeling. Compared with the California-specific CALGEM model, EDGAR42 generally shows a similar spatial distribution of CH_4 emissions. The CALGEM model estimates higher total emissions for the Central Valley (Regions 6, 8 and 12) than EDGAR42, mainly due to the higher estimates of dairy emissions. As shown in Tables 1 and 2, for Regions 7 and 10, which include the San Francisco Bay Area and the Southern California region, respectively, the EDGAR42 model estimates significantly higher CH_4 emissions than the CALGEM model.

Because there is no specific emission estimate for wetlands from CARB, wetland CH_4 emissions (not included in EDGAR42) for the CALGEM prior emission model were taken from monthly averages of the Carnegie-Ames-Stanford-Approach CH_4 (CASA-CH_4) model from Potter et al. [2006]. Also, seasonally varying (monthly) CH_4 emissions for crop agriculture were taken from the DeNitrification-DeComposition (DNDC) model output (assuming the 1983, high irrigation case) described by Salas et al. [2006]. The crop agriculture sector was scaled to the CARB 2008 inventory (0.54 Tg CO_2eq yr^{-1}) using the seasonal pattern from DNDC. As shown in Table 2, EDGAR42 provides an emission sector for agriculture (i.e., agricultural soils).

2.3. Atmospheric Transport Modeling

We use the coupled WRF-STILT (Weather Research and Forecasting and Stochastic Time-Inverted Lagrangian Transport) model for particle trajectory simulations [Lin et al., 2003; Skamarock et al., 2008; Nehrkorn et al., 2010]. The WRF-STILT model has been used to

constrain GHG emissions in many studies including airborne measurement-based (e.g., Gerbig et al., 2003; Kort et al., 2008) and tower measurement-based (e.g., Zhao et al., 2009; Jeong et al., 2012a; Jeong et al., 2012b) inversions. An ensemble of 500 STILT particles are run backwards in time for 7 days driven with meteorology from the WRF model (version 3.2.1) [Skamarock et al., 2008]. Hourly predicted mixing ratios based on WRF-STILT are aggregated into 3-hourly averages for inverse modeling.

The WRF model simulations closely follow those described in Jeong et al. [2012a, 2012b] with some modifications, which are summarized here. We use version 3.2.1 of the WRF model [Skamarock et al., 2008] instead of WRF2.2. Five domains (d01 – d05) of 36, 12, 4, and two 1.3 km resolutions were used in the WRF simulations. The 4-km domain (i.e., d03) was configured to represent most of California with the two 1.3-km nested domains (d04 and d05) that cover the San Francisco Bay Area and the metropolitan area of Los Angeles, respectively. In this study, we used the WRF meteorology within the d01, d02 and d03 domains to drive the STILT model because the GHG measurement sites are located in the Central Valley. The WRF model was run with two-way nesting instead of one-way nesting used in Jeong et al. [2012a, 2012b]. As in Jeong et al. [2012a, 2012b], 50 vertical levels were employed to resolve planetary boundary layer (PBL) heights over complex terrain features of California. Initial and boundary meteorological conditions were provided by the North American Regional Reanalysis (NARR) dataset [Mesinger et al., 2006]. All simulation durations were 30 hours including 6 hours of model spin up. The model also incorporated 3-D analysis nudging every three hours in the 36-km domain.

As an extension beyond the previous work, we ran the WRF model multiple times to evaluate different combinations of surface model and boundary layer schemes. The specific

combination of land surface models (LSMs) and PBL schemes that yielded the best comparison with PBL heights retrieved from the wind profilers [Bianco and Wilczak, 2002; Bianco et al., 2008] in the Central Valley varied with season and location. Here, we evaluated the WRF meteorology using data for the Sacramento (SAC), Chowchilla (CCL), Chico (CCO) and Lost Hills (LHS) sites shown in Figure 2 (see Section 2.5 for details on evaluation). For late spring through early fall, the combination of the five-layer thermal diffusion LSM (5-L LSM hereafter) and the Mellor-Yamada-Janjic (MYJ) PBL scheme [Mellor and Yamada, 1982; Janjić, 1990] performed best. For example, for the summer month of June 2010 (due to profiler data availability, 2010 data are used for some sites and months), the 5-L LSM and MYJ combination (root mean square (RMS) errors = 280 - 290 m) performed better than the Noah LSM and MYJ combination (RMS errors = 400 - 450 m) for the SAC and CCL sites. This is likely due to the fact that the 5-L LSM actively manages soil moisture as a function of season and land cover types that include irrigated soils. Thus, we use the 5-L LSM during the months of April – September that were identified as the period of the year with strong evapotranspiration (California Irrigation Management Information System, http://wwwcimis.water.ca.gov/cimis/data.jsp). The 5-L LSM scheme uses a fixed season-dependent value for the irrigated soil (i.e., irrigated cropland and pasture category) to generate an accurate boundary condition for soil moisture and hence energy balance. The one exception is that of the LHS site during late spring - early fall where the 5-L LSM and Yonsei University (YSU) PBL scheme combination performed better than the 5-L LSM and MYJ combination. For example, for June 2010, the RMS error for the 5-L LSM and MYJ combination (526 m) was significantly larger than that of the 5-L LSM and YSU combination (359 m). We speculate that the 5-L LSM may overestimate soil moisture at the LHS site, reducing PBL height in a manner

that is compensated for by the overestimation of PBL height by the YSU scheme. However, we lack the data to test this hypothesis at this time. For late fall through early spring, the Noah LSM and MYJ combination performed well because the more complicated Noah LSM handles the energy balance better when precipitation is the dominant source of moisture.

2.4. Bayesian Inverse Model

The inversion approach expands on earlier efforts by Zhao et al. [2009] and Jeong et al. [2012a; 2012b], and we express the model-measurement relation through a linear model:

$$\mathbf{c} = \mathbf{K}\boldsymbol{\lambda} + \mathbf{v}, \qquad (1)$$

where $\mathbf{c}$ is the measurement vector ($n \times 1$, n = number of measurements), which represents 3-hour mean, background-subtracted CH_4 mixing ratios, $\mathbf{K} = \mathbf{FE}$ (an $n \times k$ matrix, k is the number of regions or sources), $\mathbf{F}$ is the footprint ($n \times m$, m is the number of grid cells of 0.1° × 0.1°), $\mathbf{E}$ is emissions ($m \times k$), $\boldsymbol{\lambda}$ is a $k \times 1$ state vector for scaling factors, and $\mathbf{v}$ is a vector representing the model-data mismatch with a covariance matrix $\mathbf{R}$ ($n \times n$), i.e., $\mathbf{v} \sim N(\mathbf{0}, \mathbf{R})$ where N denotes the normal distribution. We model $\mathbf{R}$ as a diagonal matrix to represent the total variance associated with all error sources following Gerbig et al. [2003], Zhao et al. [2009], Göckede et al. [2010], and Jeong et al. [2012a; 2012b]. The uncertainty analysis, which constructs the $\mathbf{R}$ matrix, is presented in detail in the following section. Depending on the month and measurement site, we estimated the errors to be 20 - 233 ppb (~30 – 60% of the background-subtracted mean mixing ratio) to fill the diagonal elements of $\mathbf{R}$. Following the Gaussian assumptions, the posterior estimate for $\boldsymbol{\lambda}$ is

$$\boldsymbol{\lambda}_{post} = \left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{K} + \mathbf{Q}_{\lambda}^{-1}\right)^{-1}\left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{c} + \mathbf{Q}_{\lambda}^{-1}\boldsymbol{\lambda}_{prior}\right) \qquad (2)$$

where $\boldsymbol{\lambda}_{prior}$ is the *a priori* estimate for $\boldsymbol{\lambda}$ (initially set to one for all elements), and $\mathbf{Q}_{\lambda}$ is the error covariance matrix ($k \times k$) for $\boldsymbol{\lambda}$. The corresponding posterior covariance for $\boldsymbol{\lambda}$ is

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

We apply the inversion method at a monthly temporal scale solving for $\boldsymbol{\lambda}_{post}$ for each month. We relax our assumption on prior uncertainty to 70% from the 50% uncertainty used in Jeong et al. [2012a]. We use this relaxed prior uncertainty because this analysis estimates CH_4 emissions for a much larger region with a higher uncertainty than that (i.e., central California) of Jeong et al. [2012a]. The inverse modeling approach is applied in two phases as in Bergamaschi et al. [2005] and Jeong et al. [2012a; 2012b]. After a first inversion, the second (final) inversion uses data points that are accepted by applying the selection criteria $|\mathbf{c}_i - (\mathbf{K}\boldsymbol{\lambda})_i|^2 < \alpha \mathbf{R}_i$, where α is a fixed value ($\alpha = 3$). The outlier removal rates are 4.7 – 5.1% of a total of 1659 (i.e., total size of n) observations depending on the inverse analysis. As in the first inversion, the final inversion is performed using the original *a priori* emission maps, and therefore the first inversion is used as a data selection tool for the atmospheric observations.

