



HAL
open science

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.

► To cite this version:

Yu Yan Cui, Jerome Brioude, Wayne M Angevine, Jeff Peischl, Stuart A Mckeen, et al.. Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The San Joaquin Valley. *Journal of Geophysical Research: Atmospheres*, 2017, pp.3686-3699. 10.1002/2016JD026398 . hal-03135265

HAL Id: hal-03135265

<https://hal.univ-reunion.fr/hal-03135265v1>

Submitted on 8 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

3

4 Yu Yan Cui^{1,2}, Jerome Brioude^{1,2,3}, Wayne M. Angevine^{1,2}, Jeff Peischl^{1,2}, Stuart A.
5 McKeen^{1,2}, Si-Wan Kim^{1,2}, J. Andrew Neuman^{1,2}, Daven K. Henze⁴, Nicolas Bousseres⁴,
6 Marc L. Fischer⁵, Seongeun Jeong⁵, Hope A. Michelsen⁶, Ray P. Bambha⁶, Zhen Liu^{6,7},
7 Gregory W. Santoni⁸, Bruce C. Daube⁸, Eric A. Kort⁹, Gregory J. Frost², Thomas B.
8 Ryerson², Steven C. Wofsy⁸, Michael Trainer²

9

10 ¹Cooperative Institute for Research in Environmental Sciences, University of Colorado,
11 Boulder, CO, USA.

12 ²Chemical Sciences Division, Earth System Research Laboratory, NOAA, Boulder, CO,
13 USA.

14 ³Laboratoire de l'Atmosphere et des Cyclones, UMR8105, CNRS-Meteo France-
15 Universite La Reunion, La Reunion, France.

16 ⁴Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA.

17 ⁵Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory,
18 Berkeley, CA, USA.

19 ⁶Combustion Research Facility, Sandia National Laboratories, Livermore, CA, USA.

20 ⁷Now at Ramboll Environ US Corporation, Novato, CA, USA.

21 ⁸Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA.

22 ⁹Department of Climate and Space Sciences and Engineering, University of Michigan,
23 Ann Arbor, MI, USA.

24

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

26

27

28 **Key Points:**

- 29 • Estimate methane emissions in the San Joaquin Valley using inverse modeling
30 and a mass-balance approach
- 31 • Methane emissions are estimated to be greater than the bottom-up
32 inventory by a factor of 1.7
- 33 • Livestock largely account for differences between the optimized and prior
34 methane emission estimates^[1]_[SEPP]
35

36

37 **Index terms:**

- 38 0322 Constituent sources and sink
- 39 3260 Inverse theory
- 40 0345 Pollution: urban and regional
- 41 0365 Troposphere: composition and chemistry

42

43 **Keywords:**

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

46

47 **Abstract**

48 We quantify methane (CH₄) emissions in California's San Joaquin Valley (SJV) using
49 four days of aircraft measurements from a field campaign during May-June 2010 together
50 with a Bayesian inversion method and a mass-balance approach. For the inversion
51 estimates, we use the FLEXible PARTicle dispersion model (FLEXPART) to establish
52 the source-receptor relationship between sampled atmospheric concentrations and surface
53 fluxes. Our prior CH₄ emissions estimates are from the California Greenhouse Gas
54 Emissions Measurements (CALGEM) inventory. We use three meteorological
55 configurations to drive FLEXPART and subsequently construct three inversions to
56 analyze the final optimized estimates and their uncertainty (one standard deviation). We
57 conduct May and June inversions independently, and derive similar total CH₄ emissions
58 estimates for the SJV: 135±28 Mg/hr in May and 135±19 Mg/hr in June. The inversion
59 result is 1.7 times higher than the prior estimate from CALGEM. We also use an
60 independent mass-balance approach to estimate CH₄ emissions in the northern SJV for
61 one flight when meteorological conditions allowed. The mass-balance estimate provides
62 a confirmation of our inversion results, and these two independent estimates of the total
63 CH₄ emissions in the SJV are consistent with previous studies. In this study, we provide
64 optimized CH₄ emissions estimates at 0.1° horizontal resolution. Using independent
65 spatial information on major CH₄ sources, we estimate that livestock contribute 75–77%
66 and oil/gas production contributes 15–18% of the total CH₄ emissions in the SJV.
67 Livestock explain most of the discrepancies between the prior and the optimized
68 emissions from our inversion.

69

70 **1. Introduction**

71 Methane (CH₄) is the second most significant greenhouse gas. It has a large
72 global-warming potential and mediates global tropospheric chemistry. Globally, more
73 than 60% of total CH₄ emissions are attributed to human activities [EPA, 2015], such as
74 the natural gas and petroleum industries, domestic livestock operations, landfills, rice
75 cultivation, and coal mining. Reducing CH₄ from human activity is important for
76 reducing risks associated with climate change. As the most populous state of the US and
77 a major CH₄ emitter, California enacted State Assembly Bill 32
78 (<http://www.arb.ca.gov/cc/ab32/ab32.htm>) in 2006 to reduce greenhouse gas emissions to
79 1990 emission levels by the year 2020, and to reduce greenhouse gas emissions to 40
80 percent below 1990 levels by year 2030. Achieving this goal requires accurate accounting
81 of the magnitude and source attribution of CH₄ emissions.

82 The Central Valley covers about 14% of California's total land area and is the
83 leading dairy-farming and most productive agricultural region in California. Twenty
84 percent of US milk production occurs in California, mostly in the Central Valley
85 (<http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1103>).
86 The California Greenhouse Gas Emissions Measurements (CALGEM,
87 <http://calgem.lbl.gov>) project found that the Central Valley is the California region with
88 the highest CH₄ emissions [Zhao *et al.*, 2009; Jeong *et al.*, 2012; Jeong *et al.*, 2013]. The
89 San Joaquin Valley (SJV), the southern portion of the Central Valley, contains a variety
90 of potential CH₄ sources of anthropogenic origin, including approximately 2 million head
91 of cattle and calves [National Agricultural Statistics Service, 2013], more than 75,000
92 active oil wells, and many cities.

93 Current bottom-up inventories of CH₄ sources in the SJV are quite uncertain. The
94 Emission Database for Global Atmospheric Research (EDGAR) version 4.2 global
95 emission inventory at 0.1° × 0.1° horizontal resolution (<http://edgar.jrc.ec.europa.eu>)
96 reports that the CH₄ emissions from livestock in the SJV are 26.7 Mg/hr. However, a
97 bottom-up study from CALGEM at 0.1° × 0.1° horizontal resolution calculated CH₄
98 emissions from livestock in the San Joaquin Valley to be 60.4 Mg/hr, more than twice
99 that of EDGAR version 4.2 [Jeong *et al.*, 2013]. The SJV is also a significant region for
100 petroleum and natural gas production. A new bottom-up study from Jeong *et al.* [2014]
101 reports 3 to 7 times higher emissions from petroleum and natural gas production than the
102 California Air Resources Board (CARB) 2013 Oil and Gas Industry Survey Results and
103 2014 greenhouse gas emissions inventory.

104 To improve emissions quantification, atmospheric measurements have
105 increasingly been used to constrain the bottom-up emissions estimates. In the SJV, there
106 are ongoing studies using the tower measurements to estimate CH₄ emissions [Zhao *et al.*,
107 2009; Jeong *et al.*, 2013, 2016]. Current satellite data have been used to constrain CH₄ in
108 California, but CH₄ emissions estimates using satellite observations over the Central
109 Valley remain difficult because of the scarcity of observations [Wecht *et al.*, 2014;
110 Bousseret *et al.*, 2016].

111 A field campaign named the California Research at the Nexus of Air Quality and
112 Climate Change (CalNex, Ryerson *et al.*, 2013) took place in California during May and
113 June 2010. During CalNex, the NOAA WP-3 aircraft collected intensive measurements,
114 including CH₄ mixing ratios, over the South Coast Air Basin and the Central Valley. To
115 identify contributions from individual source categories, the aircraft flew close to

116 emission sources with extensive horizontal and vertical coverage. The CalNex aircraft
117 measurements provide a good opportunity to conduct a top-down estimate of the CH₄
118 emissions in these regions of California [Peischl *et al.*, 2013; Cui *et al.*, 2015]. The large
119 spatial coverage of the aircraft enables sampling of multiple CH₄ sources distributed
120 across the complex terrain of the SJV, providing a useful complement to ground-based
121 and remote-sensing measurements.

122 This study uses a mesoscale inverse modeling technique to estimate CH₄
123 emissions in the SJV based on aircraft measurements from CalNex. This mesoscale
124 inverse modeling system has already been employed to estimate CH₄ emissions in the
125 South Coast Air Basin of California [Cui *et al.*, 2015] using measurements from the same
126 campaign. The mass-balance approach [White *et al.*, 1976], an independent top-down
127 method, is applied in part of the SJV to provide confirmation of the inverse modeling
128 results. We compare our top-down CH₄ emissions estimates to three different inventories.
129 We also compare our results with another inversion analysis of the same region using
130 tower measurements [Jeong *et al.*, 2013, 2016].

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

134

135 **2. Methods**

136 In this section, we describe the atmospheric measurements of CH₄ mixing ratios
137 from the National Oceanic and Atmospheric Administration (NOAA) WP-3 aircraft. We
138 describe the prior CH₄ emission inventories, the construction of our atmospheric transport

139 model used to build the source-receptor relationships, and the design of our Bayesian
140 inverse modeling. The mass-balance approach, which provides an independent estimate
141 of CH₄ emissions based on the aircraft measurements, is described.

142

143 **2.1 Measurements**

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

153 CH₄ mixing ratios observed by the NOAA P-3 aircraft were measured once per
154 second using wavelength-scanned-cavity-ring-down spectroscopy (WS-CRDS; Picarro
155 1301 m) [*Peischl et al.*, 2012, 2013]. The precision of the 1-Hz CH₄ measurement is ±
156 1.4 ppbv, and accuracy is estimated at ±1.2 ppbv. We aggregate these observations into
157 30-s averages for use in the inversion framework, which, at a ground speed of
158 approximately 100 m s⁻¹, correspond to segments of about 3 km horizontally (Figure 2).
159 This aggregated dataset provides the receptor points in our backward trajectory
160 simulations from the atmospheric transport models described in Section 2.3 and is used in
161 an inverse-modeling analysis.

162

163 **2.2 Prior emission inventory**

164 A prior inventory provides critical information for Bayesian inversion modeling,
165 particularly when atmospheric measurements alone cannot fully constrain the spatial
166 distribution of the emissions sources. Inaccurate representation of the spatial distribution
167 of emissions sources in a prior limits the performance of inverse modeling [*Xiang et al.*,
168 2013]. Therefore, we need to select the best available inventory for the prior input. We
169 compared three available CH₄ inventories: a recent gridded top-down inventory based on
170 the US EPA National Emissions Inventory (NEI 2011, [https://www.epa.gov/air-](https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data)
171 [emissions-inventories/2011-national-emissions-inventory-nei-data](https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data)) [*Ahmadov et al.*,
172 2015], a recent gridded bottom-up inventory designed to be consistent with the US EPA
173 Inventory of US Greenhouse Gas Emissions and Sinks (GHGI) for 2012 [*Maasackers et*
174 *al.*, 2016], and a gridded bottom-up inventory from CALGEM designed to match the
175 CARB inventory for 2008 [*Jeong et al.*, 2012, 2013]. These three inventories provide
176 annual average CH₄ emissions estimates.

177 The spatial distributions of the three inventories are shown in Figure S1, and their
178 total CH₄ emissions for the SJV and its D1 and D2 sub-regions are listed in Table 1. The
179 three inventories' SJV total CH₄ emissions estimates range from 68-107 Mg/hr. We find
180 distinct variations between the three inventories' spatial distributions of CH₄ emissions
181 from livestock and active oil and gas wells. CALGEM, developed by *Zhao et al.* [2009]
182 and *Jeong et al.* [2012], relies on more detailed local information about source locations
183 and activity to generate the gridded CH₄ emissions estimates, compared with the other
184 two inventories based on EPA's NEI and GHGI. For example, CALGEM's spatial

185 distributions for livestock and oil/gas sources are based on the California Department of
186 Water Resources land-use survey database [*Salas et al.*, 2009] and the California
187 Department of Conservation's Division of Oil, Gas, and Geothermal Resources database
188 (http://www.conservation.ca.gov/dog/pubs_stats/annual_reports/Pages/annual_reports.aspx
189 x), respectively. Among the three inventories considered, CALGEM contains the most
190 accurate spatial distributions for the major CH₄ sources in the Central Valley, and we
191 therefore use CALGEM as the foundation of our prior inventory. We also update the
192 oil/gas source sector of CALGEM in the SJV according to emissions from *Jeong et al.*
193 [2014]. The CALGEM inventory is available at 0.1° × 0.1° spatial resolution, and we
194 optimize the inventory at the same resolution.

