Supporting Information for:

Methane Emissions from Natural Gas Infrastructure and Use in the Urban Region of Boston, Massachusetts

Kathryn McKain, Adrian Down, Steven M. Raciti, John Budney, Lucy R. Hutyra,
Cody Floerchinger, Scott Herndon, Thomas Nehrkorn, Mark S. Zahniser,
Robert B. Jackson, Nathan Phillips, Steven C. Wofsy

Table of Contents

S1 Measurement Methods 3
S1.1 Methane Measurements 3
S1.2 Measurement Site Considerations 3
S2 Ethane-Methane Ratio Analysis 4
S2.1 Details of Atmospheric Analysis 4
S2.2 Details of Pipeline Analysis 5
S2.3 Robustness of Results to Assumptions 5
S3 Modeling Framework Description 6
S3.1 WRF-STILT Configuration 6
S3.2 Prior Inventories 6
S3.2.1 Natural Gas Consumption 6
S3.2.2 Methane Emissions 7
S3.3 Error Estimation for Optimized Methane Emissions 8
S4 Robustness of Emission Results 9
S4.1 WRF Validation 9
S4.2 Simple Mass Balance Model 9
S4.3 Modeling Framework Variants 10
References 10
Fig. S1 Hourly Average Methane Concentrations 14
Fig. S2 Daily Afternoon Average Methane Concentrations 14
Fig. S3 Methane Instrument Drift 15
Fig. S4 Optimized Methane Emission Results from Individual Measurement Sites 16
Fig. S5 COP Corner-sampling Method 17
Fig. S6 Ethane and Methane Measurements 18
Fig. S7 Pipeline Map 18

Fig. S8 Pipeline Gas Composition 19
Fig. S9 WRF Model Domains 20
Fig. S10 Powerplant Map 20
Fig. S11 Monthly NG Consumption by Sector 21
Fig. S12 Prior Emissions Maps 22
Fig. S13 Observation-Model Comparisons, Scatterplots 23
Fig. S14 Observation-Model Comparisons, Diurnal Plots 25
Fig. S15 WRF Model Biases and Errors, Temperature and Wind Speed, Maps 26
Fig. S16 WRF Model Bias, Temperature and Wind Speed, Diurnal Averages 27
Fig. S17 Optimzed Methane Emissions from Alternative Model Configurations 28
Table S1 Literature Review of Top-Down Studies of Urban Methane Emissions 29
Table S2 Measurement Site Information 30
Table S3 Meathane Instrument Performance Statistics 30
Table S4 Inventory Estimates of Methane Emissions 31

S1 Measurement Methods

S1.1 Methane Measurements

Methane (CH_4) was measured continuously at four sites with four different models of Picarro cavity ring down spectrometers (1, Table S2), at intervals of 2-7 seconds, depending on the instrument model. All models measured the same spectroscopic feature of CH_4. Carbon dioxide (CO_2) and water vapor (H_2O) (and carbon monoxide at COP only) were also measured at each site. A small fraction of CH_4 data points were removed when the instrument's optical cell pressure and temperature were outside 0.1 torr and 0.005 ˚C of their set-points to ensure compatibility with the instrument's spectroscopic fit parameters. Sample gas streams were not dried, so empirical H_2O correction factors (2) were used to convert measured to dry molar fractions of CH_4. Instrument-specific H_2O correction factors were derived for all instruments except for the BU instrument, for which correction factors from the literature (2) were applied.

A two-point linear calibration equation was calculated for each instrument in the field. Compressed air cylinders with known, approximately ambient CH_4 concentrations ("surveillance standards") were measured by each instrument for four minutes every eight hours (Fig. S3) to quantify the intercept (null value) of the calibration equation and to track long-term drift (Table S3). Data from the first ~90 seconds of each surveillance measurement period were disregarded to ensure the surveillance measurement had equilibrated. Measured offsets were smoothed over a multi-day moving window before being applied to correct ambient CH_4 measurements (Fig. S3).

All calibration and surveillance standards were manufactured by Scott-Marin (Riverside, CA) using natural air (3) and tied to scales defined by the National Oceanic and Atmospheric Administration (NOAA; 4) and the World Meteorological Organization (5). Surveillance standards were calibrated against NOAA primary standards before and after field deployment and were all found to have changed ≤ 0.2 ppb for CH_4 (Table S3). One surveillance tank was used at each site for the entire year of measurements.

Total analytical uncertainty is approximated as the sum of measurement precision, uncertainty in calibration and surveillance tank values (Table S3), and uncertainty in the H_2O correction. Long-term drift was not included in the calculation of total analytical uncertainty because it was captured and corrected for in the data processing. Uncertainty in the H_2O correction equation is estimated as ±2 ppb at H_2O concentrations up to 3.4% (2), which sufficiently captures the maximum ambient measured H_2O concentration. Therefore total analytical uncertainty for one year of hourly average CH_4 measurements among the four sites was ≤ ~3 ppb (95% CI), < 0.2% of ambient concentrations.

S1.2 Measurement Site Considerations

Atmospheric measurements made in urban environments are vulnerable to disproportionate influence from nearby sources because large sources may occur at high densities in such environments. Furthermore, for practical reasons, urban measurements are often made from rooftops, which may not be optimally positioned to sample free-stream flow (6). The two urban measurement sites in this study (BU and COP) were located on the tops of buildings, so special precautions were taken to ensure that the data were not contaminated with signals from very-near sources, in particular rooftop building vents. The finding that the two urban measurement sites yielded annual average emission estimates for the region that were not significantly different (Fig. S4) is a strong indicator that the average enhancements measured at both sites were not dominated by signals from very near sources.

The BU measurement was made from a 2-meter tower mounted in the center of the flat rooftop of a 6-story building. The building is located in a neighborhood comprised of both shorter and taller buildings,

and a few small bathroom vents are located on the same roof. The measured C_2H_6/CH_4 in one of these vents was similar to that of pipeline gas, suggesting that the excess CH_4 in the vent derived largely from NG leaks into the sewer system. Signals from the sewer vents on the BU roof were characterized by very short periods (1-3 s) of clearly elevated concentrations ("spikes"). During the afternoon hours used in the analysis, these spikes accounted for < 0.2% of the BU observations. These signals were eliminated from the BU CH_4 data by trimming the lower and upper (to ensure no bias was imposed) 5% of the data in each hour prior to calculating hourly averages.

The COP site was located on a much taller building that stands well above surrounding structures, so contamination from nearby building sources was less of a concern. However, the building itself has many large vents on its rooftop for bathrooms, air heating and cooling systems, and a restaurant. Airflow over a building leads to the formation of small-scale turbulence and a zone of low pressure at the top of the building, which can entrain or aspirate contaminated air emitted from the building (6).