2.5. Uncertainty Analysis

The uncertainty in the model-measurement differences controls the relative weighting of the prior flux estimates and the measured data in the inversion, adjusting posterior CH_4 emissions relative to *a priori* emissions. Following Gerbig et al. [2003], Zhao et al. [2009], Göckede et al. [2010], and Jeong et al. [2012a], the model-measurement mismatch matrix, $\mathbf{R}$ (an $n \times n$ matrix), is represented as the linear sum of uncertainties from several sources and modeled as a diagonal matrix:

$$\mathbf{R}_i = \mathbf{S}_{part} + \mathbf{S}_{aggr} + \mathbf{S}_{bkgd} + \mathbf{S}_{transPBL} + \mathbf{S}_{transWIND},$$

where the particle number error (S_{part}) is due to the finite number of released particles at the receptor location while the aggregation error (S_{aggr}) arises from aggregating heterogeneous fluxes within a grid cell into a single average flux. The background error (S_{bkgd}) is due to the uncertainty in estimating the background contribution to the CH_4 measurements at the receptor. $\mathbf{S}_{transWIND}$ and $\mathbf{S}_{transPBL}$ represent the uncertainty in CH_4 mixing ratios caused by the errors in wind speeds and directions, and the errors in PBL heights, respectively. For the aggregation error ($\mathbf{S}_{aggr}$), we adopt the result from Jeong et al. [2012a] and use 11% of the background-subtracted mean mixing ratio. The background error ($\mathbf{S}_{bkgd}$) is estimated by combining (in quadrature) the RMS error in the estimation of the 3-D curtain (similar to that used in Jeong et al. [2012b]) and the standard error of 500 WRF-STILT background samples. Average values for S_{bkgd} were calculated for each month during September 2010 – June 2011. Recall that for each simulation time, 500 particles are released from the measurement location and tracked backwards in time for 7 days, and each particle is associated with a background value at its final location. Each background value also has an uncertainty estimate that is the time-, height-, and latitude-dependent RMS error of the residuals of the data that were used to construct the background curtain. We compute the mean RMS error over the 500 particles for each observation. The background errors were estimated to be 17 – 25 ppb depending on the season and measurement site. Only observation time points for which more than 80% of the particles reached the western boundary of the domain (130°W) were included in the study (an average of ~85% retained after the filtering with summer having the highest of >95%).

To estimate the uncertainty in predicted CH_4 mixing ratios due to errors from modeled PBL heights ($\mathbf{S}_{transPBL}$) and winds ($\mathbf{S}_{transWIND}$), we evaluated WRF model errors in winds and PBL

heights and then calculated the RMS difference in CH_4 mixing ratios obtained from simulations with and without input of an additional stochastic component of wind and PBL errors in STILT. As described previously, we evaluated PBL heights (*Zi*) and winds at four profiler sites (Figure 2): CCO, SAC, CCL, and LHS. The radar wind profiler can retrieve data in two different modes (high and low resolutions) with vertical resolutions of 60 m and 105 m, respectively. PBL heights used in this study were estimated from sub-hourly vertical velocity and returned signal strength (signal-to-noise ratio) data using the algorithms and qualitative analysis following Wyngaard and LeMone [1980], Bianco and Wilczak [2002], and Bianco et al. [2008]. The wind profiler can detect PBL heights from about 150 m to 4000 m with an accuracy of ±200 m [Dye et al., 1995]. Hourly wind (0000 – 2300 LST) and *Zi* (0800 – 1700 LST, available only during daytime) measurements from the closest profiler to the GHG measurement site were used to evaluate WRF simulations. For example, most relevant to the ARV GHG measurement site, we compared *Zi* from WRF with measurements from the LHS profiler. For the MAD and TRA GHG sites, we used wind profiler data from the CCL site. As in Zhao et al. [2009] and Jeong et al. [2012a, 2012b], we assume that the RMS scatter in predicted versus measured *Zi* can be represented as the sum of squares of measurement uncertainty (~ 200 m, Dye et al. [1995]) and WRF model uncertainty. In other words, the model uncertainty is estimated by computing the model-data RMS scatter using hourly data and subtracting an estimated measurement error (~200 m) in quadrature. When the model-data RMS error is less than 200 m, we use the calculated RMS error value for the model uncertainty. For comparison between WRF and profiler measurements, we used data for May 2010, June 2010, October 2010, and January 2011 to represent spring, summer, fall, and winter seasons, respectively. Due to data availability we used 2010 data for spring and summer except for the CCO site for which May and June 2011 data

were used. For the LHS site, we used September 2010 data for fall because the LHS profiler data were not available after September 2010. Thus, we used the result from the CCL site for the LHS site after September 2010. Based on 2008 data at the Sacramento profiler used for Jeong et al. [2012a, 2012b], we note that the RMS values in PBL depth comparison (predicted vs. measured) are high (310 to 415 m) during winter and relatively low (160 – 220 m) during summer, showing seasonal variation. The error analysis obtained in the current study exhibits similar seasonal variation (i.e., high winter error vs. low summer error) in the PBL error, suggesting posterior emission estimates likely capture variations in seasonal emissions and annual total emissions, though it is possible that transport uncertainties over the three-month seasonal periods may vary somewhat from those determined from the individual months.

The WRF simulated *Zi* was generally consistent with the measured *Zi* based on the best fit slopes (~ unity) of predicted (WRF) vs. measured (wind profiler) *Zi*. For some cases, there were slight biases based on the regression analysis of predicted vs. measured *Zi*. During June, the CCO site showed a slightly higher best-fit slope of 1.29±0.13 than unity, while the LHS site yielded a slightly lower slope of 0.8±0.04 than unity. However, when we compared the mean diurnal cycles of predicted and measured *Zi*, we found no obvious bias at the two sites. Furthermore, we calculated the difference (predicted – measured) between the predicted and measured *Zi* means which were -64±86 m (95% C.I., due to a large enough sample size (>150) uncertainty estimation based on both t and normal distributions yielded the same C.I.) and 4±60 m for the LHS and CCO sites, respectively. This indicates that the mean biases are only ~5% and ~1% of the measurement means (1222 m and 539 m) for LHS and CCO, respectively and are well within the expected measurement accuracy (~200 m) of the wind profiler [Dye et al., 1995]. Also, the result in a t test for two means showed that the difference between predicted and

measured *Zi* for both sites was not significant: t(df = 358) = 1.47 with p-value = 0.14 and t(df = 295) = -0.14 with p-value = 0.89 for LHS and CCO, respectively. Based on this analysis we are reasonably confident in assuming that random errors dominate in the following analysis.

Following Jeong et al. [2012a, 2012b], we computed CH_4 mixing ratios (C_{CH4}) based on the perturbation in *Zi* (20% decrease) to estimate the sensitivity of C_{CH4} to *Zi* (i.e., dC_{CH4}/dZi) as a first order approximation. By reducing original *Zi* from the WRF model by 20%, we obtained perturbed CH_4 mixing ratios, which are compared with the original (normal) CH_4 mixing ratio to compute dC_{CH4}. Similarly, we computed *dZi* by comparing the perturbed *Zi* and normal *Zi*. Then, we calculated the monthly mean dC_{CH4}/dZi (in units of ppb/m), which represents the gradient of CH_4 mixing ratios with respect to *Zi*. Finally, we applied the inferred RMS errors (in units of m) in the WRF-STILT model to dC_{CH4}/dZi to estimate errors (in ppb) associated with *Zi* for each season and each site. Within a given season, monthly PBL uncertainty was obtained by scaling the uncertainty value for each representative month (a total of four months) in proportion to the background-subtracted mean mixing ratio. The estimated uncertainties ranged from ~5 ppb to over 200 ppb depending on the season and site, yielding large errors during winter and relatively small errors during summer. For instance, the ARV and MAD sites with the mean background-subtracted mixing ratio of ~500 ppb in January showed large errors associated with *Zi* (~200 ppb). In June, the uncertainties due to *Zi* errors in the ARV and MAD sites were relatively small (56 and 35 ppb, respectively) although the mean mixing ratios were also low (125 and 105 ppb).

Uncertainty in modeled CH_4 mixing ratios due to errors in modeled winds was estimated by comparing WRF-simulated winds and measured winds from the four wind profiler sites (Figure 2) for a total of four selected months as in the case of *Zi*. Following Jeong et al. [2012a] and Newman et al. [2013], when we compared WRF-simulated winds with profiler-measured winds

at the available levels of profilers near the surface (~200 m above mean sea level), the RMS errors in the wind U and V components varied depending on the season and measurement location. For the SAC profiler site (most relevant to WGC), the RMS errors for the wind U/V components were 3.42 (best-fit slope of predicted vs. measured with standard error = 1.00±0.03) / 2.95 (1.13±0.02), 2.89 (1.39±0.11) / 4.96 (1.41±0.11), 3.37 (1.04±0.04) / 3.11 (1.15±0.02), and 2.87 (0.98±0.03) / 2.88(1.05±0.03) m s^{-1} for October, January, May and June, respectively. For the CCL site (most relevant to MAD and TRA), we used data for October and January only because profiler data were not available for spring and summer 2011. The RMS errors for the U/V components were 3.77 (fit slope = 0.96±0.03) / 3.48 (fit slope = 1.04±0.03) and 2.76 (1.01±0.04) /2.91 (1.32±0.05) m s^{-1} for October and January (later we used the SAC site results for the other months to perform STILT ensemble runs). We evaluated winds at the CCO site for the months of May and June 2011 when CH_4 measurements were made at the STB site near the CCO site. The wind U/V RMS errors were 4.22 (fit slope = 1.03±0.04) / 5.99 (fit slope = 1.14±0.03) and 3.17 (0.95±0.03) / 4.45 (1.06±0.03) m s^{-1} for May and June, respectively. Since profiler wind data for the LHS site were not available after early September 2010, we used results from either SAC or CCL sites to run the STILT model for error quantification. For January when WRF overestimated wind speeds relative to profiler winds, we removed outliers (data points corresponding to > 2 standard deviation of hourly measured wind speed for the month) to avoid biases in inverse analyses. To estimate the effect of uncertainty in CH_4 mixing ratios due to winds ($\mathbf{S}_{transWIND}$) and particle number ($\mathbf{S}_{part}$), we ran the STILT model 10 times and computed ensemble predicted mixing ratios for a given site and month (a total of four selected months as in the *Zi* case). Based on 10 ensemble runs, we estimated the RMS difference about the mean of the ensemble mixing ratios for each model time step and use the monthly average