195 Similar to *Cui et al.* [2015], our study adjusts the magnitude of total CH₄
196 emissions in each grid cell of the prior annual average inventory, without differentiating
197 source sectors. When we calculate the contributions from different source sectors
198 independently, we require extra spatial information. Figure 1 (B) presents the spatial
199 information for the two dominant CH₄ sources in the SJV: dairies (an important
200 livestock-related activity across the SJV) and active oil/gas wells [*Jeong et al.*, 2013].
201 Like CALGEM, we obtained the spatial information for livestock sources from *Salas et*
202 *al.* [2008], and the spatial distribution of the active oil and gas wells was taken from
203 California's Department of Conservation Division of Oil, Gas, and Geothermal
204 Resources database
205 (http://www.conservation.ca.gov/dog/pubs_stats/annual_reports/Pages/annual_reports.aspx
206 x). Livestock sources are highly concentrated in both the D1 and D2 sub-regions. Oil and
207 gas production is mainly found in the southern part of D1. In the SJV, the oil and gas

208 production sector has much larger CH₄ emissions than oil/gas processing, transmission
209 and distribution [Jeong *et al.*, 2014].

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

220

221 **2.3 Atmospheric transport modeling**

222 Following Cui *et al.*, [2015], the FLEXPART-WRF Lagrangian model version 3.1
223 [Brioude *et al.*, 2013] is used to calculate source-receptor relationships, a.k.a. footprints.
224 The surface footprints (s m² kg⁻¹) represent the residence time within a surface layer
225 (below 100 m above ground level) weighted by the atmospheric density. We conducted
226 three atmospheric transport simulations using FLEXPART driven by three different
227 meteorology configurations from the Weather Research Forecasting Model (WRF)
228 (Table 2). The three WRF meteorological fields have a 4 x 4 km horizontal grid spacing.
229 The first and second meteorology configurations (WRF1 and WRF2) are from Angevine
230 *et al.* [2012]. The third WRF configuration (WRF3) is from Kim *et al.* [2016]. Using

231 measurements from the same field campaign, WRF1 and WRF2 have been used to
232 estimate nitrous oxide emissions in the Central Valley [*Xiang et al.*, 2013], and WRF3
233 has been used to estimate ozone in the Los Angeles region [*Kim et al.*, 2016]. Detailed
234 information on evaluations of planetary boundary layer height (PBLH), wind speed, and
235 wind direction from the three transport models can be found in *Angevine et al.* [2012] and
236 *Kim et al.*, [2016]. Here we show model evaluations using observations from the four
237 flights in Figures S2-S4 and Table S1.

238 Correlations between any of the three CH₄ simulations with differing
239 meteorological configurations are no larger than the correlations between any model
240 simulation and the observations. Therefore, the three model simulations can be treated as
241 independent representations of the meteorology. Each model is used in our inverse
242 modeling system to derive the posterior emissions estimates, and the final optimized
243 emissions estimates are based on the mean value from the three estimates. Three
244 meteorological models can only represent part of the phase space of model uncertainties.
245 A complete estimate of transport model uncertainty would require a larger ensemble and
246 more comprehensive characterization [*Angevine et al.*, 2014].

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

252 Figure 4 presents the mean vertical profiles of CH₄ mixing ratios in 100-m vertical
253 intervals over the SJV from the aircraft measurements and from the three transport

254 models using the CH₄ prior inventory. The error bars represent the standard deviations
255 among the three different transport models. There is no obvious bias in the simulated
256 vertical mixing. There is a small bias in simulating CH₄ in the upper part of the mean
257 profile, but the bias is statistically insignificant as it is smaller than the uncertainty range
258 of the CH₄ background determination (see next section). There is a systematic low bias in
259 the modeled CH₄ concentrations below 1600-1800 m above sea level (ASL), which is
260 attributed to a bias in the prior emissions estimates as shown below.

261

262 **2.4 Bayesian inverse modeling**

263 We perform a 4-dimensional (three spatial dimensions in the model plus time)
264 inversion using a Bayesian framework by minimizing a cost function assuming
265 lognormal distributions for the observed enhancements and surface fluxes [*Brioude et al.*,
266 2011]. The cost function used in the inversion framework is

$$267 J = \frac{1}{2}(\ln(y_0) - \ln(Hx))^T R^{-1}(\ln(y_0) - \ln(Hx)) + \frac{1}{2}\alpha(\ln(x) - \ln(x_b))^T B^{-1}(\ln(x) -$$

268 $\ln(x_b)),$

269 where y_0 is the measured time series of CH₄ mixing ratio enhancement above defined
270 background, H is the source-receptor relationship matrix calculated by FLEXPART-
271 WRF, R and B are the error covariance matrices of the model-observation mismatch and
272 the prior information, respectively, x_b is the prior emission inventory, and x is the
273 posterior emission inventory to be determined. The parameter α [*Henze et al.*, 2009]
274 balances the errors of both covariance matrices in the minimization of the cost function to
275 calculate the best estimates of emissions.

276 The surface emissions optimization applied in this study is based on the inverse

277 modeling framework applied in *Cui et al.* [2015]. Most CH₄ mixing ratio enhancements
278 were measured below 2.0 km altitude ASL during the four flights. To reduce the potential
279 uncertainty in the transport models' ability to distinguish between the PBL and the free
280 troposphere, we focus on the measurements (i.e., receptor points) below 1.5 km ASL
281 (Figure 2). Choosing a threshold of 2.0 km or 1.0 km ASL does not significantly affect
282 our results.

283 For each flight, we plot the histogram of the observed CH₄ mixing ratios below
284 1.5 km ASL on the upwind side of the domain. We choose the mode of this distribution
285 as the background value. Based on the width of this distribution, we estimate a 10 ppbv
286 uncertainty in the background mixing ratio for each flight.

287 The NOAA P-3 flights over the SJV flew close to surface sources, so that the
288 measurements were obtained within hours from the time of emission. Therefore, it is
289 reasonable to assume that photochemical loss of CH₄ can be neglected. Hence, CH₄ is
290 treated as a passive tracer in our mesoscale inverse system.

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

297 The \mathbf{R} and \mathbf{B} covariance matrices are assumed to be diagonal matrices. \mathbf{R} is
298 calculated by the addition in quadrature of the 30-second aggregation uncertainty (i.e.,
299 the standard deviation of a 30-s interval, 10 ppbv for the mean value), the background

300 uncertainty (10 ppbv), and the uncertainty of each transport model (50% [Angevine et al.,
301 2014], 50 ppbv for the mean value) in simulating CH₄ enhancements above background.
302 The largest uncertainty in \mathbf{R} is that of the transport models. We assume a larger
303 uncertainty in the models in this study than in the Los Angeles basin [Cui et al., 2015]
304 because of the inherent difficulty in modeling the transport within the complex terrain of
305 the Central Valley.

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

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

321 For each sub-region, the total emissions estimate is calculated by summing the
322 emissions estimates of the clusters in the region. The total uncertainty estimate for each

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

332

333 **2.5 Mass-balance approach**

334 CH₄ emission fluxes were determined using the mass-balance approach [*White et*
335 *al.*, 1976] for comparison with the inversions. In this study, we use this approach to
336 quantify CH₄ emissions using measurements made both upwind and downwind of the
337 emission sources. We estimate the total CH₄ emissions from the D2 sub-region of the
338 SJV when favorable meteorological conditions were observed, including steady
339 horizontal winds, and a well-developed PBL that was well mixed vertically. The
340 uncertainties associated with the assumptions of the technique are included. The details
341 of the mass-balance approach are described in *Peischl et al.*, [2015].

342

343 **3. Results and discussion**

344 **3.1 San Joaquin Valley CH₄ emissions estimates from the inversions**

345 We optimize the spatially resolved CH₄ emissions estimates in the SJV using the

346 mesoscale inverse modeling system with the CalNex airborne measurements (Figure 5).
347 The optimized estimates are from two independent inversions using observations in the
348 May and June 2010 flights. The May and June inversions derive similar total CH₄
349 emissions estimates for the SJV (Table 1). We estimate the total CH₄ emissions from the
350 SJV to be 135±28 Mg/hr in May 2010 and 135±19 Mg/hr in June 2010. The difference in
351 total emissions between May and June is statistically insignificant. In general, the spatial
352 patterns of the CH₄ prior inventory are consistent with those of the optimized emissions
353 estimates (Figure 5). However, the optimized emissions in May and June both indicate
354 that the magnitudes of the prior emissions in the SJV are much lower than the optimized
355 estimates (Figure 5 (B) and (D)). The highest emission rates (and the largest adjustments
356 to the prior) are seen in the region from Hanford to Visalia in the southern sub-region (D1)
357 and from Merced to Stanislaus in the northern sub-region (D2) of the SJV. Our optimized
358 estimates on average in the SJV are higher by a factor of 1.7 than the prior estimates
359 based on the CALGEM inventory.

360 The optimized total CH₄ emissions estimates from each transport model are
361 shown in Table 3. The transport model evaluations shown in Table S1 indicate that
362 WRF3 has a large (57%) bias in simulating PBLH in D2 for the May inversion case.
363 Therefore, in Table 3 we also list the overall estimates based only on WRF1 and WRF2
364 simulations. In May, using only these two simulations results in only a 10% difference in
365 estimated SJV CH₄ emissions compared with the results based on three WRF simulations;
366 differences in June are much smaller. We therefore base our main conclusions on results
367 from the three WRF simulations for both May and June.

368 To evaluate the optimized emissions, we compare the measured CH₄

369 enhancements above background and those simulated by FLEXPART-WRF using the
370 optimized emissions estimates and the prior estimates (Figure 6 and Figure 7, Table 1).
371 The FLEXPART-WRF simulation using the optimized emissions captures the
372 observations with a coefficient of determination (r^2) of 0.76 and 0.71 for the May and
373 June inversions, respectively. These correlations are higher than for the simulations using
374 the prior estimates ($r^2 = 0.49$ and 0.47 , respectively). Moreover, there is a large decrease
375 in the mean bias using the optimized emissions. The mean biases between the observed
376 and simulated CH₄ enhancements using the prior inventory in the May and June
377 inversions are -55.2 and -31.8 ppbv, respectively. In contrast, the observed-simulated
378 biases using the optimized emissions are only -9.1 and -5.5 ppbv, respectively, an 83%
379 decrease for both inversions compared to the corresponding results based on the prior
380 inventory. Additionally, the vertical profiles of CH₄ mixing ratios are well captured by
381 the models when we use the optimized CH₄ emissions estimates (Figure 4).

382 We compare optimized emissions estimates in the present study to the top-down
383 estimate from *Jeong et al.* [2013, 2016] (Table 1). The total emissions estimates for the
384 SJV in this study are similar to estimates from *Jeong et al.* [2016] (98-170 Mg CH₄/hr).
385 In this study, we use many more grid clusters than the number of grid cells in *Jeong et al.*,
386 [2013] to invert for the surface fluxes in the SJV. The total emissions estimates are
387 similar, while the partitioning of CH₄ emissions between sub-regions D1 and D2 differ
388 between our study and *Jeong et al.* [2013]. We estimate total CH₄ emissions from D1 to
389 be 80 ± 17 Mg/hr in May and 79 ± 17 Mg/hr in June (Table 1), and the total CH₄ emissions
390 from D2 to be 55 ± 18 Mg/hr in May and 56 ± 13 Mg/hr in June. The differences between
391 May and June are statistically insignificant. The estimated emissions for D1 are lower

392 than those of *Jeong et al.* [2013], while those for D2 are higher on average than those of
393 *Jeong et al.* [2013]. *Jeong et al.* [2013] only used two grid cells to represent the domain
394 of the SJV in their inversions, while we substantially improved the spatial resolution by
395 aggregating 4544 grid cells ($0.1^\circ \times 0.1^\circ$) into 2024 clusters. The difference in spatial
396 resolution between the two studies results in different transport and emissions estimates.

397

398 **3.2 San Joaquin Valley CH₄ emissions estimates from the mass-balance approach**

399 We use the same CalNex aircraft measurements and an independent mass-balance
400 approach to derive CH₄ emissions from the SJV. We determined emissions in the
401 northern SJV sub-region (D2) using measurements from the May 12 flight, the only day
402 with favorable meteorological conditions in the Central Valley during CalNex.