To avoid sampling air emitted from building vents at COP, four sample inlets were placed at the corners of the building, two stories below the top of the building, and each corner was sampled sequentially for five minutes each. The concept of the COP corner-sampling method is that, at any given time, at least one corner will represent uncontaminated, upwind conditions. The upwind corner(s) was selected as that with the lowest average concentration of CH_4, CO_2, or CO in each 20-minute, 4-corner sequence (Fig. S5). Equal lengths of tubing were used for the four sample lines to ensure equal delay times among them and a bypass pump continuously flushed all four sample lines.

Harvard Forest measurements were made from a tower in a mixed-deciduous forest in Petersham, Massachusetts (7). Methane was measured sequentially at eight heights (0.3, 0.8, 4.5, 7.5, 12.7, 18.3, 24.1, and 29 m) on the tower, for 4 minutes each and for 8 minutes at the highest (29 m) level. Methane concentrations were typically slightly depleted at the lowest measurement heights due to an oxidative soil sink (8). Only measurements from 29 m were used in this study.

To test whether the non-continuous sampling at COP and HF reduced the representativeness of the datasets, we simulated the COP sub-sampling routine with the continuous BU dataset by randomly sub-selecting one 5-minute period in each 20-minute interval. The hourly averages generated from the sub-selected dataset were not significantly different than hourly averages generated from the full dataset. Total data coverage for the year at the four sites was > 94%, with gaps due to maintenance, power outages, etc.

S2 Ethane-Methane Ratio Analysis

S2.1 Details of Atmospheric Analysis

To quantify the relationship between the atmospheric C_2H_6 and CH_4 measurements, we used χ^2 minimization (9, equation given below) of a straight-line fit (b=slope, a=intercept) to 5-minute medians of 1 hertz data points (Fig. 4, x and y), with errors in each variable at each time point (i=1:N) characterized by the standard error of the mean (σ).

$$\chi^2(a,b) = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_{yi}^2 + b^2\sigma_{xi}^2}$$

For the 2012-13 period, data from afternoon (11-17 h EST) hours only were used, but for the 2014 period, data from all hours were used because the C_2H_6 signal-to-noise ratio was smaller (Fig. S6). Quantification

of covariance (slope of the regression line) between C_2H_6 and CH_4 measurements from each day using short, 5-minute median, intervals was adopted to eliminate any potential influence of building emissions (section 1.2) and because C_2H_6 measurements were not available at the background stations. If background C_2H_6 measurements had been available, then quantification of the atmospheric C_2H_6 to CH_4 covariance using longer intervals, and optimization of C_2H_6 emissions in an inversion framework, may have been possible. The current framework assumes that background concentrations did not vary substantially during individual days, supported by the tight correlation ($R^2 > 0.75$) between observed C_2H_6 and CH_4 on approximately half of the days. Days with large shifts in wind direction sometimes did not have consistent C_2H_6/CH_4 and were rejected by the $R^2 > 0.75$ criterion. We used daily fits to 5-minute intervals to calculate the average ratio in order to limit the influence of possible autocorrelation between points from individual hours or days. 95% confidence intervals were calculated by bootstrapping daily fit slopes with 1,000 iterations and sampling with replacement.

S2.2 Details of Pipeline Analysis

Three major pipelines, Tennessee (TGP), Algonquin (ALG), and Maritimes and Northeast (MNE), and a liquefied natural gas (LNG) import terminal, supply NG to the Boston region (Fig. S7). Hourly gas quality data were collected from the informational postings for each pipeline (10-11). Daily median C_2H_6 and CH_4 ratios were calculated for each pipeline using hourly data (Fig. S8) from the gas quality measurement stations closest to Boston (Fig. S7). The three pipelines delivered the following fractions of NG consumed in Massachusetts in 2012: 65% TGP, 30% ALG, and 5% MNE (12). Seasonal average pipeline C_2H_6 and CH_4 ratios were calculated for the same time period as the atmospheric ratios using the daily median pipeline ratios, weighted by the fractional contributions of each pipeline (Fig. S8). 95% confidence intervals were estimated by bootstrapping daily weighted median ratios with1,000 iterations and sampling with replacement.

Gas composition in the pipelines is measured using industry standard methods (13-14), but uncertainties due to sampling and measurement error are not reported. Additionally, the representativeness of the measured relative to the lost gas is not known. The approach described above to estimate the mean pipeline C_2H_6 and CH_4 ratio is intended to yield an aggregate estimate that is robust to sporadic erroneous and/or unrepresentative measurements.

S2.3 Robustness of Results to Assumptions

Central C_2H_6/CH_4 values calculated from both the atmospheric and pipeline data were insensitive to the underlying assumptions. Slopes from χ^2 optimization (section 2.1) were 10-15% larger than those obtained from an ordinary least squared regression because variance in the CH_4 measurements cannot be neglected. The atmospheric ratios reported in the main text were based on daily fits to data with $R^2 > 0.75$ and from afternoon hours only (for the 2012-13 period only). The same calculation using five-minute means instead of medians, a less stringent R^2 cutoff, and data from all hours instead of afternoon hours yielded both larger and smaller ratios that varied by < 10% and were not significantly different from those reported in the main text.

We believe the relative contribution of NG to CH_4 emissions observed in the late spring of 2014 can be applied to the spring and summer months of the year prior because the average air temperature and observed ΔCH_4 in May-June were very similar between 2013 and 2014 and because the observation interval approximates the midpoint of the March-August period for which we adopted the observed NG fraction.

The pipeline C_2H_6/CH_4 reported in the main text was based on a weighted average between the three pipeline companies with contributions of 65% for TGP, 35% for ALG, and 5% for MNE. However, these fractional contributions are only valid for the entire state of Massachusetts in 2012, and we have no information about the relative sensitivity of the atmospheric measurements to the three pipelines and individual gas quality measurement locations. Therefore, we also calculated the pipeline ratio using equal contributions from TGP and ALG, and no contribution from MNE. The result was not significantly different than the ratio given in the main text. The high correlation between atmospheric C_2H_6 and CH_4, and the close correspondence of the atmospheric and pipeline gas ratios, support the hypothesis that NG was the major source of enhancements for both gases.