RMS as the combined uncertainty due to wind and particle number errors. Following the method in Zhao et al. [2009], Jeong et al. [2012a; 2012b] and Lin and Gerbig [2005] we propagated a stochastic component due to the wind velocity error, which was estimated from the model-data wind comparison, through STILT. As in Jeong et al. [2012a, 2012b], we adopted the set-up from Lin and Gerbig [2005] where they used 240 min, 120 km, and 900 m for correlation time scale (i.e., time scale for the temporal correlation to decay to zero), horizontal correlation scale and vertical correlation scale, respectively. This approach yielded a mixing ratio variation of 1 – 15 ppb depending on the season and site. As with the *Zi* case, the errors due to winds were higher during winter (8 – 15 ppb) than during summer (~ 2 ppb).

Following Zhao et al. [2009] and Jeong et al. [2012a; 2012b], we assumed that all of the errors are independent. The errors were combined in quadrature to yield a total expected model-data mismatch error, and the total error for each site is summarized in Table 3. Depending on the month and measurement location, the errors ranged from 20 to 233 ppb, which are approximately 30 – 60% of the background-subtracted mean mixing ratio. The total error was particularly large (100 – 233 ppb) during winter in the ARV and MAD sites where the background-subtracted mean mixing ratio was also high (220 - 520 ppb).

3. Results

3.1. CH_4 Mixing Ratios

Figure 3 shows the 3-hourly measured mixing ratio, background mixing ratio and predicted (before inversion) mixing ratio using the CALGEM prior model for the five network sites. Predicted mixing ratios are shown only for the data points used in the final inversion during the well-mixed periods (noon – afternoon). For inverse analyses, we use data during 12 – 17 hours

(LST) except for winter (11 – 16 LST) during which we found that PBL tends to collapse earlier in WRF simulations than in wind profiler measurements. Based on a sensitivity test to the outlier removal, we find that there is no significant difference in the posterior emissions for the Central Valley between the first and final inversions based on the CALGEM prior (33.01±1.97 vs. 32.87±2.09 Tg CO_2eq yr^{-1}). This suggests that the small amount of data (~ 5%) that were removed do not significantly affect the posterior emissions.

Overall, the predicted mixing ratios at all sites show underestimation of CH_4 compared to the measurements although the prediction captures the synoptic variation of the measured mixing ratios (Figure 3). The minimum measured mixing ratios approximate the predicted background CH_4 well, suggesting that the estimated background mixing ratios are reasonable and there is no significant bias in the measured mixing ratios. In order to examine systematic biases in background values, we also computed the intercept from the linear regression (predicted vs. measured) for each month after subtracting background values from measured CH_4 mixing ratios. We found no significant bias in this comparison except for January, March and April, which showed intercepts (in regression coefficients) of −28.65±19.39 (= standard error), −4.78±2.92, and −4.75±2.17, respectively. These values are small compared to the background-subtracted mean mixing ratio for the corresponding month, and did not affect inversion results significantly. The result also shows that there is a clear seasonal variation in CH_4 mixing ratios with high variability, particularly in winter while ARV and MAD show high variability throughout the seasons. The comparison result in STB indicates that the prior emissions from rice agriculture during late spring and early summer are significantly lower than actual emissions. The CALGEM prior estimate for crop agriculture based on the DNDC model suggests that CH_4 emissions from rice agriculture in Region 6 become strong starting in June with an

emission sum of 3.4 Tg CO_2eq yr^{-1} and peaking in August with emissions equating to 4.6 Tg CO_2eq yr^{-1}. We discuss more on rice emissions later in Section 3.3. For WGC, the predicted mixing ratios are significantly lower than the measurements, showing similar results to those shown in Jeong et al. [2012a].

3.2. Footprints

We present the first analysis of footprints that constrain most of California's Central Valley and its surrounding areas across different seasons. When footprints for all five sites are combined, the sensitivity of the measurement sites to surface emissions is significantly improved, as compared to the result with one site only. Figure 4 shows the average footprint from the multiple sites during September 2010 – June 2011, including the average footprint (May – June 2011) using a single site (i.e., WGC) for comparison. The significance of the multi-site network is clear in the figure where the averaged footprint from a single tower shows limited sensitivity while the footprint from the multiple sites shows strong sensitivity in the entire Central Valley. Although the measurement network significantly expands the area that is constrained, the network in the Central Valley shows limited ability to constrain CH_4 emissions in the Southern California region due to weak sensitivity.

Because measurements at STB were available only during May – June 2011, STB footprints were not simulated for other seasons. There is a clear seasonal pattern for the distribution of footprints, which is important to attribute mixing ratios to different emission sources for each season. Overall, the seasonal footprints are strong in the north-south direction in the Central Valley although footprints are strong in the west-east direction near the WGC site for some seasons. Depending on the season, footprints allow for constraining important urban emissions (e.g., South Coast Air Basin).

3.3. Bayesian Inverse Analysis

Bayesian inverse analysis was conducted using two independent prior emission models: CALGEM and EDGAR42 emission models. Using each emission model, we performed Bayesian inversion to estimate optimized emissions for 1) the 14 regions defined in Figure 1 (region analysis) and 2) individual emission source sectors (source analysis). For the region analysis, we solve for a total of 14 scaling factors (i.e., dimension of $\boldsymbol{\lambda} = 14 \times 1$) for each month including the region outside California for both emission models. The dimensions of $\boldsymbol{\lambda}$ for the source analysis (for each month) are 9×1 (i.e., 8 sectors and outside California) and 17×1 (16 EDGAR sectors and outside California) for the CALGEM and EDGAR42 cases, respectively.

Table 4 summarizes the chi-square linear analysis results where we show the best-fit slopes (with standard error) of predicted vs. measured CH_4 mixing ratios before (prior) and after (posterior) Bayesian region inversion [Press et al., 1992]. The best-fit slopes were obtained using the data from all sites for a given month, reflecting the aggregate regression of predicted vs. measured mixing ratios. The posterior results in Table 4 were obtained by applying Equations 1 and 2 in Section 2.4. Predicted CH_4 mixing ratios using the CALGEM emission model are typically 30 – 50% of measurements before inversion while EDGAR42-based prior mixing ratios are 20 – 40% of measurements. After inversion, the posterior CH_4 mixing ratios based on the CALGEM emission model are consistent with the measurements for most of the months, while the posterior mixing ratios from EDGAR42 are still lower than the measurements. To further examine the low best-fit slopes of posterior predictions vs. measurements based on EDGAR42, we conducted an inversion using 100% uncertainty in the prior. We find that the inversion still yields best-fit slopes of posterior predictions vs. measurements that are lower than unity (0.80–0.94) although the best-fit slopes based on the 100% uncertainty assumption are slightly higher

than those of the 70% assumption in the prior uncertainty. This suggests that the spatial distribution of CH_4 emissions in California is not well represented by EDGAR42. The counterpart source analysis showed a similar result where EDGAR-based mixing ratios are lower than those of the CALGEM case.

Inversions are performed at the monthly temporal scale, and inferred CH_4 emissions are reported by season (five bi-monthly seasons during September 2010 – June 2011) for the regions and source sectors where the total emissions are significant and footprints show sensitivity (Figure 5). Figure 5(a) shows the Bayesian region analysis result (solving $\boldsymbol{\lambda}$ for each region) using the CALGEM prior emission model. Overall, the inversion results show that actual CH_4 emissions are higher than the prior emissions for most of the regions. In particular, the posterior (optimized) emissions are significantly higher than the prior in the Central Valley (Regions 6, 8, and 12) where measurements are made and thus the emissions are well constrained. For Region 10 (Southern California region), the posterior uncertainties are only slightly reduced, suggesting that the measurements in the Central Valley weakly constrain the emissions in Region 10 (see Table 5 for details). The significantly higher posterior emissions in the San Joaquin Valley (Regions 8 and 12) suggest that emissions from the livestock source sector are significantly higher than the prior. Note that livestock emissions from the CALGEM emission model account for 87% (4.33 Tg CO_2eq) and 84% (6.77 Tg CO_2eq) of the total emissions in Regions 8 and 12, respectively. The results also show that there is a clear seasonal variation in CH_4 emissions. For example, in Region 6 where high emissions are expected from rice agriculture, the posterior emissions are high during the early fall and late spring - early summer seasons. We discuss more on rice emissions later in the section.