403 On the May 12 flight, the upwind transect in San Joaquin County (Figure 1 (C))
404 resulted in a CH₄ flux of 28 ± 19 Mg/hr (1-sigma uncertainty) coming mainly from the
405 nearby Sacramento Valley. The downwind transect in Merced County resulted in a flux
406 of 97 ± 45 Mg/hr. The difference between the upwind and downwind transects, 69 ± 47 Mg
407 CH₄/hr, represents the estimated emissions from sub-region D2, assuming the upwind
408 sources were constant while the wind traveled from the upwind transect to the downwind
409 transect. Details of the mass-balance calculation are given in Table 4. Within the stated
410 uncertainties, the mass-balance emissions estimate agrees with our inversion in D2
411 (55 ± 18 Mg CH₄/hr in May). Therefore, an independent method purely based on the
412 measurements confirms our optimized inversion results.

413 We did not conduct a mass-balance analysis for the southern SJV region (D1) in
414 this study because CH₄ surface emissions from D2 strongly influenced CH₄ in D1 (Figure

415 8). In addition, the nighttime Fresno eddy [Bao *et al.*, 2008] complicates the application
416 of a mass-balance approach to the flights over D1, such as leading to a build-up of CH₄
417 enhancements in the entire domain the following day and violating the steady wind
418 assumption. Therefore, favorable conditions for mass-balance estimates in D1 are
419 difficult to obtain during CalNex. Similarly, winds over the D1 and D2 regions during the
420 June flights had a westerly component that transported emissions through the eastern
421 edge of the San Joaquin Valley and beyond the extent of the downwind flight legs, so we
422 could not carry out mass-balance estimates using the June flights. These limitations to
423 using the mass-balance approach in the SJV show the value of inverse modeling
424 estimates for the region.

425

426 **3.3 Major source contributions in the San Joaquin Valley**

427 Livestock sources (including dairies and animal feeding operations) are the largest
428 source of CH₄ emissions in both sub-regions of the San Joaquin Valley. Livestock and
429 oil/gas production sources are rarely collocated in the same 0.1° grid cell. In the few
430 cases where a grid cell contains more than one CH₄ source, the source type of the cell is
431 determined by the dominant source. Combining our optimized 0.1° resolution CH₄
432 emissions estimates (Figure 5) and the locations of two major sources (Figure 1 (B)), we
433 estimate the CH₄ emissions from livestock sources in the SJV to be 103±29 Mg/hr and
434 105±25 Mg/hr for May and June, respectively (Table 5), which are higher than the prior
435 CH₄ emissions by a factor of 1.8. Livestock emissions contribute 75–77 % of the total
436 CH₄ emissions in the SJV according to our optimized results on average. Our estimates
437 are consistent with the analysis of Jeong *et al.* [2016], who estimate SJV CH₄ emissions

438 from the livestock source sector are 81-177 Mg/hr. Moreover, our finding for livestock
439 sources is consistent with the analysis of *Johnson et al.* [2016], who estimate a factor of 2
440 higher emissions from a top-down approach compared with the CALGEM inventory.

441 Active oil/gas wells are mainly located in the southern SJV (Figure 1 (B)). We
442 estimate CH₄ emissions in the SJV from the active oil/gas wells to be 24±11 Mg/hr in
443 May and 21±7 Mg/hr in June (Table 4), which are higher than the prior CH₄ emissions by
444 a factor of 1.6. On average, the wells emissions contribute 15–18% of the total CH₄
445 emissions in the SJV according to our optimized results. Our results are in agreement
446 with the *Jeong et al.* [2014, 2016] estimates of 19 Mg/hr from oil and natural gas
447 production in the SJV.

448 We also calculate the fractional adjustment in each of the two sources relative to
449 the fractional change between the prior and optimized estimates of the SJV total CH₄
450 emissions. On average, livestock sources explain 82-86% of the discrepancy between our
451 prior and optimized estimates, while oil/gas production explains 13-18% of the
452 discrepancy.

453

454 **4. Conclusions**

455 Using airborne measurements collected during the CalNex 2010 study, we apply a
456 mesoscale inverse model to perform a top-down estimate of CH₄ emissions in the San
457 Joaquin Valley of California. Our optimized estimates of total CH₄ emissions in the San
458 Joaquin Valley in May 2010 (June 2010) are 135±28 (135±19) Mg CH₄/hr. Our
459 optimized CH₄ emissions estimates are higher by a factor of 1.7 than the prior estimates
460 based on CALGEM.

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

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

476 Compared with the prior CALGEM inventory, our optimized estimates for
477 CH₄ emissions from livestock sources are higher by a factor of 1.8, while our optimized
478 CH₄ emissions from oil/gas production are higher by a factor of 1.6. Livestock are the
479 most important source of CH₄ emissions in the SJV, and we find that livestock sources
480 explain most of the discrepancies between the prior and our optimized CH₄ emissions
481 estimates. Our use of high-frequency aircraft observations and a model with high spatial
482 resolution allow us to distinguish signals from livestock and oil/gas sources and to
483 provide a quantitative top-down constraint on the emissions from these sectors.

484 To validate our optimized emissions estimates, we also conduct a mass-balance
485 estimate for one flight and one sub-region as an independent approach. Our optimized
486 estimates are in agreement with the mass-balance estimate within the combined
487 uncertainty of the two approaches. The mass-balance method using aircraft observations
488 can be used to estimate emissions from a region under favorable meteorological
489 conditions, but such conditions do not always occur. For instance, no mass-balance
490 estimates could be performed for the southern SJV in this study. Mesoscale inverse
491 modeling therefore offers a reliable, complementary technique for quantifying emissions
492 from multiple CH₄ sources over a large area.

493 Our inversions based on high quality aircraft measurements provide estimates of
494 CH₄ emissions in the San Joaquin Valley that agree with previous inversion calculations
495 based on tall tower observations. These independent top-down estimates confirm that
496 major CH₄ sources in the Valley are underestimated by the CALGEM prior inventory.
497 This study shows that applying an inverse model to tower and aircraft measurements to
498 assess and improve emissions estimates can inform bottom-up inventories and could
499 ultimately be useful in evaluating emissions reduction strategies.

500

501 **Acknowledgments:**

502 FLEXPART-WRF model is available at the official FLEXPART website
503 (<http://flexpart.eu>). NOAA P-3 data are available and can be downloaded at
504 <http://www.esrl.noaa.gov/csd/projects/calnex>. The optimized emission inventory is
505 available online as supporting information in NetCDF format. The lognormal Bayesian
506 inverse software was developed at NOAA/ESRL/CSD and CIRES. The WRF
507 initial/boundary data were provided by ERA-Interim and NOAA/NCEP. U.S. EPA NEI
508 2011 provided information that was compared to our prior inventory. We thank NOAA's
509 High Performance Computing Program for their support in running FLEXPART-WRF.
510 This work was supported in part by NOAA's Atmospheric Chemistry, Carbon Cycle, and
511 Climate Program. J.B., D.K.H., N.B., and M.T. acknowledge support from the NOAA
512 Climate Program Office (CPO) (NA14OAR4310136). M.L.F. and S.J. acknowledge
513 support from the California Energy Commission Public Interest Environmental Research
514 Program to LBNL under contract no. DE-AC02-05CH11231. Z.L., R.P.B., and H.A.M
515 are supported under the Laboratory Directed Research and Development program at
516 Sandia National Laboratories. Sandia is a multi-mission laboratory managed and operated
517 by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Company, for the
518 United States Department of Energy's National Nuclear Security Administration under
519 contract DEAC04-94AL85000.

520 **References:**

521

522 Ahmadov, R., S. McKeen, M. Trainer, R. Banta, A. Brewer, S. Brown, P.M. Edwards,
523 J.A. de Gouw, G.J. Frost, J. Gilman, D. Helmig, B. Johnson, A. Karion, A. Koss, A.
524 Langford, B. Lerner, J. Olson, S. Oltmans, J. Peischl, G. Pétron, Y. Pichugina, J.M.
525 Roberts, T. Ryerson, R. Schnell, C. Senff, C. Sweeney, C. Thompson, P. Veres, C.
526 Warneke, R. Wild, E.J. Williams, B. Yuan, and R. Zamora, Understanding high
527 wintertime ozone pollution events in an oil and natural gas producing region of the
528 western US, *Atmospheric Chemistry and Physics*, doi:10.5194/acp-15-411-2015,
529 2015.

530

531 Angevine, W. M., L. Eddington, K. Durkee, C. Fairall, L. Bianco, and J. Brioude (2012),
532 Meteorological model evaluation for CalNex 2010, *Mon. Weather Rev.*, 140, 3885–
533 3906, doi:10.1175/MWR-D-12-00042.1.

534

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

538

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

543

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

547

548 Bousserez, N., D. K. Henze, B. Rooney, A. Perkins, K. J. Wecht, A. J. Turner, V. Natraj,
549 J. R. Worden (2016), Constraints on methane emissions in North America from future
550 geostationary remote sensing measurements, *Atmos. Chem. Phys.*, 16, 6175–
551 6190, doi:10.5194/acp-16-6175-2016.

552

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

556

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

561

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

566

567 Cui, Y. Y., J. Brioude, S. A. McKeen, W. M. Angevine, S.-W. Kim, G. J. Frost, R.
568 Ahmadov, J. Peischl, N. Bousserez, Z. Liu, T. B. Ryerson, S. C. Wofsy, G. W.
569 Santoni, E. A. Kort, M. L. Fischer, and M. Trainer (2015), Top-down estimate of
570 methane emissions in California using a mesoscale inverse modeling technique: The
571 South Coast Air Basin. *J. Geophys. Res. Atmos.*, 120, 6698–6711.
572 doi: 10.1002/2014JD023002.
573

574 Dudhia, Jimmy, (1996), A multi-layer soil temperature model for MM5 the Sixth
575 PSU/NCAR Mesoscale Model Users' Workshop.
576
577

578 EPA (2015), Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013,
579 [http://epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2015-Main-](http://epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2015-Main-Text.pdf)
580 [Text.pdf](http://epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2015-Main-Text.pdf)
581

582 Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with explicit
583 treatment of entrainment processes. *Mon. Wea. Rev.*, 134, 2318–2341.
584

585 Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012),
586 Seasonal variation of CH₄ emissions from central California, *J. Geophys. Res.*, 117,
587 D11306, doi:10.1029/2011JD016896.
588

589 Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak and M. L.
590 Fischer (2013), A Multi-tower Measurement Network Estimate of California's
591 Methane Emissions. *J. Geophys. Res.*, doi: 10.1002/jgrd.50854.
592

593 Jeong, S., D. Millstein and M. L. Fischer (2014), Spatially Explicit Methane Emissions
594 from Petroleum Production and the Natural Gas System in California, *Env. Sci.*
595 *Technol.*, doi: 10.1021/es4046692.
596

597 Jeong, S., et al. (2016), Estimating methane emissions in California's urban and rural
598 regions using multi-tower observations, *J. Geophys. Res. Atmos.*, 121,
599 doi:10.1002/2016JD025404.
600

601 Johnson et al., (2016), Investigating seasonal methane emissions in northern California
602 using airborne measurements and inverse modeling, *J. Geophys. Res. Atmos.*, doi:
603 10.1002/2016JD025157.
604

605 Kalnay, E., M. Kanamitsu, and W. E. Baker (1990), Global numerical weather prediction
606 at the National Meteorological Center, *Bull. Amer. Meteor. Soc.*, 71, 1410–1428, doi:
607 10.1175/1520-0477(1990)071<1410:GNWPAT>2.0.CO;2.
608

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

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

614

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

617

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

621

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

625

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

632

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

639

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

645

646 Ryerson, T. B., A. E. Andrews, W. M. Angevine, T. S. Bates, C. A. Brock, B. Cairns, R.
647 C. Cohen, O. R. Cooper, J. A. de Gouw, F. C. Fehsenfeld, R. A. Ferrare, M. L.
648 Fischer, R. C. Flagan, A. H. Goldstein, J. W. Hair, R. M. Hardesty, C. A. Hostetler, J.
649 L. Jimenez, A. O. Langford, E. McCauley, S. A. McKeen, L. T. Molina, A. Nenes, S.
650 J. Oltmans, D. D. Parrish, J. R. Pederson, R. B. Pierce, K. Prather, P. K. Quinn, J. H.
651 Seinfeld, C. J. Senff, A. Sorooshian, J. Stutz, J. D. Surratt, M. Trainer, R. Volkamer,
652 E. J. Williams, S. C. Wofsy (2013), The 2010 California Research at the Nexus of Air
653 Quality and Climate Change (CalNex) field study, *J. Geophys. Res. Atmos.*, 118,
654 5830–5866, doi:10.1002/jgrd.50331.