S3 Modeling Framework Description

S3.1 WRF-STILT Configuration

The Stochastic Time-Inverted Lagrangian Transport (STILT v656, 15) model was driven with customized meteorological fields from the advanced research version of the Weather Research and Forecasting model (WRF v3.4.1, 16-17). Meteorological fields were generated at four gridded horizontal resolutions (1, 3, 9, and 27 km) in a two-way nested arrangement centered on Boston (Fig. S9). All WRF domains had 42 vertical levels. Initial and lateral boundary conditions were provided by the North American Regional Reanalysis (18). Overlapping 30-hour forecasts were initialized every 24 hours, at 00 UTC, and the first 6 hours of each forecast were discarded to allow for spinup. Grid nudging was used in the outer-most domain only and not within the planetary boundary layer (PBL). Additional details of the WRF configuration used in this study are given in Nehrkorn et al. (19, case "Turb-U"). Section 4.1 describes comparisons between WRF-simulated and measured meteorological parameters.

STILT was run in time-reverse mode in which an ensemble of 500 particles was released every hour from each of the urban measurement sites and transported backward in time for 10 days according to the WRF meteorology. The majority of particles reached the study boundary (Fig. 1) in < 8 hours and the median travel time was < 3 hours. Background values generated from NHT measurements were assigned to particles that exited the coastal portion (at 20-140°) of circular boundary (~22% of particles) and background values from HF were assigned to all other particles.

S3.2 Prior Flux Fields

S3.2.1 Natural Gas Consumption

In order to understand emissions results as a fractional loss rate of NG delivered within the modeling framework, it was necessary to create a spatially explicit map of NG consumption (Fig. 6). Consumption is an appropriate estimator of net gas flows through the study area because all of the pipelines entering the study area terminate inside or very near the study area boundary (Fig. S7). Reports of NG consumption by state, month, and sector were obtained from the Energy Information Administration (EIA, 20) for the study area and time period. Volumes of NG were converted to masses of CH_4 using the ideal gas law by assuming industry standard temperature and pressure (60 ˚F and 1 atm) and 97% CH_4 content, based on the gas quality measurements (section 2.2, Fig. S7), giving 1 scf NG = 1.16 moles CH_4= 18.6 grams CH_4.

Natural gas consumption in electric power, residential, commercial, industrial, vehicle fuel, and pipeline and distribution use sectors (20), were all included in the consumption estimate because we were not able to distinguish emissions from individual sectors with the atmospheric data. Monthly and state consumption by the entire electric power sector and portions of the industrial and commercial sectors (17% and 6%, respectively) was spatially allocated to individual power production facilities (21, Fig.

S10). A small number of plants had missing values in 2013, which were filled according to each plant's relative consumption of the state, sector, and monthly total in 2012. Filled values accounted for 1.5% of the total consumption by power production facilities.

Consumption by the residential and commercial sectors (excluding the portion accounted for by individual electric power production facilities) was spatially allocated using a parcel-level database for Massachusetts of residential and commercial building square footage and the fuel type used in each building for space and water heating (22). This dataset was constructed from multiple state and local government data sources such as Registry of Deeds, Land Court data, Town Clerk data and tax assessor information. Vehicle fuel consumption was spatially allocated to compressed natural gas (CNG) and LNG fueling station locations (23). Industrial sector consumption (excluding the portion accounted for by individual power production facilities) and pipeline and distribution use was spatially allocated using commercial building square footage data.

For the study area outside of Massachusetts, census data on the number of housing units with NG at the blockgroup scale (24) were used in place of the building square footage database, which was only available for Massachusetts. Within Massachusetts, the R^2 value for the 1 km^2 gridded census and square footage datasets was 0.8 for residential buildings and 0.7 for residential and commercial buildings together, demonstrating that the census dataset was a reasonable substitute for the square footage dataset where the former was not available. Figure S11 gives monthly average NG consumption by sector for Massachusetts and for the 90-km radius study area, as spatially allocated using the methods described above.

Uncertainty estimates for state monthly NG consumption are supplied by EIA (25). Monthly standard errors for total consumption in Massachusetts are available for 6 of the 12 months in the study period, with NAs reported for the remaining months, and range from ± 0.4 to 1.1 %. Using the largest monthly standard error value, multiplying by 1.96 to estimate the 95% confidence interval, and summing them in quadrature, leads to 95% confidence interval estimates of ± 3.7 % for three-month seasonal totals and ± 7.4 % for the annual total. We did not estimate uncertainties in the spatial allocation of state total consumption to the study area, or uncertainties in spatial distribution within the study area, because no independent dataset is available for comparison and because it is unknown how well the spatial distribution of consumption approximated the spatial distribution of emissions.

S3.2.2 Methane Emissions

Two CH_4 emission priors, EDGAR v4.2 FT2010 (26) and one created for this study (Table S4), were tested as inputs to the modeling framework. EDGAR is a global product that uses simplified methods (e.g. scaling by population density) to spatially disaggregate emissions. To take advantage of locally available data and knowledge, a customized emission prior with five anthropogenic and biogenic source categories (described below) and 1 km^2 spatial resolution was developed for the study domain (Fig. S12). The custom prior was not meant to be exhaustive, but rather was created to provide detailed emission estimates for key sectors with improved spatial resolution and accuracy. Emission results using the customized prior are given in the main text. Emission results using EDGAR are discussed in section 4.3 and shown in Fig. S18. Both priors were adopted as temporally invariant.

Wetlands. Data on wetlands' location, size, and type were obtained from the National Wetlands Inventory (27) for the four states in the study region. These wetland inventories are based on aerial photography and have a mapping unit of 0.4 – 1.2 hectares. Average CH_4 emission rates for each type were taken from Bridgham et al. (28), which calculated mean emission rates from > 100 studies. An average emission rate of 7.6 g CH_4 m^{-2} yr^{-1} was applied to the freshwater wetlands (emergent and forested/shrub) and an emission rate of 1.3 g CH_4 m^{-2} yr^{-1} was applied to saltwater wetlands (estuarine and

marine). Areas of open water (rivers, lakes, deepwater marine) were not included in the wetlands emission layer due to a lack of emissions data from these areas. The total wetland area in the study domain of the prior was 1,900 km^2, ~11% of the land area.

Enteric Fermentation. Methane emissions from ruminant livestock were spatially allocated to counties according to county-level headcounts of cattle and calves from the USDA 2007 Agricultural Census (29). An emission factor of 117 kg CH_4 $head^{-1}$ yr^{-1}, the Environmental Protection Agency (EPA) emission factor for mature dairy cattle in the North Atlantic Region (30), was multiplied by the cattle count in each county to yield a total average emission rate. Emissions from animals other than cattle were not included because cattle accounted for the majority of livestock emissions in the study area.

Transportation. Methane emissions from transportation were estimated using per-mile emission factors by vehicle type and model year (31-32), state-level data on vehicle fleet composition (33), and a database of vehicle miles traveled per road (34). Gately et al. (34) provide a complete methodological description for an analogous emissions model for CO_2.