We also performed Bayesian region analysis based on the EDGAR42 emission model (Figure 5(b)) and compare the result with the CALGEM case (shown in Figure 5(a)). In Table 5, we summarized annual CH_4 emissions from the Bayesian region analyses based on the CALGEM and EDGAR42 prior emission models, including the aggregated uncertainty for the entire state as well as that of each region. For the uncertainty in state total emissions, we report the uncertainty using two error assumptions: uncorrelated (only including diagonal elements of the posterior error covariance matrix) and correlated (also including off-diagonal elements) errors among the regions. As shown in Table 5, the correlated error assumption yields slightly smaller aggregated uncertainty than that of the uncorrelated assumption. This is because there are anti-correlations (i.e., negative correlation coefficients, Tarantola [1987]) between some of the regions as reported by Jeong et al. [2012a] and Bergamaschi et al. [2005]. Bergamaschi et al. [2005] showed slightly smaller aggregated uncertainty in the correlated error estimation than in the uncorrelated error estimation. Hereafter, we only report more conservative uncertainty estimates (i.e., based on the uncorrelated error assumption) to consider potential uncertainties (e.g., uncaptured transport uncertainty) that we may not have identified although some of the regions may have correlated errors. Also, we note that we have not defined the correlations between regions to construct the prior error covariance for $\boldsymbol{\lambda}$. This might affect posterior uncertainty estimation and needs to be investigated further in future studies.

The region analysis results in Table 5 (also in Figures 5(a) and 5(b)) show that the current measurement network estimates annual average CH_4 emissions for the Central Valley (i.e., Regions 6, 8 and 12) to be 32.87±2.09 (prior = 15.09) Tg CO_2eq and 31.60±2.17 (prior = 10.79) Tg CO_2eq based on the CALGEM and EDGAR42 prior emission models respectively, assuming uncorrelated errors between regions. This suggests that the measurement network constrains

emissions in the Central Valley, independent of *a priori* emission models. However, the posterior emission estimates based on the EDGAR42 and CALGEM prior models are only marginally consistent in the predominantly urban regions (7 and 10) where the EDGAR42 model yields higher CH_4 emissions than those estimated with the CALGEM model: 29.11±11.59 vs. 12.92±6.08 Tg CO_2eq yr^{-1}. This is because the EDGAR42 emissions are significantly higher than the CALGEM emissions in Regions 7 and 10 (23.7 vs. 10.5 Tg CO_2eq yr^{-1}, see Tables 1 and 2), and our measurement sites in the Central Valley have relatively weak sensitivity to the urban regions. Although the results using multiple emission models help to characterize the uncertainty associated with estimating emissions at the sub-regional scale, these results demonstrate that additional measurements are required in the San Francisco Bay and Southern California areas in order to strongly constrain emissions from those urban regions.

We also estimate CH_4 emissions by inferring state-wide scaling factors for each emission source instead of each sub-region. Figure 5(c) shows the source analysis results using the CALGEM emission model. These results are consistent with those of the counterpart inverse analysis for regional emissions. For example, the source inversion suggests that actual emissions from livestock are much higher than the prior. Recall that the region analysis result showed higher posterior emissions in Regions 8 and 12 where livestock emissions are dominant (~90% of annual CH_4 emissions). Figure 5(c) indicates that CH_4 emissions from natural gas sources are generally higher than the prior. However, more measurements are required to effectively constrain natural gas emissions from the large urban areas (Regions 7 and 10), which account for 64% of the total natural gas emissions in the CALGEM model (urban ratio for natural gas in EDGAR42 = 76%).

The source analysis result also indicates that the posterior emissions for crop agriculture are higher during early fall and late spring-summer seasons than the prior, which are consistent with the region analysis (higher emissions in Region 6). Our result is similar to that of a recent study based on aircraft CH_4 measurements during the California Research at the Nexus of Air Quality and Climate Change (CalNex) period in summer 2010 [Peischl et al., 2012]. Peischl et al. [2012] estimated annual CH_4 emissions from rice cultivation to be 1.64 – 1.95 Tg CO_2eq, which is 3.0 – 3.6 times larger than the CARB 2008 inventory for rice CH_4 emissions (0.54 Tg CO_2eq yr^{-1}). This estimate by Peischl et al. [2012] is based on the rice emission study in a commercial rice field by McMillan et al. [2007] where they estimated annual CH_4 emissions of 26.1 – 31.0 g CH_4-C m^{-2} during October 2001 – October 2002. Assuming posterior emissions for July and August (not available in our study) are proportional to the prior and scaling (available) June posterior emissions according to the prior ratios of July and August to June (3.26 Tg CO_2eq / 3.59 Tg CO_2eq and 5.10 / 3.59, respectively), we find that the annual rice emission total is 1.43±0.19 Tg CO_2eq (original DNDC prior for rice = 1.34 Tg CO_2eq), which is very similar to that of Peischl et al. [2012]. The slight difference between the estimate by Peischl et al. [2012] and our estimate is possibly due to the difference in emissions during late fall and winter. CH_4 emissions during late fall and winter from McMillan et al. [2007] are not negligible while our *a priori* rice emissions based on the DNDC model described by Salas et al. [2006] are insignificant and often negative.

The source analysis results based on EDGAR42 are shown in Figure 5(d), where eight major sources (~95% of total emissions) out of a total of 16 sources are compared. While posterior emissions from livestock for the entire state are similar between the CALGEM (32.23±2.92 Tg CO_2eq) and EDGAR42 (33.60±3.72 Tg CO_2eq) models, the source analysis based on EDGAR42

shows different posterior emissions for some of the source sectors, compared to the CALGEM case. In particular, the state-wide annual CH_4 emission for solid waste (equivalent to landfill of the CALGEM model) based on the EDGAR42 prior model is 23.38±6.47 Tg CO_2eq, which is only marginally consistent with that (14.43±3.92) estimated using the CALGEM model. This is likely due to the fact that ~70% of landfill emissions are concentrated in the urban regions (Regions 7 and 10), and these urban regions are only weakly constrained by the measurements in the Central Valley.

4. Discussion and Conclusions

The current GHG network constrains annual CH_4 emissions from California's Central Valley to be 32.87±2.09 Tg CO_2eq and 31.60±2.17 Tg CO_2eq based on the CALGEM and EDGAR42 prior models respectively, showing consistency between the two independent prior emission models. However, as noted above, our region analysis estimates state total annual CH_4 emissions to be 1.51±0.20 times and 2.03±0.37 times (Table 5) the current CARB inventory (32 Tg CO_2eq; CARB, 2011) using the CALGEM and EDGAR42 priors, respectively. This suggests that uncertainty in the state total emission estimates are dominated by uncertainty in emissions from the urban regions.

To address the uncertainty in state total emissions by constraining urban emissions based on published work, we consider a range of emission estimates for the larger Los Angeles metropolitan area (hereafter SoCAB). All relevant studies in SoCAB use correlations of CH_4 to CO enhancements and CO emission inventories to estimate CH_4 emissions [Hsu et al., 2010; Wennberg et al., 2012; Peischl et al., 2013]. Here, we apply the results from Wennberg et al. [2012], which provide a more conservative estimate (0.44±0.15 Tg CH_4 yr^{-1}) for SoCAB than

those of Hsu et al. [2010] (0.38±0.10 Tg CH_4 yr^{-1}, recalculated by Wennberg et al. [2012]) and Peischl et al. [2013] (0.41±0.04 Tg CH_4 yr^{-1}). This estimate of urban emissions is 0.91 to 1.84 times the CALGEM prior for SoCAB but a factor of 0.45 to 0.91 times SoCAB emissions in EDGAR42, which taken together with results from the Central Valley suggest that CALGEM provides a superior representation of California CH_4 emissions. Here, we estimate state total CH_4 emissions as the sum of our posterior emission estimates for the Central Valley and other non-urban regions (33.24 – 37.63 Tg CO_2eq yr^{-1}) and the CALGEM prior emissions for major urban regions (10.45 Tg CO_2eq yr^{-1}) scaled by a factor ranging from 0.91 to 1.84 from the above comparison with Wennberg et al [2012]. This yields annual state total emissions of 42.75 Tg CO_2eq yr^{-1} (= 33.24 + 10.45 × 0.91) to 56.86Tg CO_2eq yr^{-1} (= 37.63 + 10.45 × 1.84), suggesting that California total emissions are 1.34 to 1.78 times the current CARB CH_4 inventory.

We also note that the primary source of uncertainty is due to under-sampling of urban regions, not temporal coverage. For example, when we compare the 10-month case (without July and August) with the full year case based on the results from Jeong et al. [2012a] that analyzed a full year of data from the Central Valley, we do not find a significant difference (13.0±2.0 vs. 14.1±2.2 Tg CO_2eq yr^{-1}) for the regions (Regions 6, 7 and 8) near the WGC tower.

In conclusion, our measurements constrain annual mean CH_4 emissions from California's Central Valley and state total emissions when combined with independent estimates from urban regions. In the future, we expect that additional tower measurements in the San Francisco Bay and Southern California areas will be effective in constraining urban emissions and that measurements of source specific tracers (e.g., CO, VOCs, and potentially CH_4 isotopes) will help separate different sources of CH_4 [Townsend-Small et al., 2012; Peischl et al., 2013].