655

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

658 Commission, PIER Energy-Related Environmental Research, CEC-500-2008-093,
659 [http://www.energy.ca.gov/2008publications/CEC-500-2008-093/CEC-500-2008-](http://www.energy.ca.gov/2008publications/CEC-500-2008-093/CEC-500-2008-093.PDF)
660 [093.PDF](http://www.energy.ca.gov/2008publications/CEC-500-2008-093/CEC-500-2008-093.PDF).
661
662 Salas, W., C. Li, F. Mitloehner, and J. Pisano (2009), Developing and applying process-
663 based models for estimating greenhouse gas and air emissions from California dairies,
664 Rep. CEC-500-2008-093, Public Interest Energy Res. Program, Calif. Energy Comm.,
665 Sacramento, Calif.
666
667 Wecht, K. J., D. J. Jacob, M. P. Sulprizio, G. W. Santoni, S. C. Wofsy, R. Parker, H.
668 Bösch, and J. Worden (2014), Spatially resolving methane emissions in California:
669 constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and
670 future (TROPOMI, geostationary) satellite observations, *Atmos. Chem. Phys.*, 14,
671 8173-8184, doi:10.5194/acp-14-8173-2014.
672
673 White, W., J. Anderson, D. Blumenthal, R. Husar, N. Gillani, J. Husar, and W. Wilson
674 (1976), Formation and transport of secondary air pollutants: Ozone and aerosols in
675 the St. Louis urban plume, *Science*, 194, 187–189, doi: 10.1126/science.959846.
676
677 Xiang, B., S. M. Miller, E. A. Kort, G. W. Santoni, B. C. Daube, Bruce C. R. Commane,
678 W. M. Angevine, T. B. Ryerson, M. K. Trainer, A. E. Andrews, T. Nehr Korn, H. Tian,
679 and S. C. Wofsy, (2013), Nitrous oxide (N₂O) emissions from California based on
680 2010 CalNex airborne measurements, *J. Geophys. Res. Atmos.*, 118, 2809–2820,
681 doi:10.1002/jgrd.50189.
682
683 Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T.
684 Nehr Korn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane
685 emissions from Central California, *J. Geophys. Res.*, 114, D16302,
686 doi:10.1029/2008JD011671.
687

688
689

Table 1. Comparison of Total CH₄ 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	-	-	-

690

691
692
693
694
695
696

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

697
698
699
700
701
702
703
704
705

^{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].

706
707
708
709
710

711
712
713
714

Table 3. Optimized CH₄ 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				

715
716
717
718
719
720
721
722
723
724
725
726
727
728

- 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).

729
730
731

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

732

733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750

751
752
753
754
755
756
757

Table 5. Prior and Optimized CH₄ Emissions from two Major Source Sectors and Their Contributions to the San Joaquin Valley.

	Prior (Mg/hr)	Livestock				Prior (Mg/hr)	Oil/Gas			
		Inversion (Mg/hr)		Contribution			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%	

758
759
760
761

- 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).

762 **Figure 1.** (A) The San Joaquin Valley (SJV) and two sub-regions, the Southern
763 SJV (D1) and the Northern SJV (D2). The background map is the prior inventory
764 of CH₄ emissions used in this study based on CALGEM, showing the annual
765 average emissions rate (unit: $\mu\text{g s}^{-1} \text{m}^{-2}$). (B) The spatial distribution of the two
766 major CH₄ sources in the SJV: livestock and active oil/gas wells. (C) Two NOAA
767 P-3 flight tracks over the SJV in May 2010. The black rectangles highlight the
768 locations of the upwind transect in San Joaquin County and the downwind
769 transect in Merced County used in the mass-balance estimate. (D) Two NOAA P-
770 3 flight tracks over the SJV in June 2010.

771 **Figure 2.** Airborne measurements of CH₄ mixing ratios (averaged over 30 s) in
772 the San Joaquin Valley, at 0-1500 m ASL and excluding measurements taken
773 over the ocean and during takeoff and landing from the Los Angeles area. Each
774 data point represents a receptor for the inverse modeling.
775

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

782 **Figure 4.** Vertical profiles of 100-m averaged measurements of CH₄
783 enhancement mixing ratios, ΔCH_4 , (measured mixing ratios in Figure 2 above a
784 background derived for each flight; see text for details), simulations of Δch_4 from
785 FLEXPART-WRF using the prior and optimized emission estimates in the San
786 Joaquin Valley for May (left) and June (right) 2010. The error bars represent the
787 standard deviations (1-sigma) of simulations from the three different transport
788 models.

789
790 **Figure 5.** Two-dimensional maps of CH₄ emissions estimates in the San Joaquin
791 Valley from this study. (A) and (C) are average optimized emissions using the
792 airborne measurements from two May flights and two June flights, respectively.
793 (B) and (D) are the corresponding differences between the optimized emissions
794 estimates and the prior emission inventory in Figure 1(A).
795

796 **Figure 6.** Airborne measurements of CH₄ enhancement mixing ratios, ΔCH_4 ,
797 (measured mixing ratios in Figure 2 above a background derived for each flight;
798 see text for details) (black line), simulations of ΔCH_4 from FLEXPART-WRF
799 based on the prior inventory (blue lines), and simulations from FLEXPART-WRF
800 based on the optimized emissions (red lines). Solid lines are average values
801 based on the three transport models, and shading represents the standard
802 deviation (1-sigma) of three transport models.
803

804 **Figure 7.** The relationship between observed and simulated CH₄ enhancement
805 mixing ratios for the May (left) and June (right) flights. The simulated data points

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

810 **Figure 8.** CH₄ enhancement mixing ratios simulated by the FLEXPART-WRF
811 model based on the optimized CH₄ emissions from the whole domain (All, green
812 lines) and due to CH₄ emissions from only one specific sub-region (either D1 or
813 D2). Flights 0507 and 0616 mainly flew over D1, but were impacted by air
814 masses from D2. Flights 0512 and 0618 mainly flew over D2 and were rarely
815 impacted by air masses from D1. The percentages shown in the titles represent
816 the contributions of emissions from this other sub-region (D1 or D2) to the overall
817 airborne measurements of CH₄ mixing ratios in each flight.

818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848

Figure 1.

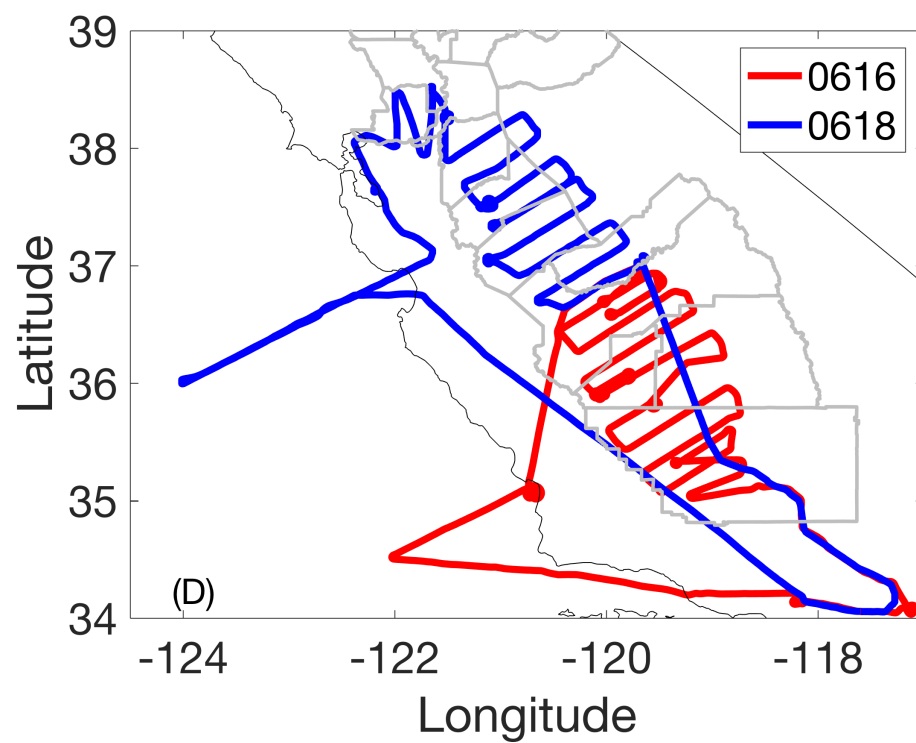
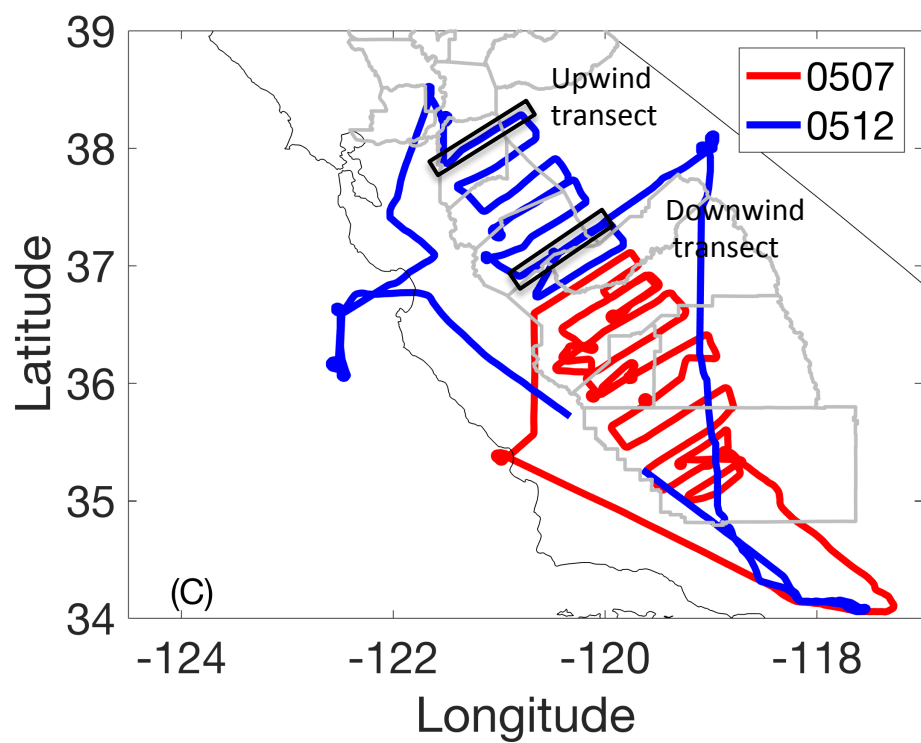
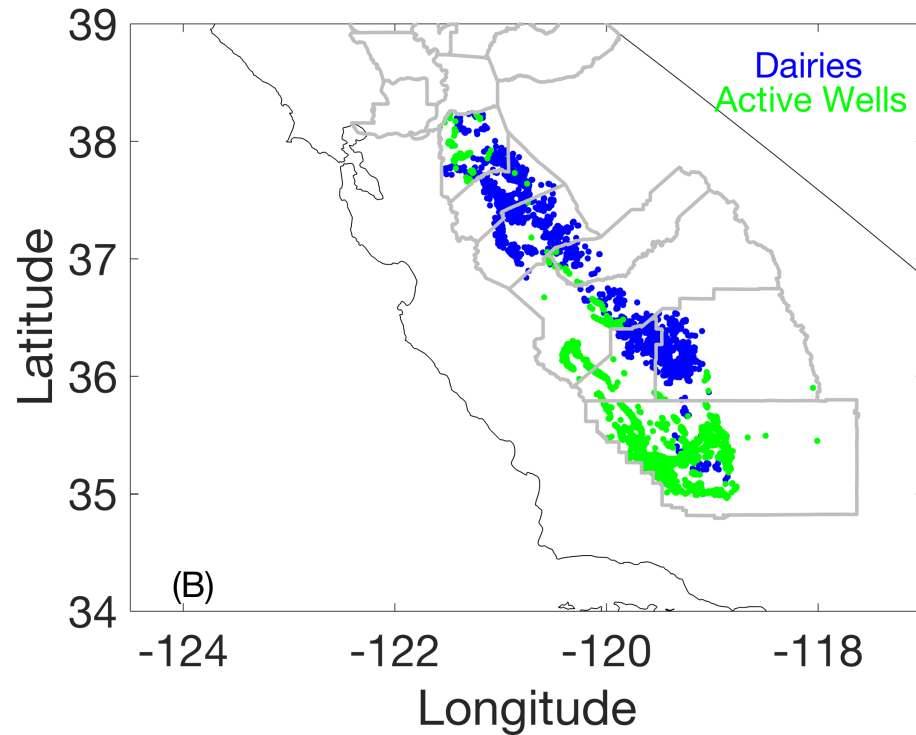
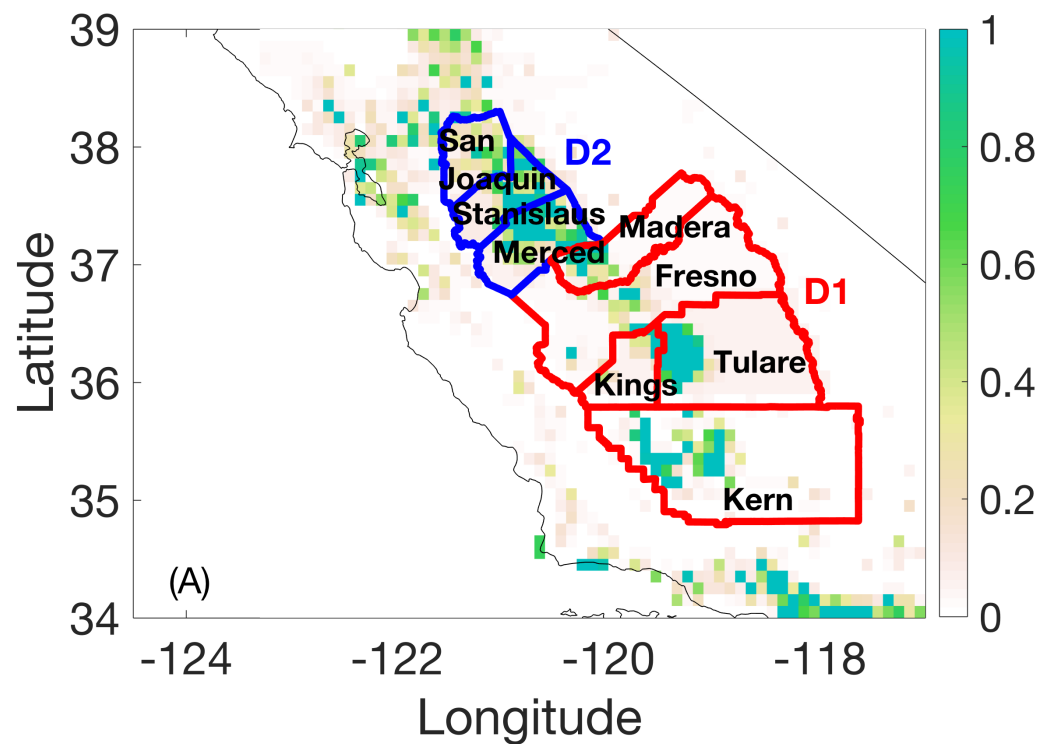
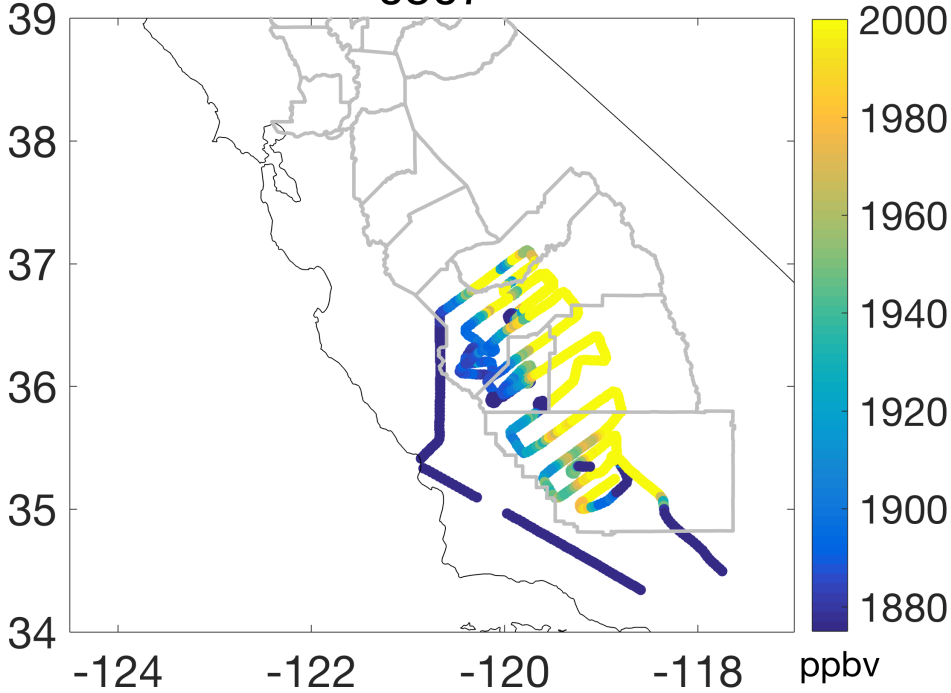
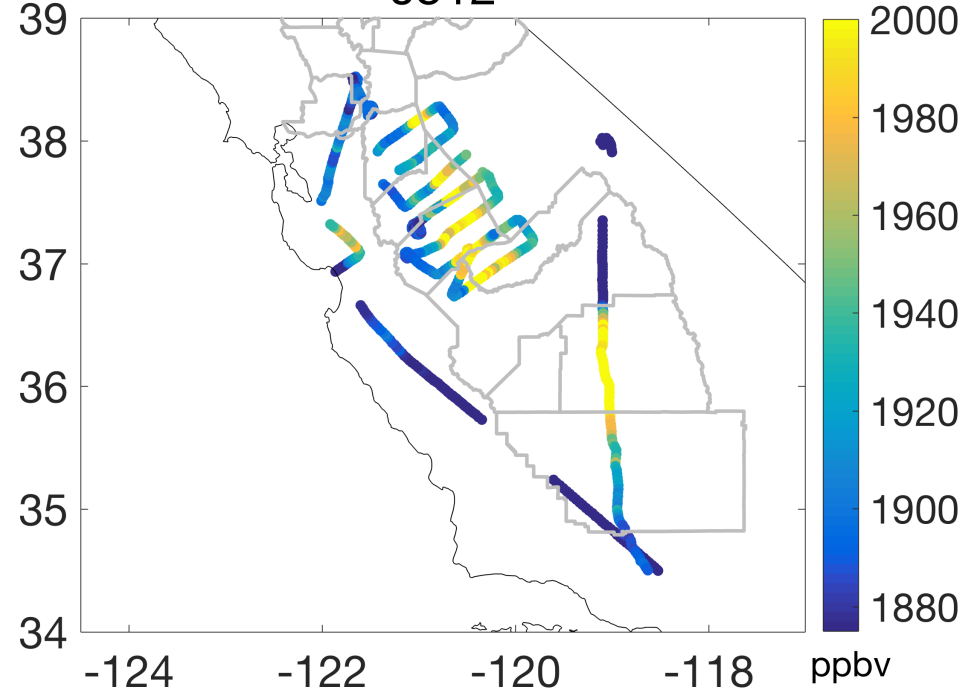