Point-sources. Annual facility-level data reported to the EPA GHG Reporting Program (35) were used to represent CH_4 emissions from the largest point sources, including landfills, waste combustion, and waste water treatment plants. Data from 2012 and 2013 were weighted according to the study time period (25% in 2012 and 75% in 2013). Emissions from NG distribution companies were not included because they were not considered point sources and were conceptually accounted for in the NG losses layer (described next).

Natural Gas Losses. Methane emissions from NG losses were spatially allocated according to the areas of residential and commercial NG consumption (section 3.2.1) and scaled so total emissions from NG in Massachusetts was equal to the state inventory estimate for 2012 (36). The prior flux fields were not used to determine the fractional contribution of the NG, nor any other, source sector, to total CH_4 emissions because the C_2H_6 measurements provided definitive attribution of CH_4 emissions from NG. Emissions results from scaling NG losses to equal 80% of total emissions, as informed by the C_2H_6 measurements, are discussed in section 4.3 and shown in Fig. S18.

S3.3 Error Estimation for Optimized Methane Emissions

Data points with model-data residuals > 3-σ of the residual distribution (< 5% of points for any individual site and season) were excluded from the emission scaling factor calculations (Fig. S13). Exclusion of outliers had no impact on the mean emission result, but led to slightly smaller confidence intervals (Fig. S18, pt. 0 vs. pt. 1) and larger R^2 values for the optimized data-model fits (Fig. S13).

Means and 95% confidence intervals on all reported estimates of ΔCH_4 and of CH_4 fluxes were generated through an end-to-end bootstrap analysis with the following steps. Distributions of possible background CH_4 concentrations at the two exterior sites (HF and NHT) were generated at 1-percentile increments between the 5^{th} and 35^{th} percentiles, over a 48-hour moving window. The lower percentile and moving window approach was employed to capture synoptic-scale variability and because near-surface nighttime observations are often affected by small nearby sources due to stratification of the nocturnal boundary layer which traps emissions near the ground. In the bootstrap, background distributions were randomly sampled each day to calculate $\Delta CH_{4,obs}$. Average hourly afternoon $CH_{4,obs}$ and $\Delta CH_{4,mod}$ values were randomly sampled separately for each day to generate average daily values of observed and modeled ΔCH_4. Daily average ΔCH_4 values were randomly sampled to generate seasonal average ΔCH_4 and inventory scaling factors to derive average CH_4 emissions. Lastly, seasonal average ΔCH_4 and CH_4 emissions were randomly sampled to generate annual averages of each. Each of these steps was

performed 1,000 times and means and confidence intervals were calculated from the resulting distributions.

S4 Robustness of Emission Results

S4.1 WRF Validation

WRF-simulated meteorological fields were compared against available meteorological measurements at National Weather Service observing sites (37) using the Model Evaluation Tools verification software (38) from the National Center for Atmospheric Research. Figure S15 shows summary statistics of average near-surface temperature and wind speed biases and errors for the year of simulations and each surface station used in the verification. For most land-based stations, WRF wind speeds were biased slightly high and temperatures were biased slightly low. Examination of the average diurnal cycle (Fig. S16) reveals that the wind speed bias is present at all hours of the day, whereas the temperature bias is due largely to stronger than observed nocturnal cooling. The latter finding is consistent with overestimated nocturnal low-level stability, which likely contributed to the over-prediction of nighttime CH_4 concentrations at BU (Fig. 3C, Fig. S14), typically located below the nocturnal boundary layer, and the modest under-prediction of nighttime concentrations at COP (Fig. 3A, Fig. S14), which was often above the nocturnal boundary layer. The emissions estimate was derived from afternoon data only and thus was not affected by this model bias.

S4.2 Simple Mass Balance Model

Methane enhancements in the urban core reflect the accumulation of emission inputs into the PBL during transit from the 90-km radius model boundary to the observation point (Fig. 1), less exchange by vertical or horizontal mixing with background air (e.g. from above the PBL, or via a sea breeze circulation). We checked our posterior emission fluxes for consistency with this basic mass balance concept using the HYSPLIT back-trajectory model (39) driven by meteorology from the North American Model mesoscale forecast system with 12 km resolution (40, NAM12).

The model equation for the mass balance model is:

$$Z_{PBL}\,\overline{\Delta CH_4}\,\overline{N_{air}} = < F_{CH4} > \tau_{transit} - K\,\Delta_h CH_4\,\overline{N_h}\,\tau_{transit}$$

where Z_{PBL} is the depth of the mixed layer (m), $\overline{\Delta CH_4}$ is the mean enhancement (mole fraction) in the PBL, $\overline{N_{air}}$ is the average air number density in the PBL (m^{-3}), $< F_{CH4} >$ is the mean surface flux ($m^{-2}\,s^{-1}$), and $\tau_{transit}$ is the transit time (s) within the PBL. The last term in the equation represents column-integrated exchange with background air during transit, where K is the exchange velocity ($m\,s^{-1}$), $\Delta_h CH_4$ is the concentration difference (mole fraction), and N_h is the mean number density (m^{-3}) for the exchange process.

The mass balance model was used to simulate ΔCH_4 at COP for afternoon hours (17-21h UTC) in September and October, 2012 using values for Z_{PBL} and $\tau_{transit}$ from HYSPLIT/NAM12. The boundary layer in most of these simulations was well-developed with transit times to the model boundary between 2 and 8 hours for > 80% of the trajectories. Unlike WRF-STILT, HYSPLIT back trajectories do not simulate exchange and the last term of the model equation was therefore neglected.

This mass balance model and our optimized fluxes yielded mean ΔCH_4 values that were ~25% higher than both observed and simulated values from WRF-STILT. This small excess is consistent with the

absence of an entrainment term in the box model. We conclude that the optimal fluxes from WRF-STILT reported in the main text are consistent with analysis using a simple mass balance approach based on independent meteorology and transport models.

S4.3 Modeling Framework Variants

The sensitivity of the emissions result was tested against several variants of the modeling framework (Fig. S17). Measured CH_4 enhancements presented in the main text (Fig. 2) were aggregated into daily points by taking afternoon (11-16 h EST) averages. This aggregation method was tested by also averaging the four lowest hourly observations and model enhancements in the period of 9-18 h EST each day (Fig. S2, Fig. S17 pt. 2). Both of these approaches to data selection and aggregation aim to focus the analysis on periods when the atmosphere is well-mixed and when the data are less variable, which maximizes the areal representativeness of the results and minimizes the influence of a possible model biases in boundary layer height. Both aggregation methods yielded comparable results.