Acknowledgements

We thank Dave Field, Dave Bush, Edward Wahl, Ken Reichl, Fabien Guerin, Yuchen Yi, and particularly Jon Kofler for assistance with measurements at WGC and analysis of data from radar wind profiler sites, John Lin, Christoph Gerbig, Steve Wofsy, Janusz Eluszkiewicz, Thomas Nehrkorn for sharing the STILT code and advice, Chris Potter and William Salas for sharing modeled CH_4 emission for use as a priori estimates, Ed Dlugokencky and Colm Sweeney for sharing data for CH_4 background estimates, Larry Hunsaker, Marc Vayssières, Joseph Fischer, and Webster Tassat for sharing CARB CH_4 emissions information, and Krishna Muriki for assistance running the WRF-STILT models on the LBNL-Lawrencium cluster. We thank Ken Davis and two anonymous reviewers for providing thoughtful comments and suggestions. We also thank Guido Franco, Jorn Horner, and Eileen McCauley for useful advice, and David Larson for institutional support at CSU East Bay. This analysis was supported by CARB (Contract number 09-348), data collection at Walnut Grove by the California Energy Commissions Public Interest Environmental Research program, with work at LBNL conducted under US Department of Energy Contract DE-AC02-05CH11231. Support for L. Bianco and J. Wilczak was provided by NOAA USWRP and HOA program funding. NOAA measurements were funded in part by the Atmospheric Composition and Climate Program and the Carbon Cycle Program of NOAA's Climate Program Office.

References

Andrews, A. E., J. D. Kofler, M. E. Trudeau, J. C. Williams, D. H. Neff, K. A. Masarie, D. Y. Chao, D. R. Kitzis, P. C. Novelli, C. L. Zhao, E. J. Dlugokencky, P. M. Lang, M. J. Crotwell, M. L. Fischer, J. T. Lee, D. D. Baumann, A. R. Desai, C. O. Stanier, S. F. J. de Wekker, D. E. Wolfe, J. W. Munger, M. J. Parker, and P. P. Tans (2013), CO_2, CO and CH_4 measurements from the NOAA Earth System Research Laboratory's tall tower greenhouse gas observing network: instrumentation, uncertainty analysis and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmospheric Measurement Techniques Discussions, 6, 1461–1553.

Bergamaschi, P., M. Krol, F. Dentener, A. Vermeulen, F. Meinhardt, R. Graul, M. Ramonet, W. Peters, and E. J. Dlugokencky (2005), Inverse modelling of national and European CH_4 emissions using the atmospheric zoom model TM5. Atmos. Chem. Phys., 5, 2431–2460.

Bianco, L., and J. Wilczak (2002), Convective boundary-layer depth: improved measurement by Doppler radar wind profile using fuzzy logic methods. J. Atmos. Oceanic Tech., 19, 1745-1758.

Bianco, L., J. Wilczak, and A. White (2008), Convective boundary layer depth estimation from wind profilers: statistical comparison between an automated algorithm and expert estimations. J. of Atmos. and Oceanic Tech., 25, 1397–1413, DOI: 10.1175/2008JTECHA981.1

Bianco, L., I. V. Djalalova, C. W. King, and J. M. Wilczak (2011), Diurnal evolution and annual variability of boundary-layer height and its correlation to other meteorological variables in California's Central Valley, Boundary-Layer Meteorology, 140, 491-511, DOI 10.1007/s10546-011-9622-4, 1-21.

CALGEM (2013), California Greenhouse Gas Emissions Measurement Project emissions maps (http://calgem.lbl.gov/prior_emission.html, document: CA-CH4-CALGEM-v2.1-20121221.zip).

CARB (2010), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report (personal communication, document: ghg_inventory_by_sector_00-08_2010-05-12.xls).

CARB (2011), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report (http://www.arb.ca.gov/cc/inventory/inventory.htm, document: ghg_inventory_by_ipcc_00-09_2011-10-26.xls).

Dye, T.S., C. G. Lindsey, and J. A. Anderson (1995), Estimates of mixing depth from "boundary layer" radar profilers. Preprints from the 9th Symposium on Meteorological Observations and Instrumentation, Charlotte, NC, March 27-31, 156-160 (STI-94212-1451).

Gerbig, C, J. Lin, S. Wofsy, B. Daube, A. E. Andrews, B. Stephens, P. S. Bakwin and C. Grainger (2003), Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. J. Geophys. Res., 108(D24), doi:10.1029/2003JD003770.

Gimson, N. R. and Uliasz, M. (2003), The determination of agricultural methane emissions in New Zealand using receptor-oriented modeling techniques. Atmos. Environ., 37, 3903-3912.

Göckede, M., A. M. Michalak, D. Vickers, D. P. Turner, and B. E. Law (2010), Atmospheric inverse modeling to constrain regional-scale CO_2 budgets at high spatial and temporal resolution. J. Geophys. Res., 115, D15113, doi:10.1029/2009JD012257.

Hofman, D. J., J. H. Butler, E. J. Dlugokencky, J. W. Elkins, K. Masarie, S. A. Montzka, and P. Tans (2006), The role of carbon dioxide in climate forcing from 1979 - 2004: Introduction of the Annual Greenhouse Gas Index. Tellus B, 58B, 614-619.

Houweling, S., T. Kaminski, F. Dentener, J. Lelieveld, and M. Heimann (1999), Inverse modeling of methane sources and sinks using the adjoint of a global transport model. J. Geophys. Res., 104(D21), 26137-26160.

Hsu, Y.-K., T. VanCuren, S. Park, C. Jakober, J. Herner, M. FitzGibbon, D. R. Blake, and D. D. Parrish (2010), Methane emissions inventory verification in southern California, Atmos. Environ., 44, 1 – 7, doi:10.1016/j.atmosenv.2009.10.002.

IPCC (1995), IPCC Second Assessment Report (SAR), "Climate Change 1995".

Janjic, Z. I. (1990), The step-mountain coordinate: physical package. Mon. Weather Rev., 118, 1429–1443.

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

Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012b), Seasonal variations in N_2O emissions from central California. Geophys. Res. Lett., 39, L16805, doi:10.1029/2012GL052307.

Kort, E. A., J. Eluszkiewicz, B. B. Stephens, J. B. Miller, C. Gerbig, T. Nehrkorn, B. C. Daube, J. O. Kaplan, S. Houweling, and S. C. Wofsy (2008), Emissions of CH_4 and N_2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations. Geophy. Res. Lett., 35, L18808, doi:10.1029/2008GL034031.

Lin, J. C., and C. Gerbig (2005), Accounting for the effect of transport errors on tracer inversions, Geophys. Res. Lett., 32, L01802, doi:10.1029/2004GL021127.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. J. Geophys. Res., 108(D16), 4493, doi:10.1029/2002JD003161.

McMillan, A. M. S., M. L. Goulden, and S. C. Tyler (2007), Stoichiometry of CH_4 and CO_2 flux in a California rice paddy, J. Geophys. Res., 112, G01008, doi:10.1029/2006JG000198.

Mellor G. L. and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems. Rev. Geophys. Space Phys. 20, 851–875.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006), North American Regional Reanalysis. Bull. Amer. Meteor. Soc., 87 (3), 343-360.

Montzka, S. A., E. J. Dlugokencky1, and J. H. Butler (2011) Non-CO_2 greenhouse gases and climate change. Nature, 476, 43-50.

Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting - stochastic time-inverted lagrangian transport (WRF-STILT) model. Meteor. Atmos. Phys., 107 (1), 51-64, doi:10.1007/s00703-010-0068-x.

Newman, S., 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 (2013), Diurnal tracking of anthropogenic CO_2 emissions in the Los Angeles basin megacity during spring 2010, Atmos. Chem. Phys., 13, 4359–4372, doi:10.5194/acp-13-4359-2013.

Peischl, J., et al. (2012), Airborne observations of methane emissions from rice cultivation in the Sacramento Valley of California, J. Geophys. Res., 117, D00V25, doi:10.1029/2012JD017994.

Peischl, J., et al. (2013), Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118, doi:10.1002/jgrd.50413.

Potter, C., S. Klooster, S. Hiatt, M. Fladeland, V. Genovese and P. Gross (2006), Methane Emissions from Natural Wetlands in the United States: Satellite-Derived Estimation Based on Ecosystem Carbon Cycling. Earth Interactions, 10, 1–12.

Press, W. H., S. A.Teukolsky, W. T. Vetterling, and B.P. Flannery (1992), Numerical Recipes in C, 2nd edition. Cambridge: Cambridge University Press.

Salas, W., P. Green, S. Frolking, C. Li and S. Boles (2006), Estimating Irrigation Water Use for California Agriculture: 1950s to Present. California Energy Commission, PIER Energy-Related Environmental Research. CEC-500-2006-057.

Skamarock, W.C., J. B. Klemp, J. Dudhia, D.O. Gill, D. M. Barker, X. Z. Huang, W. Wang, and J. G. Powers (2008), A description of the advanced research WRF version 3. Technical Note 475+STR. Mesoscale and Microscale Meteorology Division, NCAR, Boulder, Colorado.

Tarantola, A. (1987), Inverse Problem Theory Methods for Data Fitting and Model Parameter Estimation, 613 pp., Elsevier, New York.

Thoning, K.W., P. P. Tans, and W. D. Komhyr (1989), Atmospheric carbon dioxide at Mauna Loa Observatory, 2. Analysis of the NOAA/GMCC data, 1974 - 1985, J. Geophys. Res., 94, 8549 - 8565.