Figure 2.

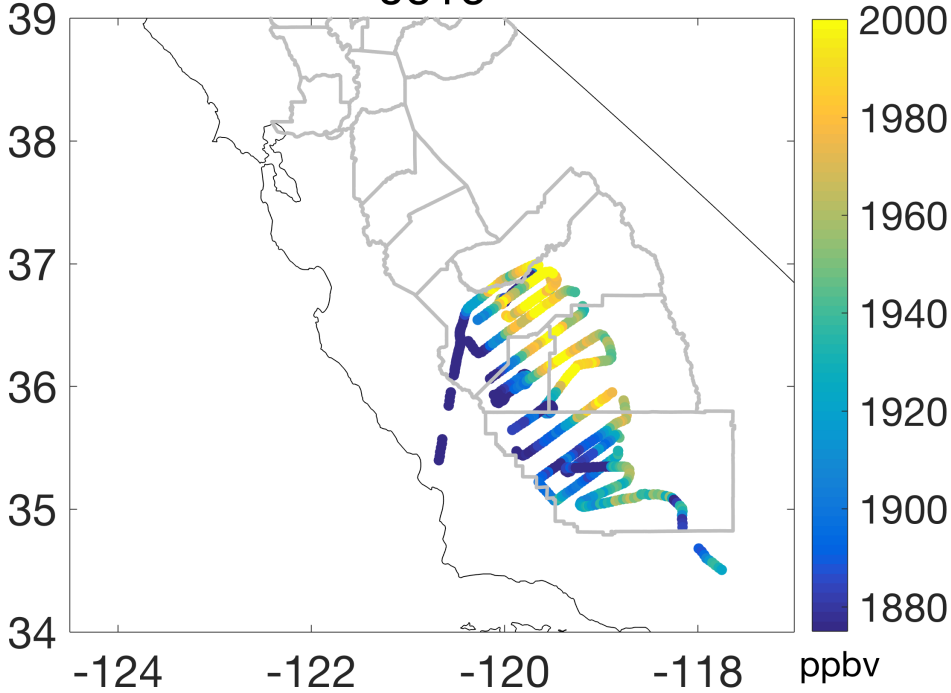
0507



0512



0616



0618

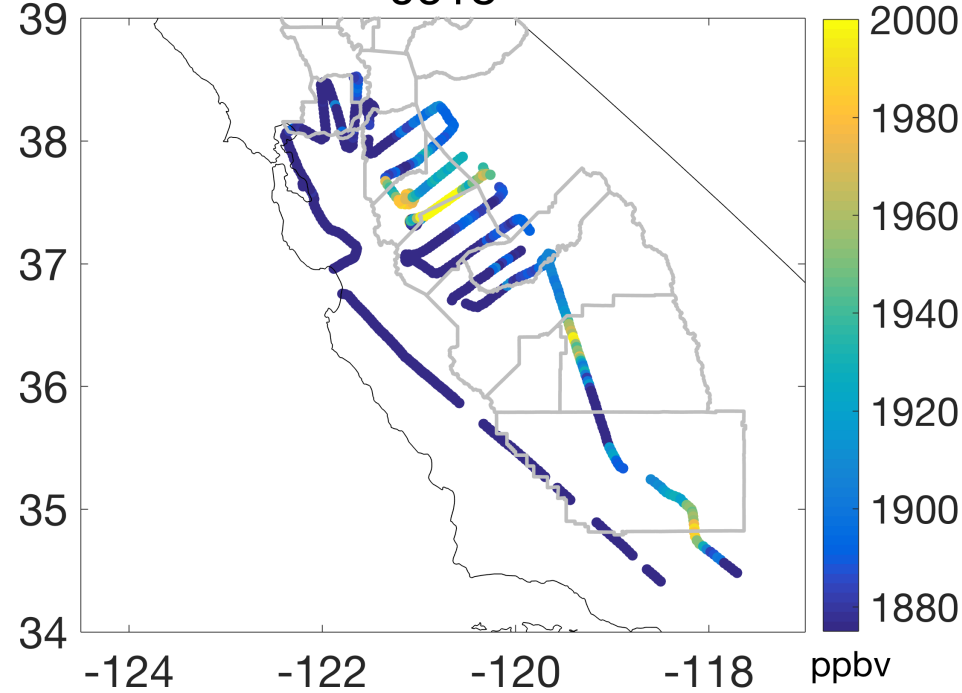


Figure 3.