Optimized emissions resulting from the use of EDGAR and the custom prior flux model at a coarser spatial resolution were not significantly different than the main result which used the custom emission inventory (section 3.2.2, Fig. S17 pts. 3-4). We also tested the null hypothesis for spatial variation of sources, using a model of uniform, constant flux over land and zero flux over water. This variant resulted in similar model-data correlations (Fig. S13) as the spatially-varying flux prior (section 3.2.2), but mean footprint-weighted, optimized emission fluxes were significantly different for the two sites. (The larger footprint from the COP site gave rise to smaller optimized fluxes). In contrast, results using our custom prior flux model yielded optimized fluxes that were statistically indistinguishable at the two sites (Fig. S4). Therefore, we rejected the null hypothesis of no significant spatial variation in emissions and adopted the custom prior model (section 3.2.2) instead.

References

1. Crosson ER (2008) A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor. *Appl Phys B* 92: 403-408.
2. Rella CW, et al. (2013) High-accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. *Atmos Meas Tech* 6: 837-860.
3. Nara H, et al. (2012) Effect of air composition (N_2, O_2, Ar, and H_2O) on CO_2 and CH_4 measurement by wavelength-scanned cavity ring-down spectroscopy: calibration and measurement strategy. *Atmos Meas Tech* 5: 2689-2701.
4. U.S. National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Global Monitoring Division, Central Calibration Laboratory, www.esrl.noaa.gov/gmd/ccl/ccl.html.
5. World Meteorological Organization (2009) 14th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques WMO/TD - No. 1487 (Helsinki, Finland, 10-13 September 2007). GAW Report No. 186.
6. Prasad K, Bova A, Whetstone JR, Novakovskaia E (2013) Greenhouse gas emissions and dispersion 1. Optimum placement of gas inlets on a building rooftop for the measurement of greenhouse gas concentration. NIST Special Publication 1158. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1158.pdf.
7. Urbanski S, et al. (2007) Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest. *J Geophys Res* 112: G02020.
8. Curry CL (2007) Modeling the soil consumption of atmospheric methane at the global scale. *Global Biogeochemical Cycles* 21: GB4012.
9. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2002) Modeling of Data. *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (Cambridge Univ. Press, New York ed. 2), chapter 15.

10. Algonquin Gas Transmission LLC, Link System Informational Postings, Gas Quality, Hourly Chromatograph Postings (Spectra Energy, Houston TX), http://infopost.spectraenergy.com/infopost.
11. Tennessee Gas Pipeline Company LLC, Informational Postings, Gas Quality, Hourly Gas Quality Data (Kinder Morgan, Houston, TX), http://webapps.elpaso.com/PortalUI/DefaultKM.aspx?TSP=TGPD.
12. U.S. Energy Information Administration, Natural Gas Annual Respondent Query System, Annual Report on Natural and Supplemental Gas Supply and Disposition, www.eia.gov/cfapps/ngqs/ngqs.cfm?f_report=RP1.
13. Gas Processing Association (2013) Analysis for natural gas and similar gaseous mixtures by gas chromatography, GPA 2261, www.gpaglobal.org/publications/item/?id=3583.
14. ASTM International (2010) Standard test method for analysis of natural gas by gas chromatography, ASTM D1945, www.astm.org/Standards/D1945.htm.
15. Lin JC, et al. (2003) A near-field tool or simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *J Geophys Res* 108: 4493-4510.
16. Skamarock WC, Klemp JB (2008) A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J Comput Phys* 227: 3465-3485.
17. Nehrkorn T, et al. (2010) Coupled weather research and forecasting-stochastic time-inverted Lagrangian transport (WRF-STILT) model. *Meteorol Atmos Phys* 107: 51-64.
18. Mesinger F, et al. (2006) North American regional reanalysis. *Bull Am Meteorol Soc* 87: 343-360.
19. Nehrkorn T, et al. (2013) WRF simulations of the urban circulation in the Salt Lake City area for CO_2 modeling. *J Appl Meteorology* 52: 323-340.
20. U.S. Energy Information Administration (2014) Natural Gas Consumption by End Use, www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_m.htm.
21. U.S. Energy Information Administration (2013-2014) 2012 and 2013 Annual Electric Utility Data, Form EIA-923, www.eia.gov/electricity/data/eia923.
22. The Warren Group, Inc. (2010) Boston, MA, www.thewarrengroup.com.
23. U.S. Department of Energy (2014) Alternative Fuel Station Locator, www.afdc.energy.gov/locator/stations.
24. U.S. Census Bureau (2012) 2010 American Community Survey 5-year Estimates, www.census.gov.
25. U.S. Energy Information Administration, Natural Gas Monthly, Table C1, Standard Error for Natural Gas Deliveries and Price by Consumers, by State, www.eia.gov/naturalgas/monthly/archive.
26. European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL) (2013) Emission Database for Global Atmospheric Research (EDGAR), release version 4.2 FT2010, edgar.jrc.ec.europa.eu.
27. U.S. Fish and Wildlife Service, National Wetlands Inventory, www.fws.gov/wetlands.
28. Bridgham SD, Megonigal JP, Keller JK, Bliss NB, Trettin C (2007) Wetlands - Supplemental Materials. The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, eds AW King, et al., pp 177-192, http://cdiac.ornl.gov/SOCCR.
29. U.S. Department of Agriculture (2009), 2007 Census of Agriculture, www.agcensus.usda.gov.
30. U.S. Environmental Protection Agency (1998) AP 42, 5th edition, volume 1, Compilation of Air Pollutant Emission Factors, Chapter 14, Greenhouse Gas Biogenic Sources, Section 4, Enteric Fermentation, www.epa.gov/ttn/chief/ap42/ch14/index.html.
31. U.S. Environmental Protection Agency (2014) Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012, EPA-430-R-14-003, www.epa.gov/climatechange/ghgemissions/usinventoryreport.html.
32. U.S. Environmental Protection Agency (2008) Greenhouse Gas Inventory Protocol Core Module Guidance – Direct Emissions from Mobile Combustion Sources, EPA Climate Leaders, Tables 3, A-6, and A-7.