Townsend-Small, A., S. C. Tyler, D. E. Pataki, X. Xu, and L. E. Christensen (2012), Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of "fugitive" fossil fuel emissions, J. Geophys. Res., 117, D07308, doi:10.1029/2011JD016826.

Wennberg, P. O., W. Mui, D Wunch, E. A. Kort, D. R. Blake, E. L. Atlas, G. W. Santoni, S. C. Wofsy, G. S. Diskin, S. Jeong, and M. L. Fischer (2012), On the Sources of Methane to the Los Angeles Atmosphere. Environ. Sci. Technol., 46 (17), 9282 - 9289, doi:10.1021/es301138y.

Wyngaard, J. C., and M. A. LeMone (1980), Behavior of the refractive index structure parameter in the entraining convective boundary layer, J. Atmos. Sci., 37, 1573–1585.

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.

Tables

Table 1. Annual CALGEM CH_4 Emissions by Region and Sector (Tg CO_2eq)[a]

Sector	Region													Sector
	R01	**R02**	**R03**	**R04**	**R05**	**R06**	**R07**	**R08**	**R09**	**R10**	**R11**	**R12**	**R13**	**Total**
Crop agriculture	0.00	0.00	0.00	0.01	0.00	0.50	0.00	0.01	0.00	0.00	0.01	0.00	0.02	**0.54**
Landfill	0.02	0.04	0.11	0.08	0.03	0.46	0.87	0.19	0.34	4.00	0.10	0.29	0.06	**6.60**
Dairy livestock	0.00	0.00	0.01	0.10	0.01	0.36	0.08	3.79	0.02	1.71	0.03	5.77	0.01	**11.90**
Non-dairy livestock	0.03	0.10	0.11	0.06	0.17	0.19	0.12	0.54	0.11	0.64	0.07	1.00	0.03	**3.17**
Natural gas	0.00	0.01	0.04	0.02	0.01	0.33	0.33	0.10	0.05	0.91	0.02	0.11	0.03	**1.95**
Petroleum	0.00	0.00	0.05	0.00	0.00	0.03	0.05	0.02	0.07	0.19	0.00	0.71	0.00	**1.13**
Wastewater	0.00	0.09	0.02	0.01	0.00	0.03	0.17	0.08	0.06	1.33	0.01	0.11	0.01	**1.92**
Wetland	0.01	0.00	0.00	0.00	0.22	0.18	0.03	0.27	0.01	0.03	0.01	0.02	0.01	**0.79**
Region Total	**0.06**	**0.24**	**0.34**	**0.28**	**0.44**	**2.08**	**1.65**	**5.00**	**0.66**	**8.81**	**0.25**	**8.01**	**0.17**	**28.00**

[a]Assumed a global warming potential of 21 g CO_2eq/g CH_4 [IPCC, 1995].

Table 2. Annual EDGAR42 CH_4 Emissions by Region and Sector (Tg CO_2eq)

Sector	Region													Sector Total
	R01	**R02**	**R03**	**R04**	**R05**	**R06**	**R07**	**R08**	**R09**	**R10**	**R11**	**R12**	**R13**	
Agricultural waste and burning	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.02**
Energy manufacturing transportation	0.00	0.00	0.01	0.00	0.00	0.03	0.08	0.01	0.01	0.26	0.00	0.03	0.01	**0.44**
Enteric fermentation	0.05	0.24	0.24	0.14	0.35	0.37	0.31	1.38	0.27	1.18	0.15	2.44	0.09	**7.22**
Fugitive from solid	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**
Gas production and distribution	0.01	0.02	0.11	0.04	0.02	0.86	2.05	0.29	0.29	5.90	0.04	0.74	0.13	**10.50**
Industrial process and product use	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.06	0.00	0.00	0.00	**0.10**
Manure management	0.02	0.03	0.02	0.02	0.04	0.04	0.10	0.71	0.04	0.92	0.02	0.38	0.01	**2.34**
Oil production and refineries	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.03	0.15	0.00	0.12	0.00	**0.45**
Residential	0.00	0.01	0.02	0.01	0.00	0.04	0.08	0.03	0.03	0.24	0.01	0.05	0.02	**0.54**
Road transportation	0.00	0.00	0.00	0.00	0.00	0.02	0.06	0.01	0.00	0.18	0.00	0.01	0.00	**0.29**
Solid waste disposal	0.02	0.06	0.23	0.12	0.05	0.88	2.24	0.49	0.34	6.92	0.08	0.81	0.16	**12.39**
Wastewater	0.00	0.00	0.03	0.01	0.00	0.19	0.63	0.09	0.08	2.14	0.01	0.13	0.03	**3.34**
Agricultural soils	0.00	0.01	0.01	0.00	0.01	0.17	0.02	0.15	0.00	0.02	0.00	0.28	0.01	**0.68**
Non-road transportation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**
Fossil fuel fires	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**
Large scale biomass burning	0.00	0.00	0.02	0.02	0.02	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00	**0.10**
Region Total	**0.10**	**0.37**	**0.68**	**0.34**	**0.47**	**2.62**	**5.73**	**3.16**	**1.11**	**17.99**	**0.31**	**5.01**	**0.46**	**38.34**

Table 3. Summary of Estimated Model-Data Mismatch Uncertainty by Site (ppb)

Site	Month									
	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
ARV	61	48	53	144	218	112	111	76	46	61
MAD	61	64	98	150	233	100	77	53	34	41
TRA	58	57	100	110	148	73	44	35	30	32
STB	NA	NA	NA	NA	NA	NA	NA	NA	21	29
WGC	25	27	29	86	128	53	31	22	20	22

Table 4. Linear Regression Analysis (Predicted vs. Measured CH_4 Mixing Ratios) Results Before and After Bayesian Region Inversion

		Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Before Inversion (prior)[e]											
CA[a]	Slope[c]	0.47±0.02	0.56±0.04	0.49±0.03	0.51±0.04	0.38±0.04	0.33±0.03	0.37±0.02	0.43±0.02	0.45±0.02	0.53±0.04
	RMSE[d]	70	70	117	153	283	126	102	78	52	52
ED42[b]	Slope	0.28±0.03	0.43±0.05	0.31±0.02	0.35±0.04	0.26±0.03	0.29±0.03	0.24±0.02	0.26±0.01	0.31±0.02	0.42±0.04
	RMSE	90	73	137	168	312	133	114	90	61	54
After Inversion (posterior)											
CA	Slope	0.97±0.04	0.98±0.07	0.95±0.05	0.98±0.08	0.99±0.1	0.81±0.07	0.96±0.05	0.99±0.04	1.01±0.04	0.96±0.05
	RMSE	35	51	78	118	164	86	63	44	30	32
ED42	Slope	0.83±0.05	0.88±0.07	0.83±0.04	0.89±0.08	0.85±0.08	0.73±0.06	0.82±0.05	0.87±0.03	0.92±0.04	0.85±0.07
	RMSE	45	46	74	121	168	88	69	44	34	37

[a]CALGEM prior emission model

[b]EDGAR42 prior emission model

[c]Best-fit slope of predicted vs. measured mixing ratios with standard error

[d]Root mean square error in units of ppb

[e]Results after outlier removal

Table 5. Comparison of Annual Posterior CH_4 Emissions (Tg CO_2eq) between CALGEM-based and EDGAR42-based Bayesian Region Analysis

Prior Model	Emission	Region													Total
		R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	
CALGEM	Prior Emissions	0.06	0.24	0.34	0.28	0.44	2.08	1.65	5.00	0.66	8.81	0.25	8.01	0.17	**28.00**
	Prior Uncertainty[a]	0.04	0.17	0.24	0.20	0.31	1.46	1.15	3.50	0.46	6.17	0.18	5.61	0.12	**9.26**[b]
	Posterior Emissions	0.06	0.24	0.39	0.29	0.44	4.53	3.01	8.57	0.70	9.90	0.29	19.78	0.16	**48.35**
	Posterior Uncertainty[c]	0.04	0.17	0.24	0.20	0.29	0.63	1.00	1.27	0.46	6.00	0.17	1.54	0.11	**6.47**[d] **(6.27)**[e]
EDGAR42	Prior Emissions	0.10	0.37	0.68	0.34	0.47	2.62	5.73	3.16	1.11	17.99	0.31	5.01	0.46	**38.34**
	Prior Uncertainty[a]	0.07	0.26	0.48	0.24	0.33	1.83	4.01	2.21	0.78	12.59	0.22	3.51	0.32	**14.02**[b]
	Posterior Emissions	0.10	0.39	0.78	0.36	0.51	5.72	8.52	7.99	1.23	20.59	0.38	17.89	0.51	**64.97**
	Posterior Uncertainty[c]	0.07	0.26	0.47	0.24	0.33	0.96	1.97	1.20	0.77	11.43	0.21	1.53	0.32	**11.85**[d] **(11.57)**[e]

[a]70% uncertainty in priors

[b]Square root of sum of squares of prior uncertainty for each region

[c]Posterior uncertainty = 1σ

[d]This uncertainty is calculated based on the uncorrelated error assumption between the regions.