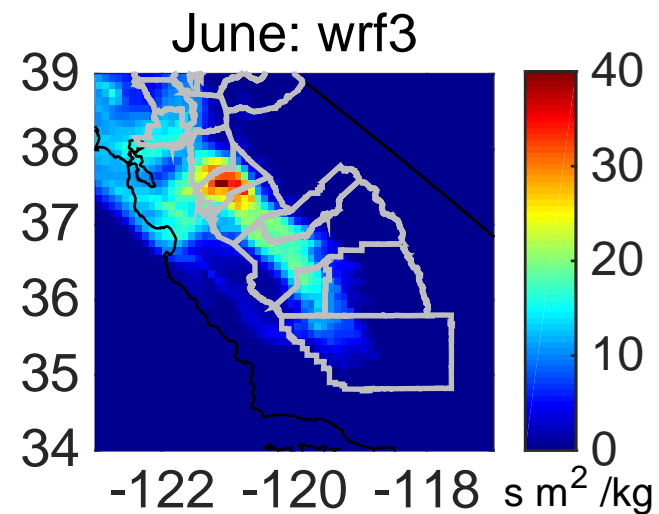
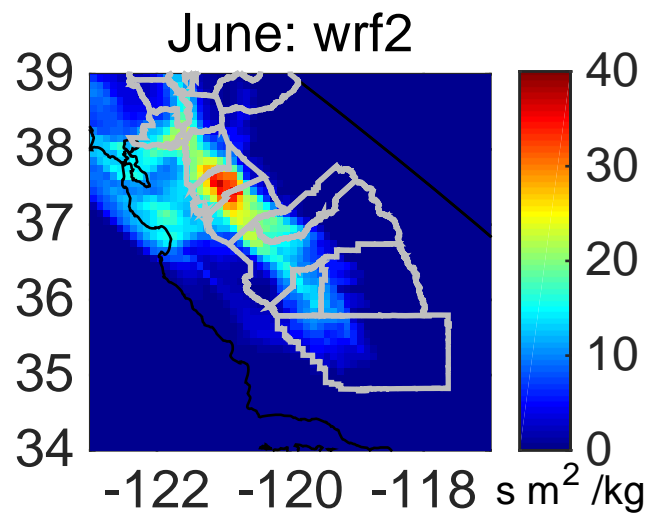
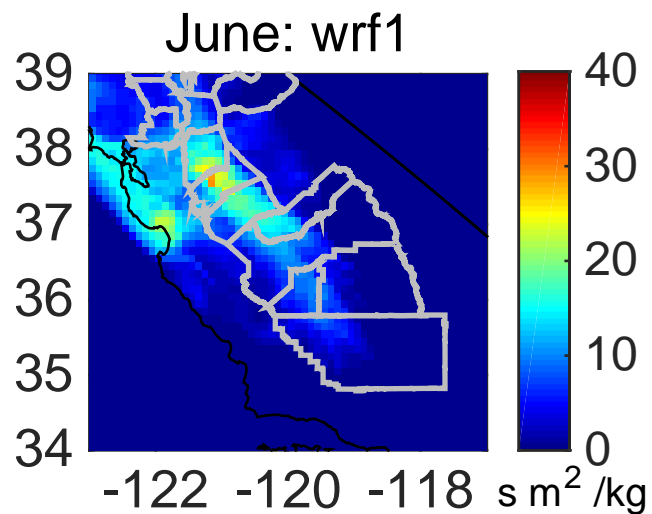
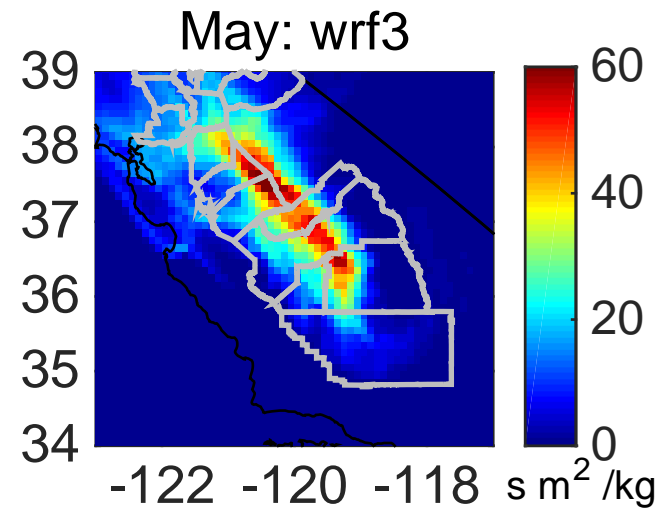
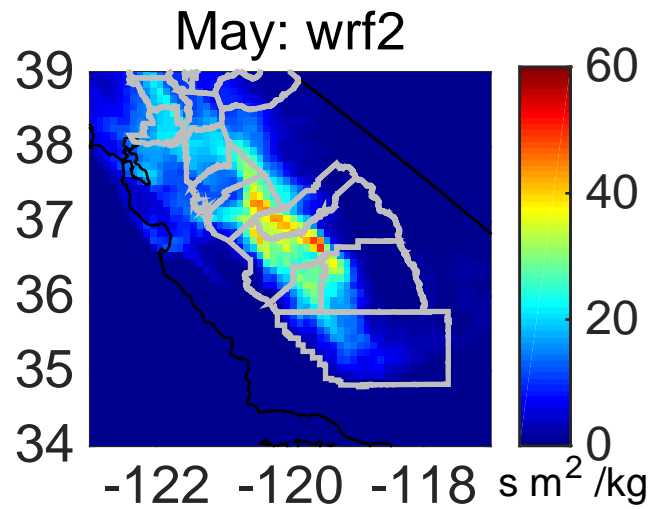
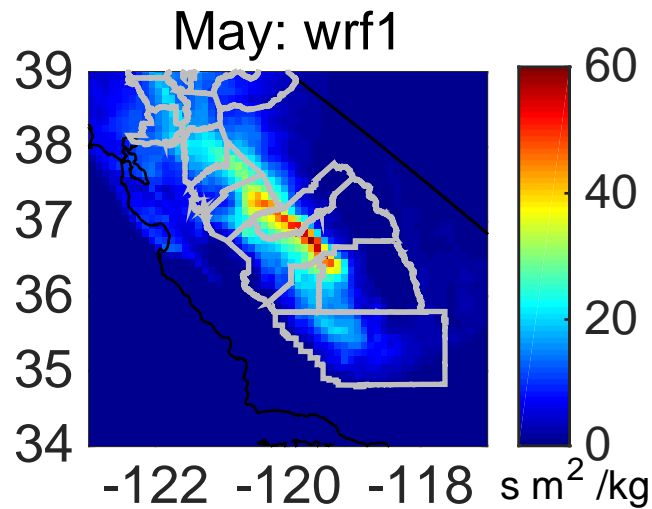
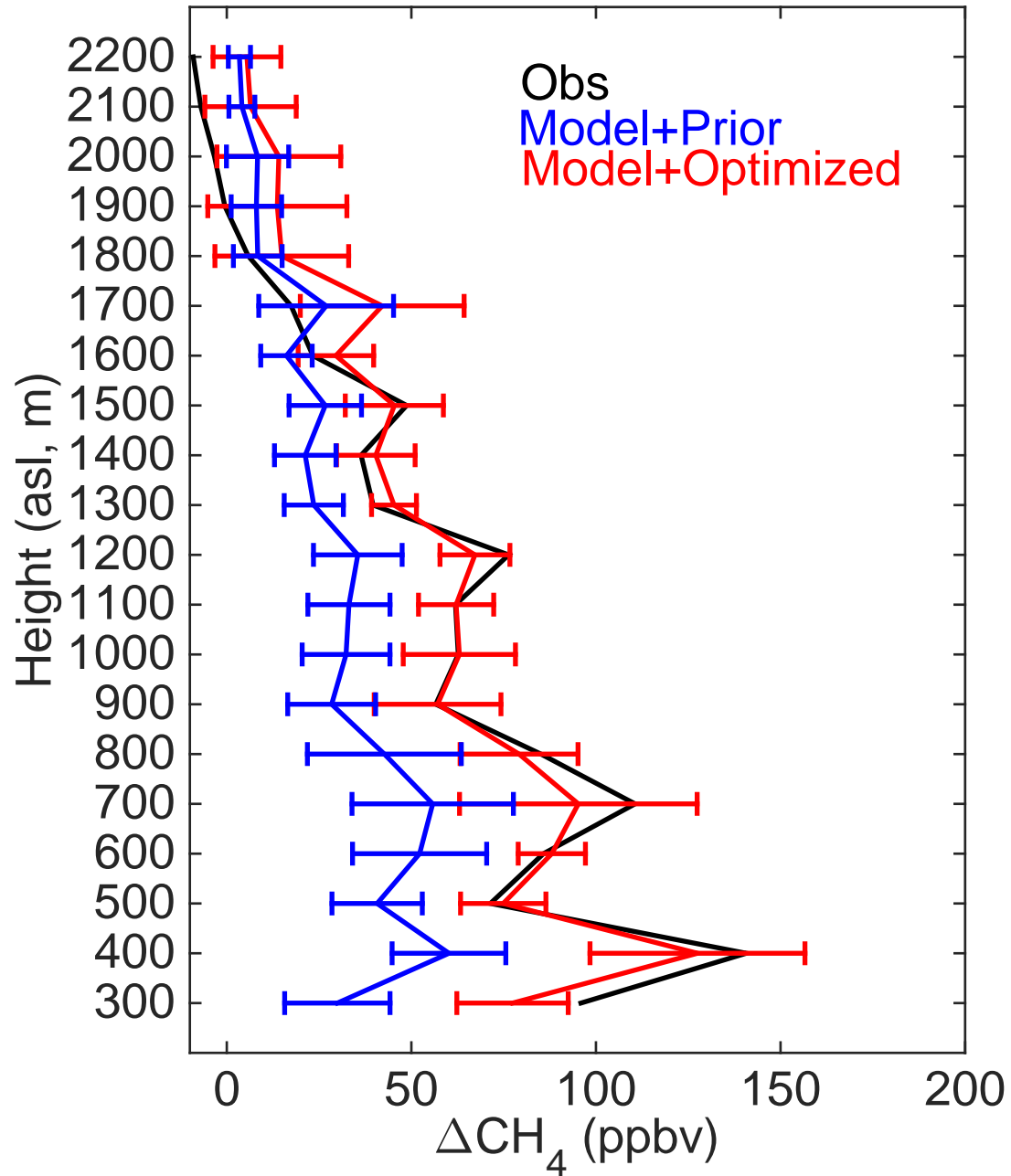


Figure 4.

May



June

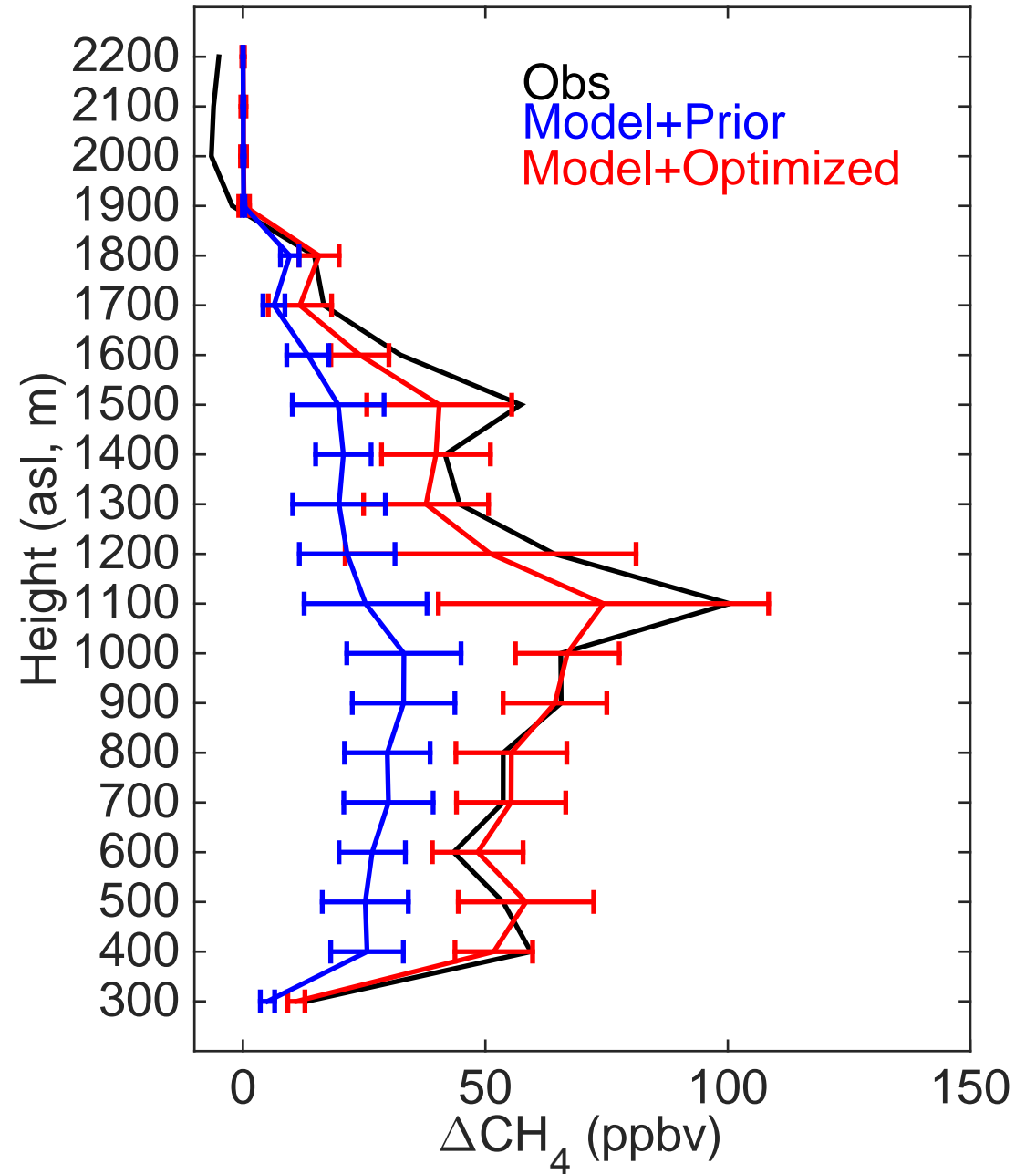


Figure 5.