33. Federal Highway Administration (2012), Highway Statistics Series, Highway Statistics 2012, Table VM-4, www.fhwa.dot.gov/policyinformation/statistics/2012.
34. Gately CK, Hutyra LR, Wing IS, Brondfield MN (2013) A bottom up approach to on-road CO_2 emissions estimates: improved spatial accuracy and applications for regional planning. *Env Sci Tech* 47: 2423-2430.
35. U.S. Environmental Protection Agency (2014) 2012 and 2013 Greenhouse Gas Emissions from Large Facilities, www.ghgdata.epa.gov/ghgp.
36. Massachusetts Executive Office of Environmental Affairs (2014) Massachusetts Annual Greenhouse Gas Emissions Inventory: 1990 through 2011, with Partial 2012 Data, www.mass.gov/eea/docs/dep/air/climate/maghginv.xls.
37. National Center for Atmospheric Research, NCEP ADP Global Surface Observational Weather Data, October 1999 – continuing, http://rda.ucar.edu/datasets/ds461.0.
38. National Center for Atmospheric Research, Developmental Testbed Center, Model Evaluation Tools, www.dtcenter.org/met/users.
39. U.S. National Oceanic and Atmospheric Administration, Air Resources Laboratory, Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) Model, www.arl.noaa.gov/HYSPLIT_info.php.
40. National Center for Environmental Prediction, North American Model, www.nco.ncep.noaa.gov/pmb/products/nam.
41. U.S. Energy Information Administration, www.eia.gov/maps/map_data/EIA_NaturalGas_InterIntrastate_Pipelines_US.zip.
42. California Air Resources Board 2005 Almanac, Appendix D.
43. Moriizumi J, NagamineK , Iida T, Ikebe Y (1996) Methane in an urban area of Nagoya, Japan, inferred from atmospheric radon-222 data. *Atmosph Environ* 30: 1543-1549.
44. Lamb BK, et al. (1995) Development of atmospheric tracer methods to measure methane emissions from natural gas facilities and urban areas. *Environ Sci Technol* 29: 1468-1479.
45. Shorter JH, et al. (1996) Methane emission measurements in urban areas in eastern Germany. *J Atmos Chem* 24: 121-140.
46. Fowler D, et al. (1996) Measurements of regional CH_4 emissions in the UK using boundary layer budget methods. *Energy Convers Mgmt* 37: 769-775.
47. Levin I, Glatzel-Mattheier H, Marik T, Cuntz M, Schmidt M (1999) Verification of German methane emission inventories and their recent changes based on atmospheric observations. *J Geophys Res* 104: 3447-3456.
48. Kuc T, Rozanski K, Zimnoch M, Necki JM, Korus A (2003) Anthropogenic emissions of CO_2 and CH_4 in an urban environment. *Applied Energy* 75: 193-2003.
49. Zinchenko AV, Paramonova NN, Privalov VI, Reshetnikov AI (2002) Estimation of methane emissions in the St. Petersburg, Russia, region: An atmospheric nocturnal boundary layer budget approach. *J Geophys Res* 107: 4416.
50. Su F, Shao M, Cai X, Zeng L, Zhu T (2003) Estimates of methane emissions in Beijing using a backward trajectory inversion model. *Chemical Speciation and Bioavailability* 14: 43-48.
51. Hsu YK, et al. (2010) Methane emissions inventory verification in southern California. *Atmosph Environ* 44: 1-7.
52. Wunch D, Wennberg PO, Toon GC, Keppel-Aleks G, Yavin YG (2009) Emissions of greenhouse gases from a North American megacity. *Geophys Res Lett* 36: L15810.
53. Mays KL, et al. (2009) Aircraft-based measurements of the carbon footprint of Indianapolis. *Environ Sci Technol* 43: 7816-7823.
54. Wennberg PO, et al. (2012) On the sources of methane to the Los Angeles atmosphere. *Environ Sci Technol* 46: 9282-9289.
55. Peischl J, et al. (2013) Quantifying sources of methane using light alkanes in the Los Angeles basin, California. *J Geophys Res* 118: 1-17.
56. Santoni GW (2013) *Fluxes of atmospheric methane using novel instruments, field measurements, and inverse modeling*, Doctoral Thesis, Harvard University.

57. Wecht KJ, et al. (2014) Spatially resolving methane emissions in California: constraints from CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI, geostationary) satellite observations. *Atmos Chem Phys Discuss* 14: 4119-4148.
58. Gioli B, et al. (2012) Methane and carbon dioxide fluxes and source partitioning in urban areas: the case study of Florence, Italy. *Environ Poll* 164: 125-131.
59. O'Shea SJ, et al. (2014) Area fluxes of carbon dioxide, methane, and carbon monoxide derived from airborne measurements around Greater London: A case study during summer 2012. *J Geophys Res* 119: 4940-4952.

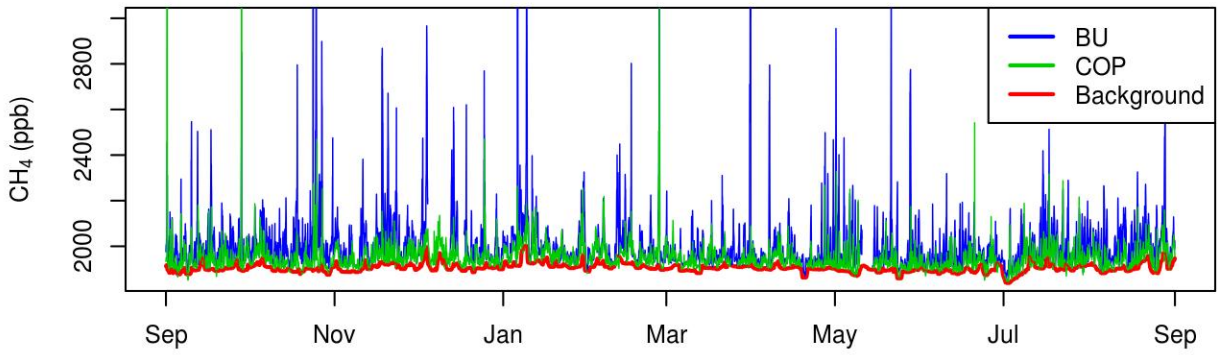

Fig. S1. Mean hourly measured CH_4 concentrations at BU and COP and in the empirical background from September 1, 2012 through August 31, 2013.