[e]This uncertainty in parentheses is calculated based on the propagation of correlated errors using the posterior error covariance matrix as:

$\sigma_e^2 = \sum_i^n (e_i \sigma_i)^2 + \sum_i^n \sum_{j(j \neq i)}^n e_i e_j \rho_{ij} \sigma_i \sigma_j$ where σ_e is the aggregated emission uncertainty, n is the number of regions (i.e., n = 13), e_i (or e_j) is the prior emission for each region, σ_i (or σ_j) is the posterior scaling factor uncertainty for each region from the posterior error covariance matrix, and ρ_{ij} is the correlation coefficient ($-1 \leq \rho_{ij} \leq 1$) between regions i and j.

Figures

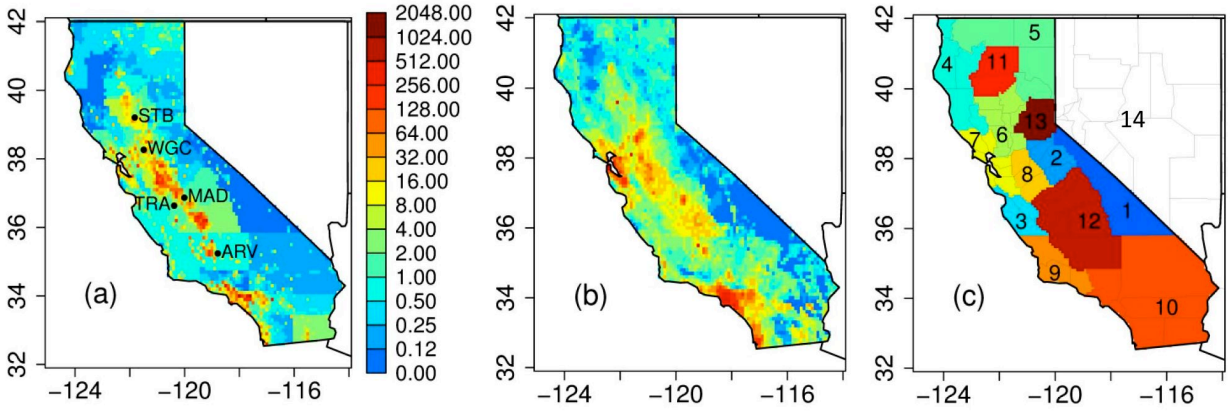

Figure 1. (a) CALGEM total CH_4 emissions (nmol m^{-2} s^{-1}) with network measurement locations (black dots), (b) EDGAR42 total CH_4 emissions (nmol m^{-2} s^{-1}), and (c) 14 sub-region classification for inverse modeling including the region outside California (Region 14).

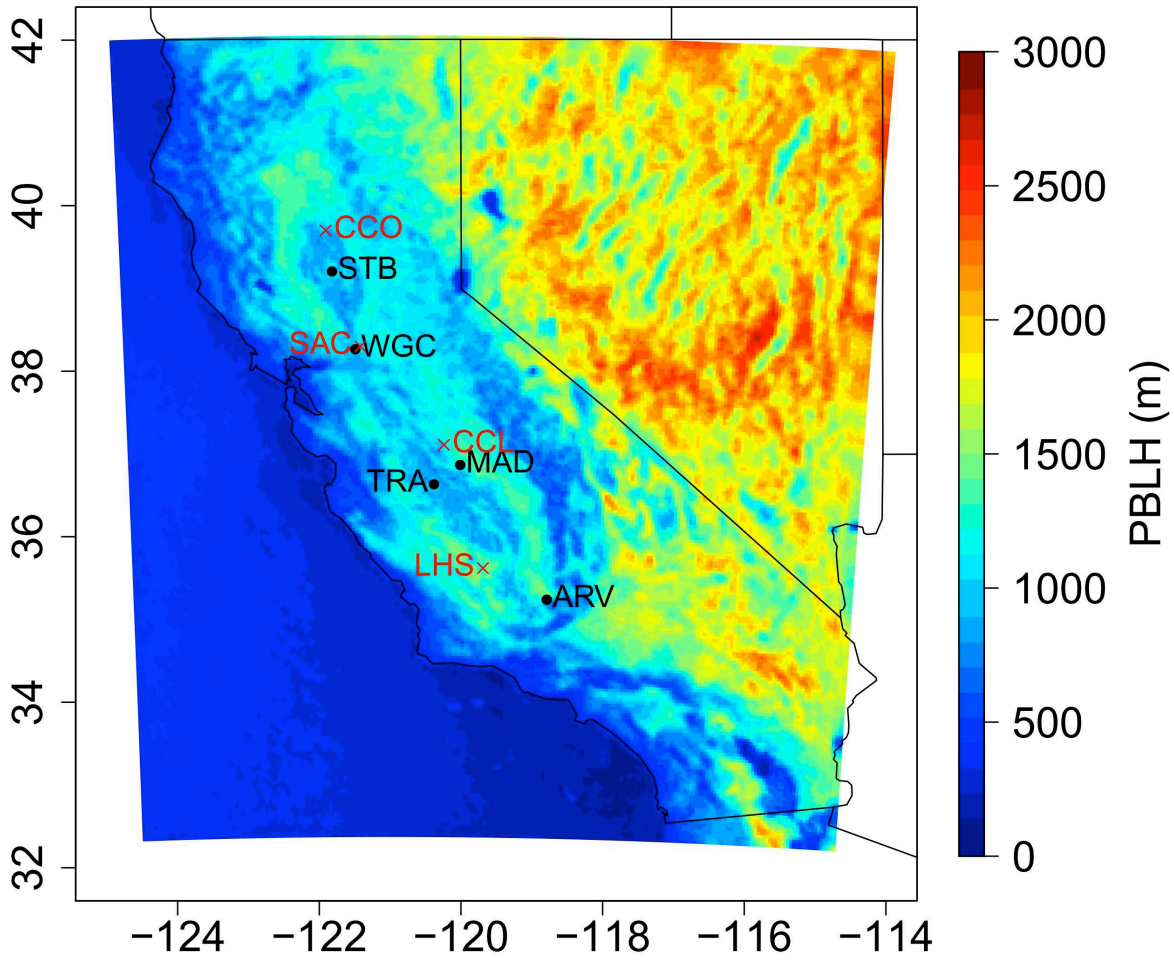

Figure 2. Location of GHG measurement sites (black) and wind profiler sites (red) in the Central Valley with predicted monthly mean PBL heights (m) for June 2011, 14:00 LST shown in color.

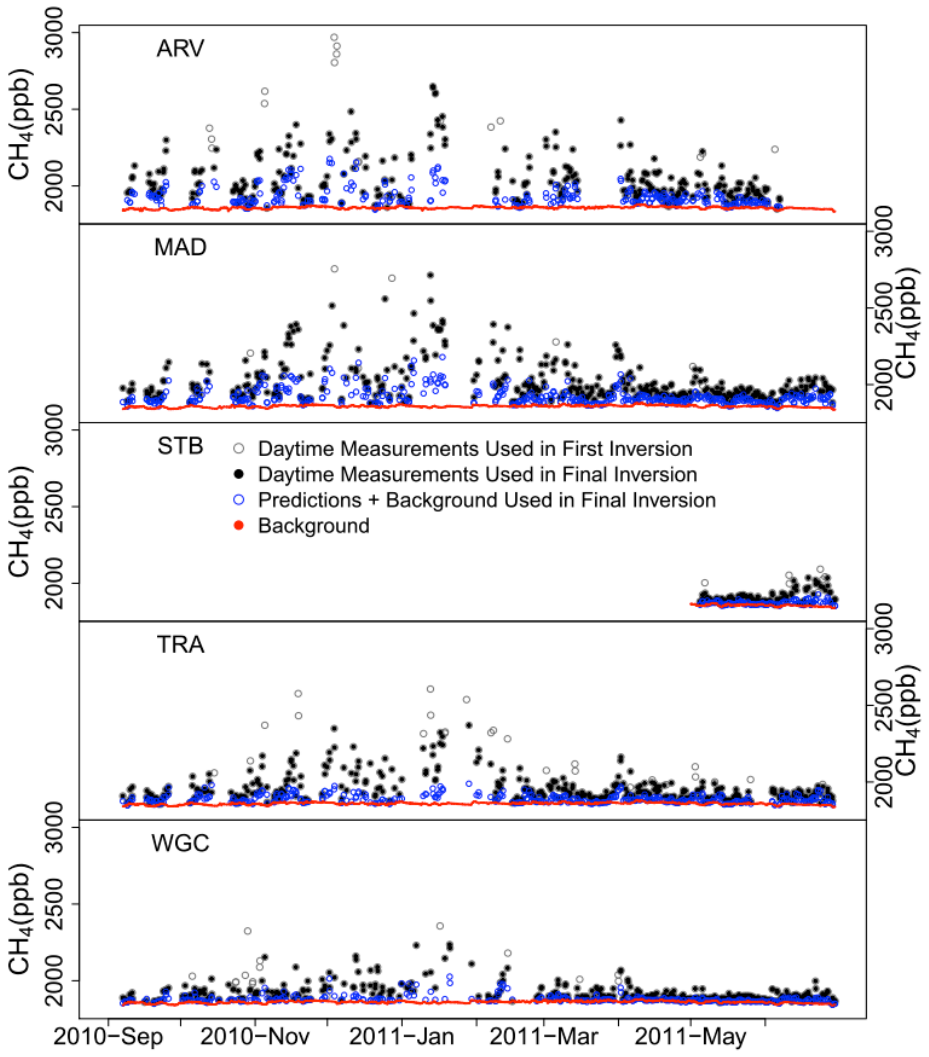

Figure 3. 3-hour mean CH_4 mixing ratio comparison: measured CH_4 mixing ratio during noon - afternoon hours used in the first inversion (gray open circle), measured CH_4 mixing ratio used in the final inversion (black filled circle), WRF-STILT predicted (before inversion) CH_4 mixing ratio using the CALGEM prior model + WRF-STILT predicted CH_4 background mixing ratio during noon – afternoon hours used for the final inversion (blue open circle), and WRF-STILT predicted CH_4 background mixing ratio using the 3-D curtain (red dots). Outliers were removed after the first inversion based on the data selection criteria described in Section 2.4.