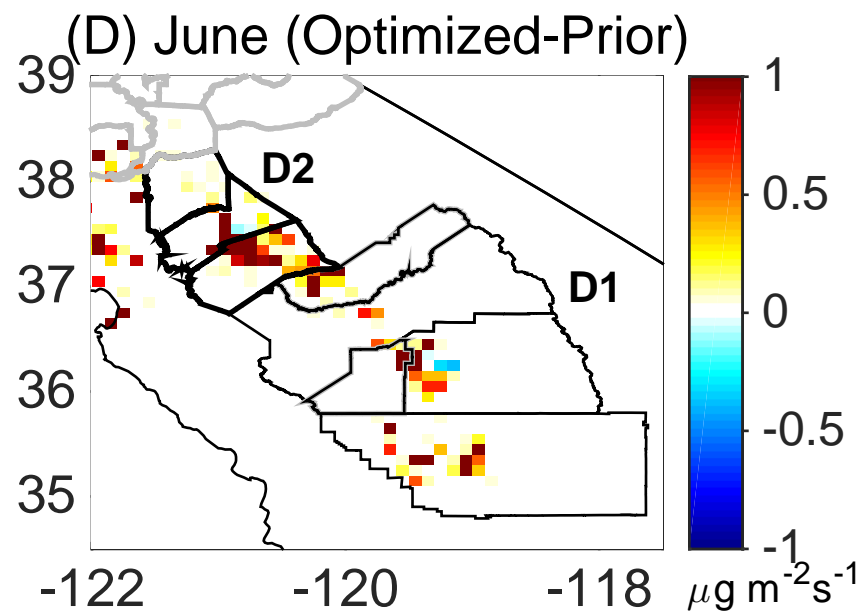
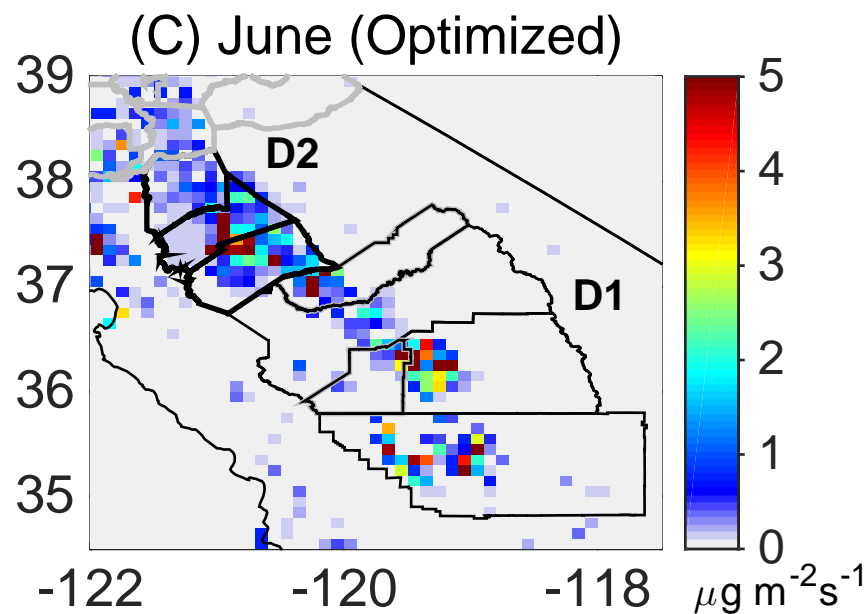
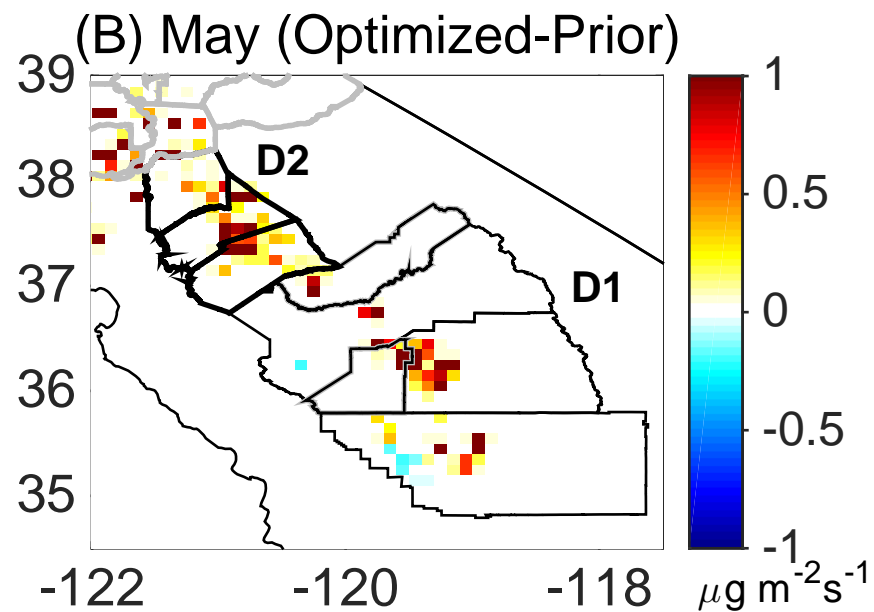
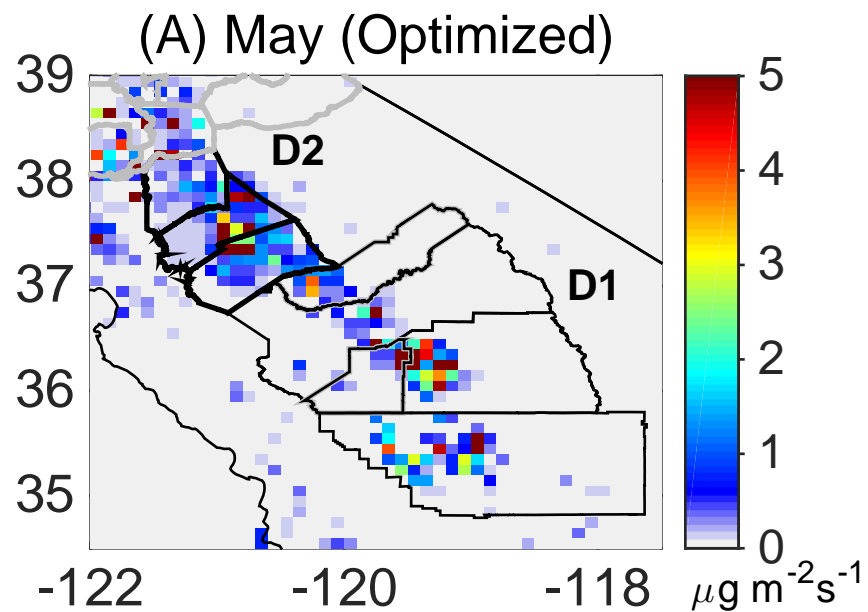


Figure 6.

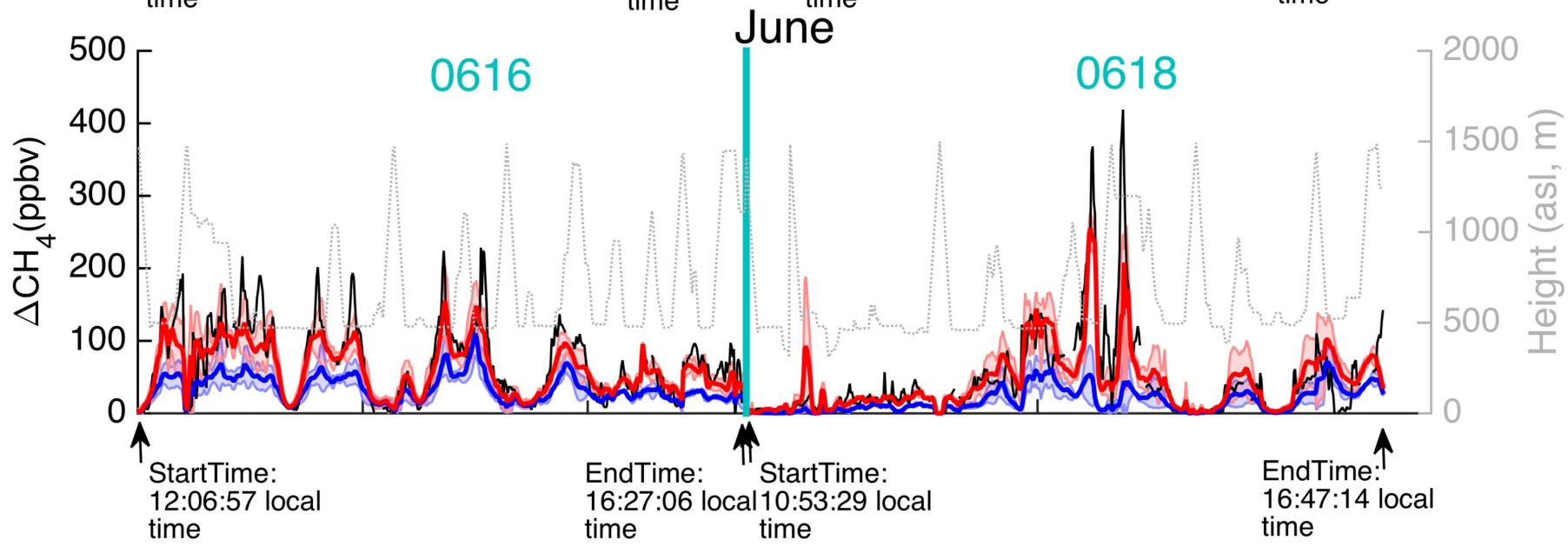
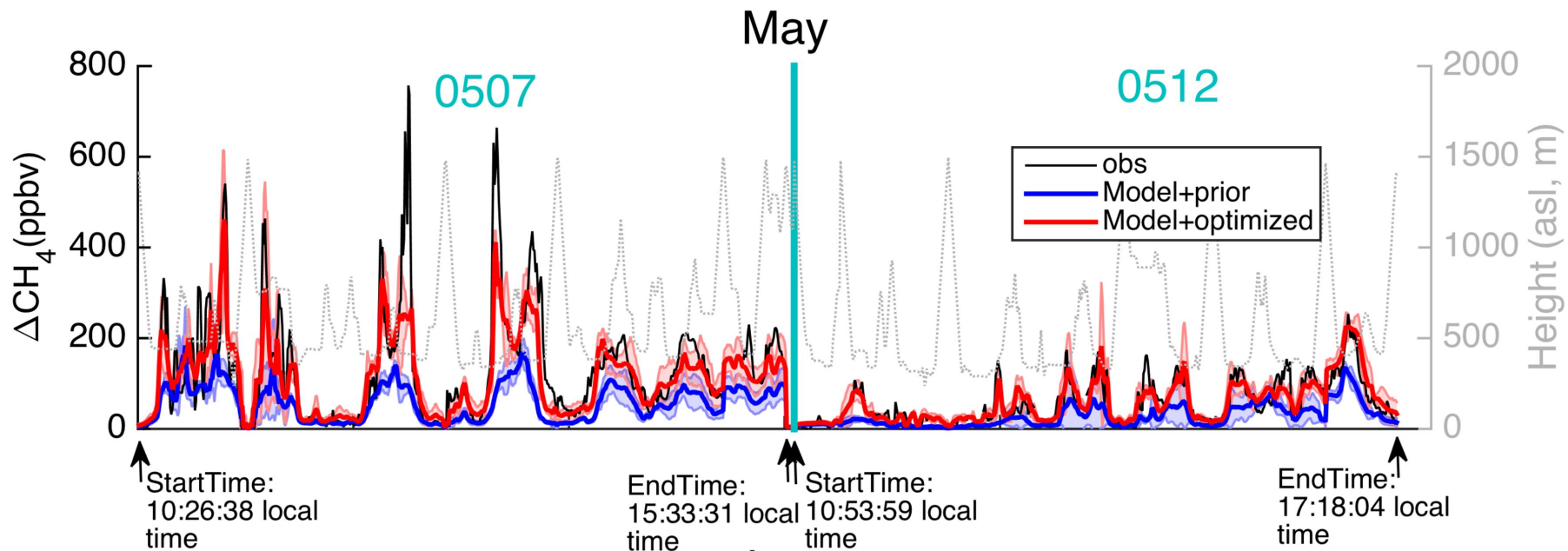


Figure 7.