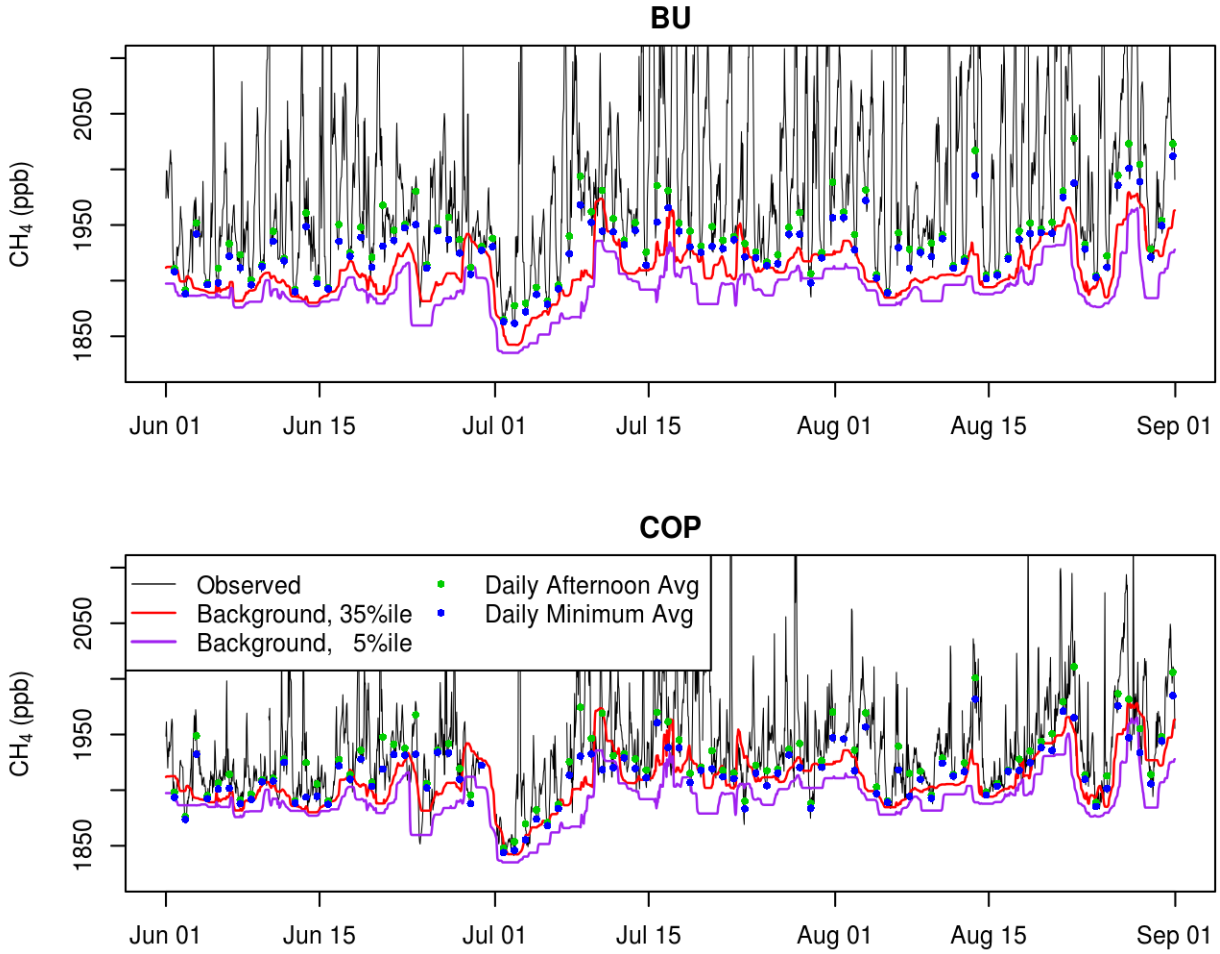

Fig. S2. Mean hourly measured CH_4 concentrations (black lines), the range of empirical background concentrations from upwind stations (red lines), and the daily afternoon average (green) and minimum (blue) points that represent mean enhancements at BU and COP, from an example period of three months in 2013.

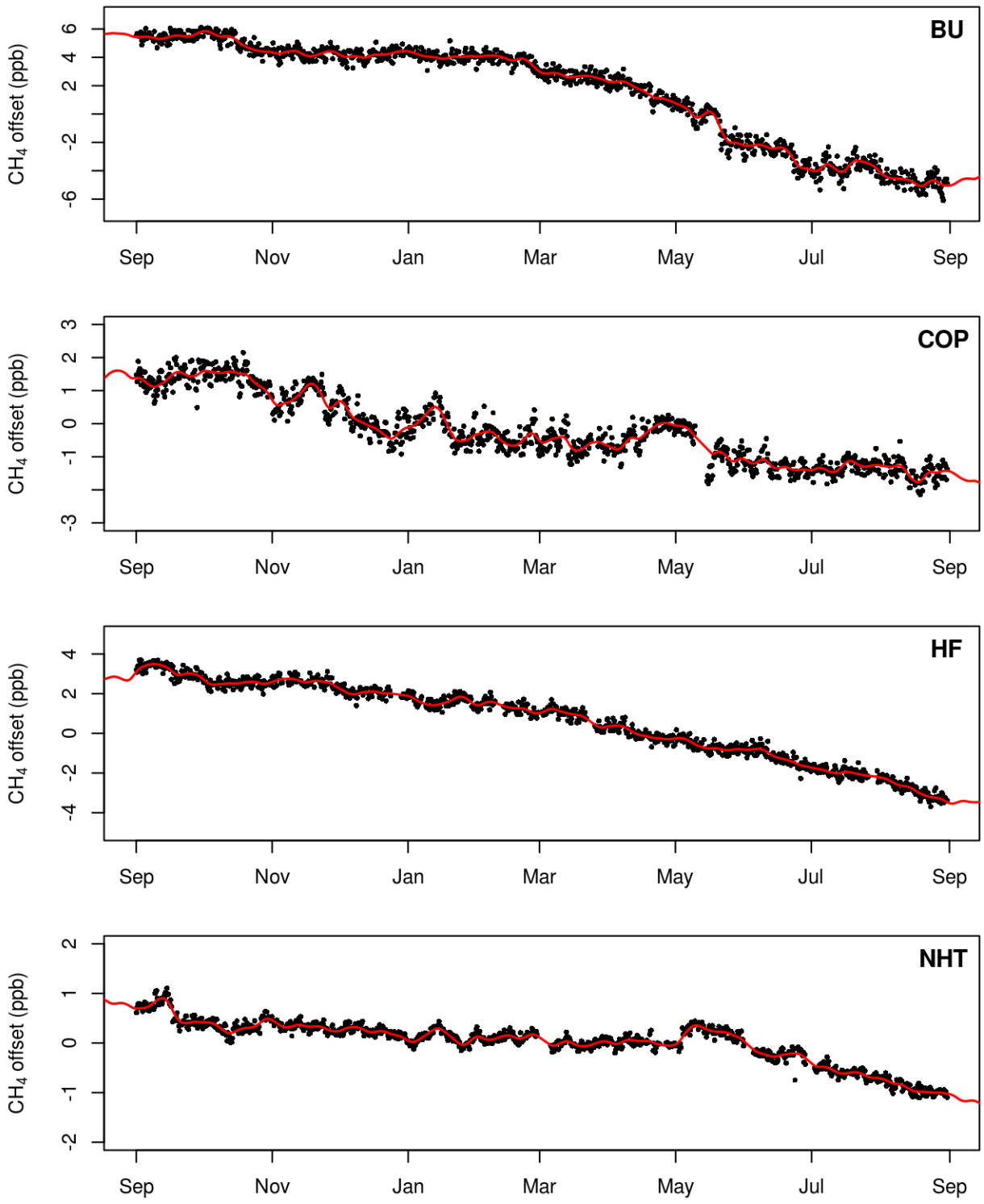

Fig. S3. CH_4 calibration intercepts, calculated as the difference between known and measured surveillance standard values for the four instruments over one year. Black points are average measured values for each ~2.5 minute measurement period and the red line is the smoothed intercept used to correct the data. All plots were shifted to have zero mean offsets to aid visualization of the drifts.