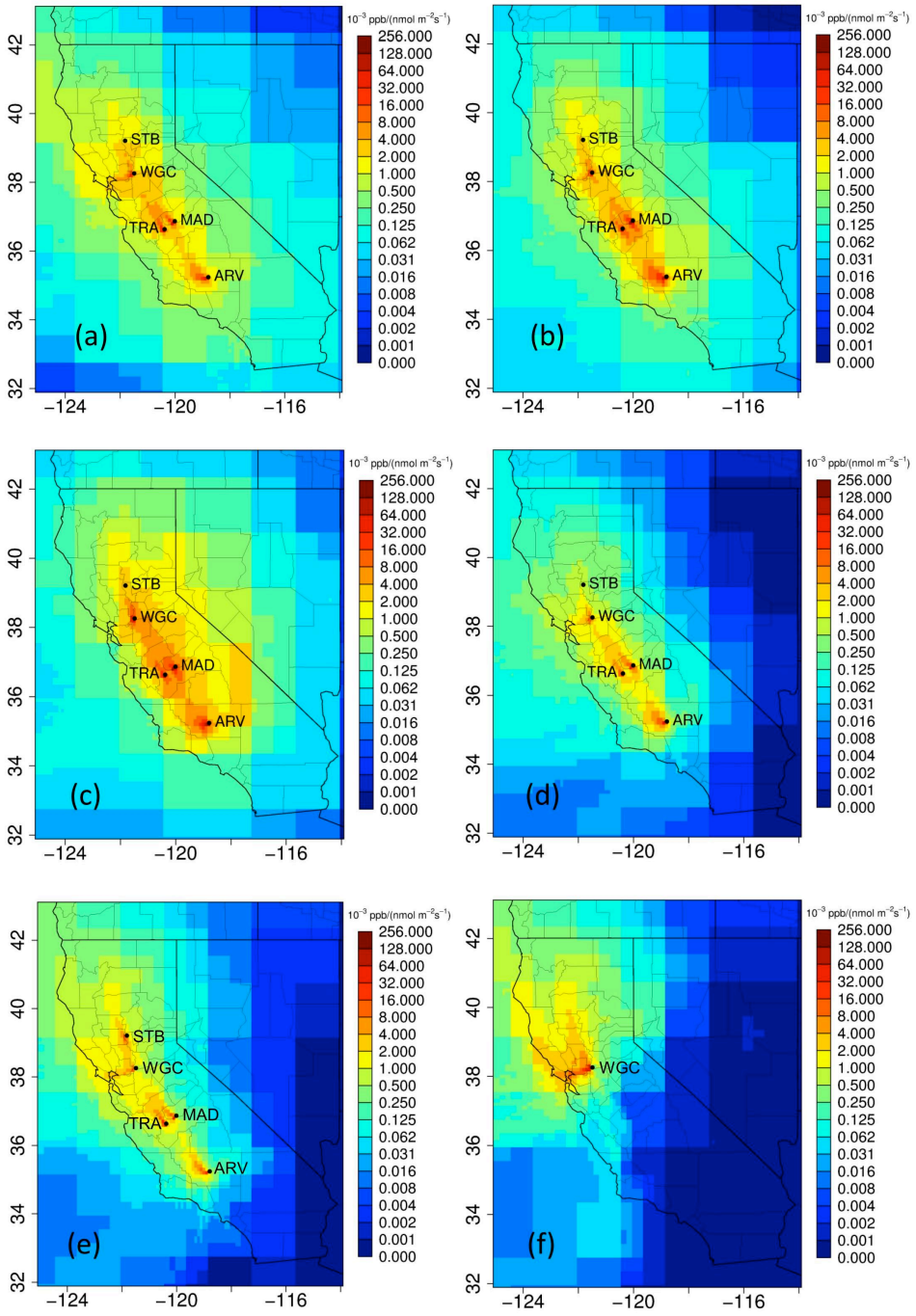

Figure 4. Seasonal mean footprints during the noon-afternoon hours for (a) September – October 2010, (b) November – December 2010, (c) January – February 2011, (d) March – April 2011, (e) May – June 2011 from all five sites, and (f) May – June 2011 from the WGC site only.

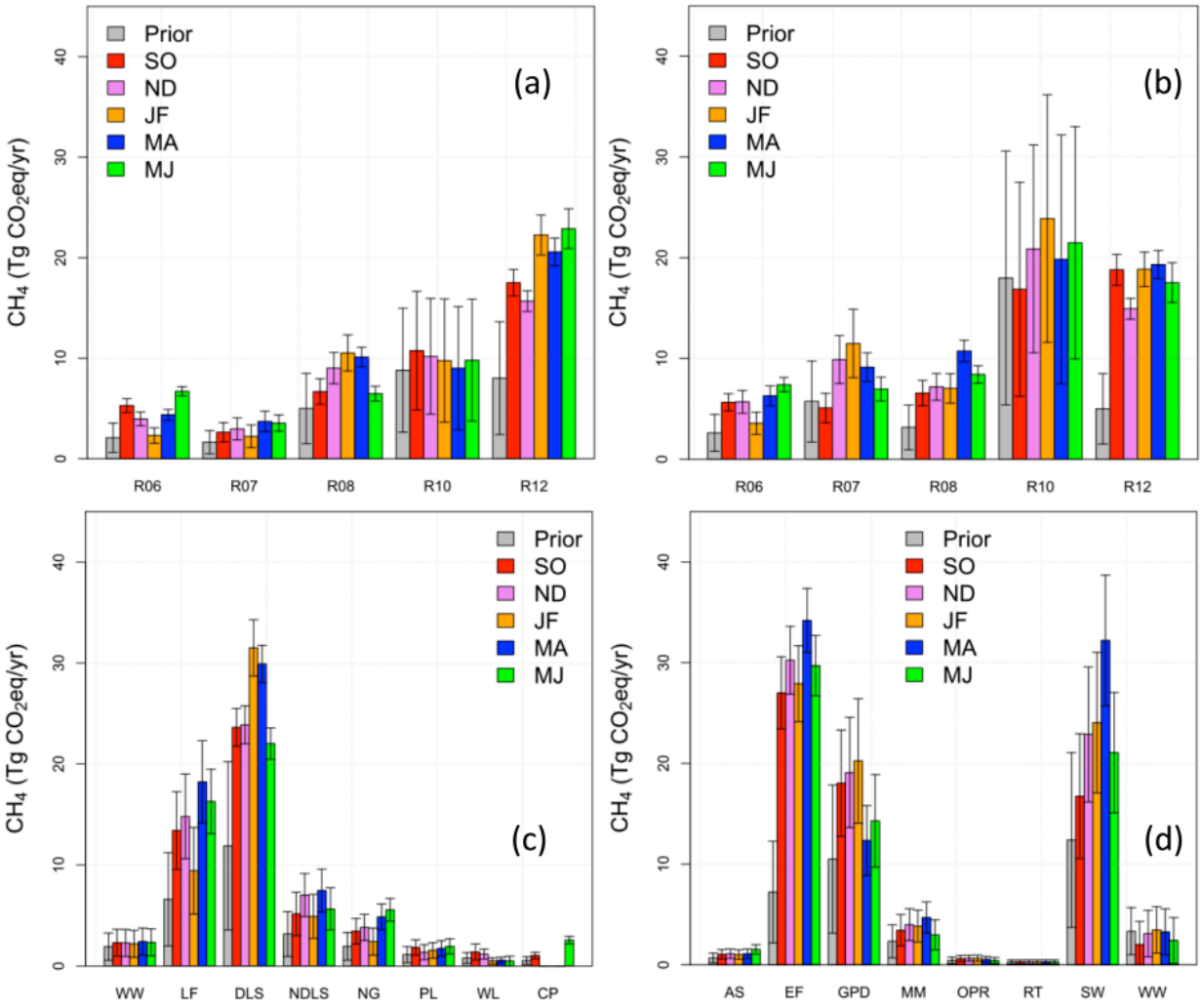

Figure 5. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) by season: (a) region analysis based on the CALGEM emission model, (b) region analysis based on EDGAR42, (c) source analysis based on the CALGEM emission model, and (d) source analysis based on EDGAR42. Only regions with significant emissions are shown. The annual mean prior (gray bar) is compared with posterior seasonal emissions (color bars). WW, LF, DLS, NDLS, NG, PL, WL and CP represent wastewater, landfill, dairy livestock, non-dairy livestock, natural gas, petroleum, wetland, and crop agriculture sources, respectively. AS, EF, GPD, MM, OPR, RT, SW, and WW represent agricultural soils, enteric fermentation, gas production and distribution, manure management, oil production and refineries, road transportation, solid waste, and wastewater, respectively.

DISCLAIMER