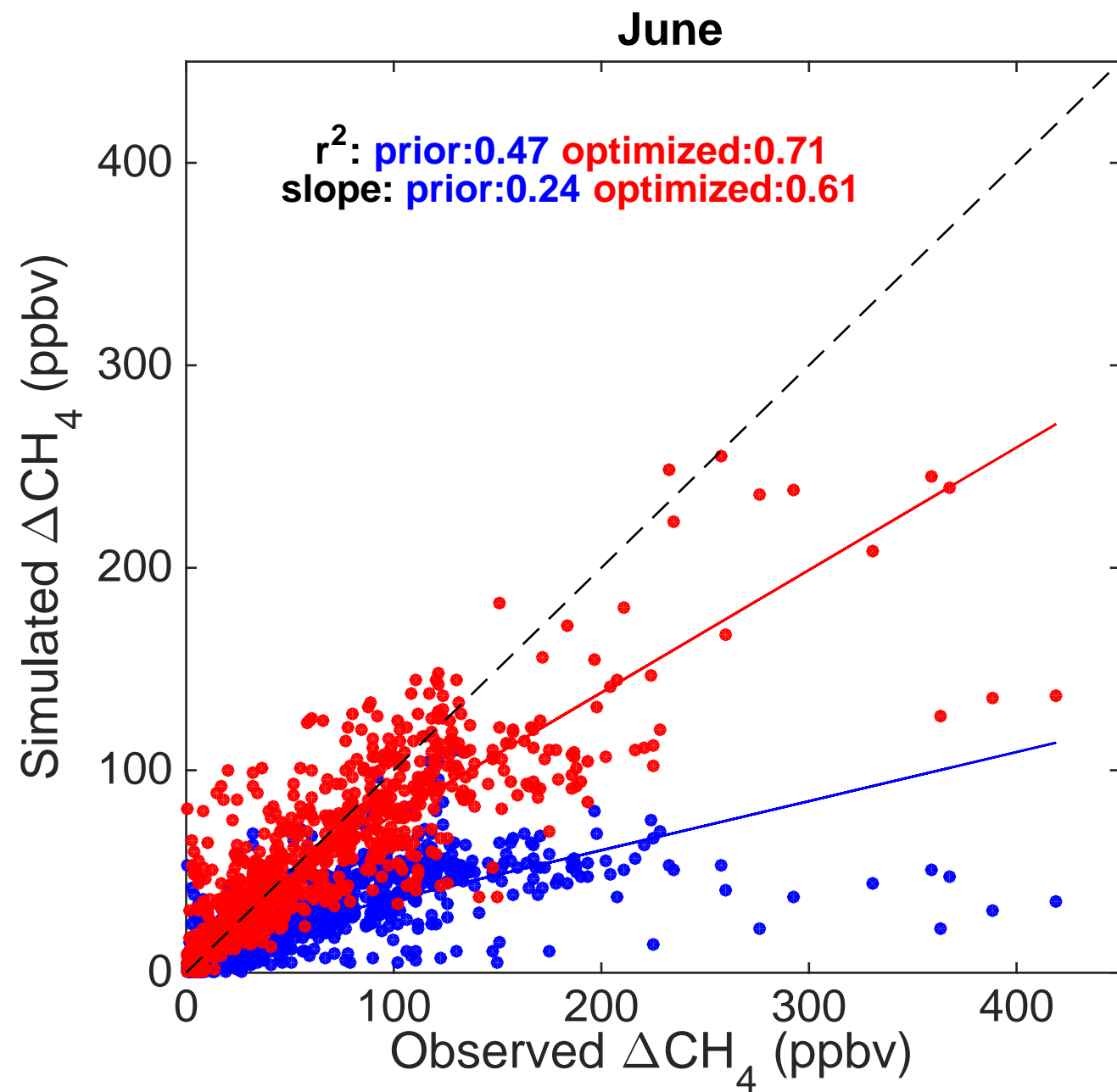
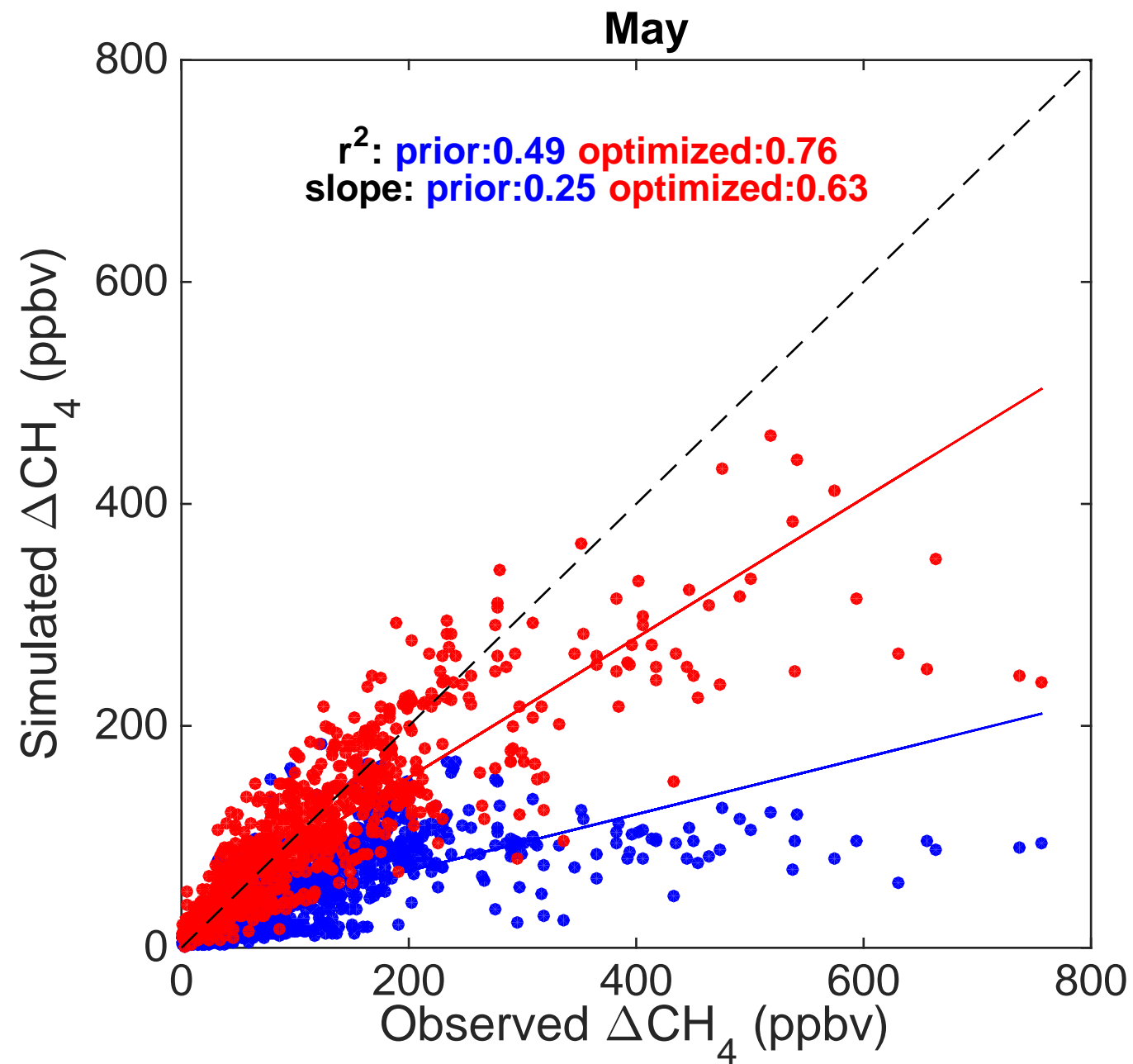
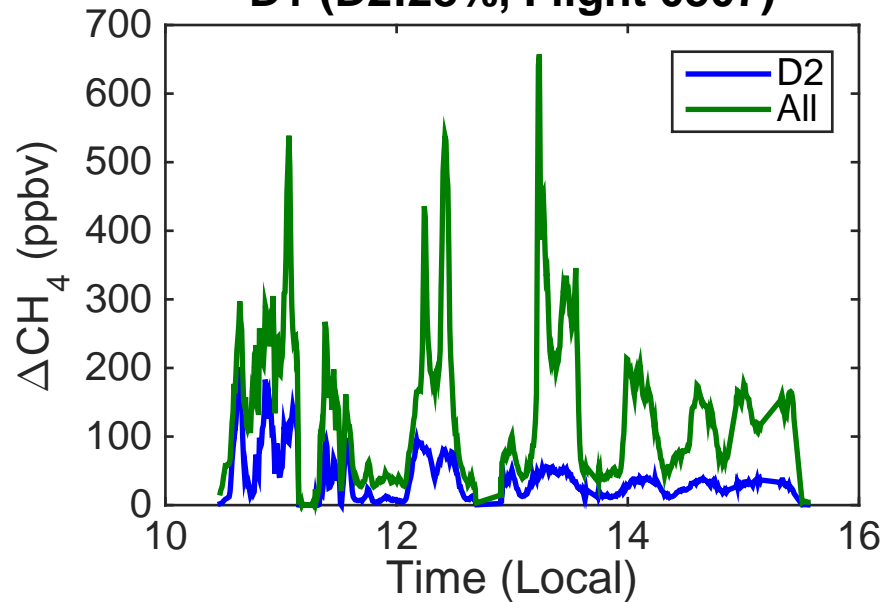
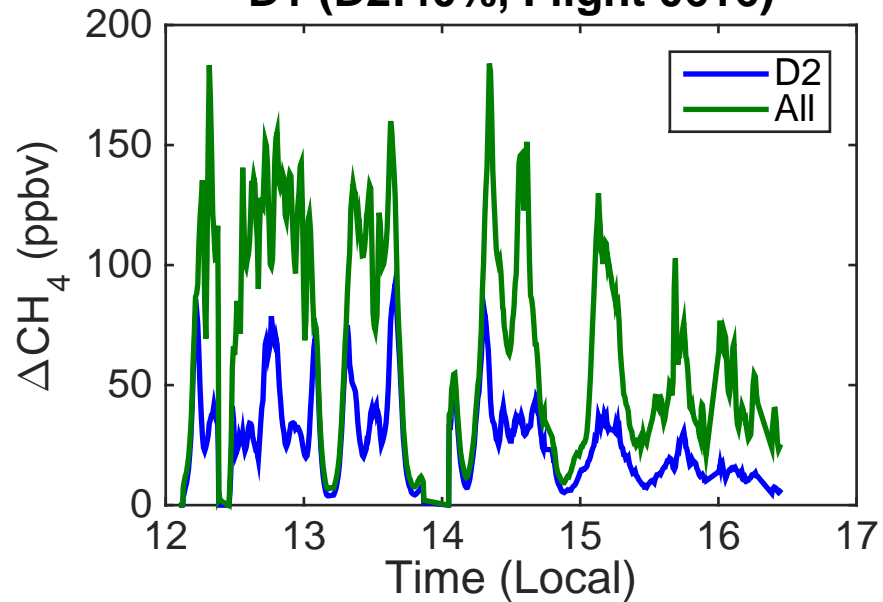


Figure 8.

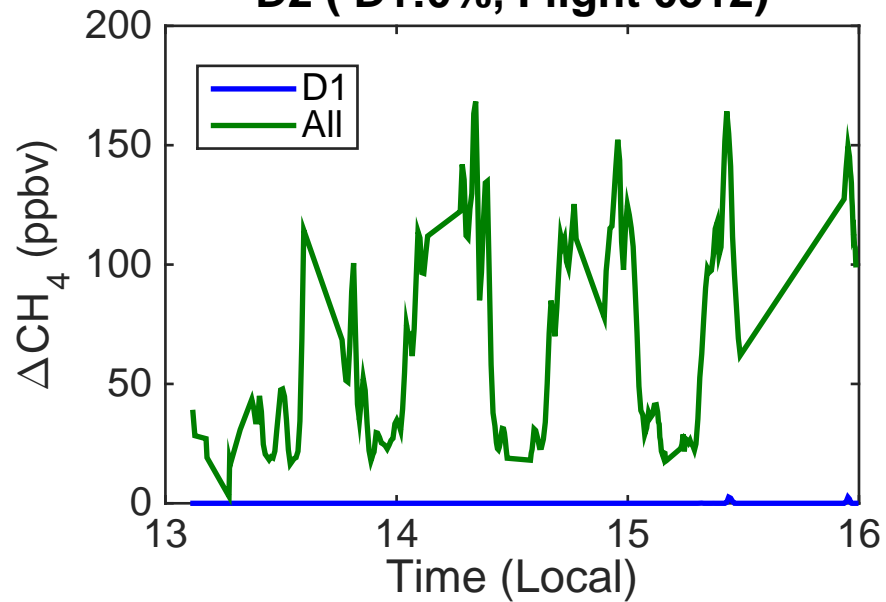
D1 (D2:28%, Flight 0507)



D1 (D2:40%, Flight 0616)



D2 (D1:0%, Flight 0512)



D2 (D1:0%, Flight 0618)