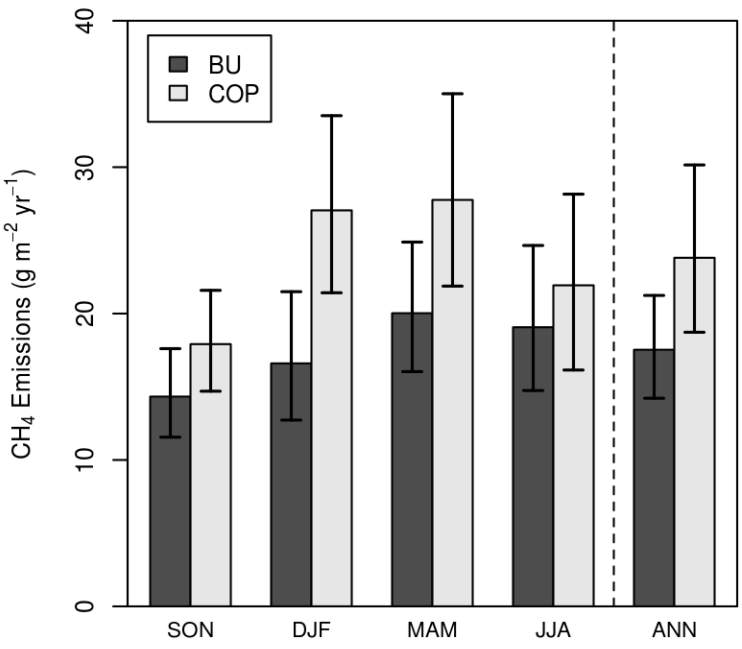

Fig. S4. Optimized CH_4 emissions seasonally and annually, derived from CH_4 observations from BU and COP individually.

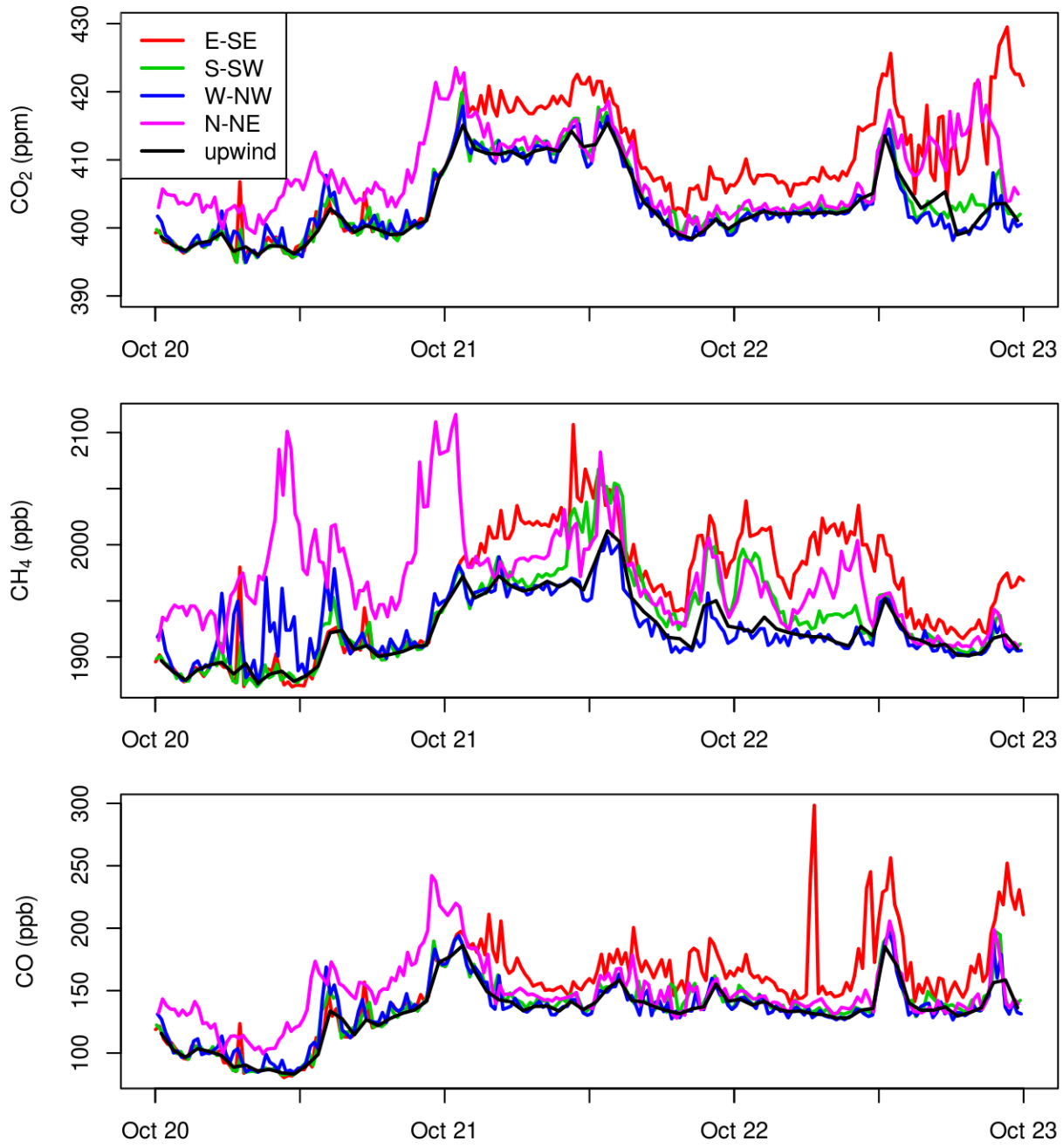

Fig. S5. Example of the COP corner-sampling method for three days in October, 2012. The colored lines show average concentrations during each 5-min sampling interval from each corner. The black line shows the average hourly inferred upwind concentration. The legend describes the orientation of each corner.

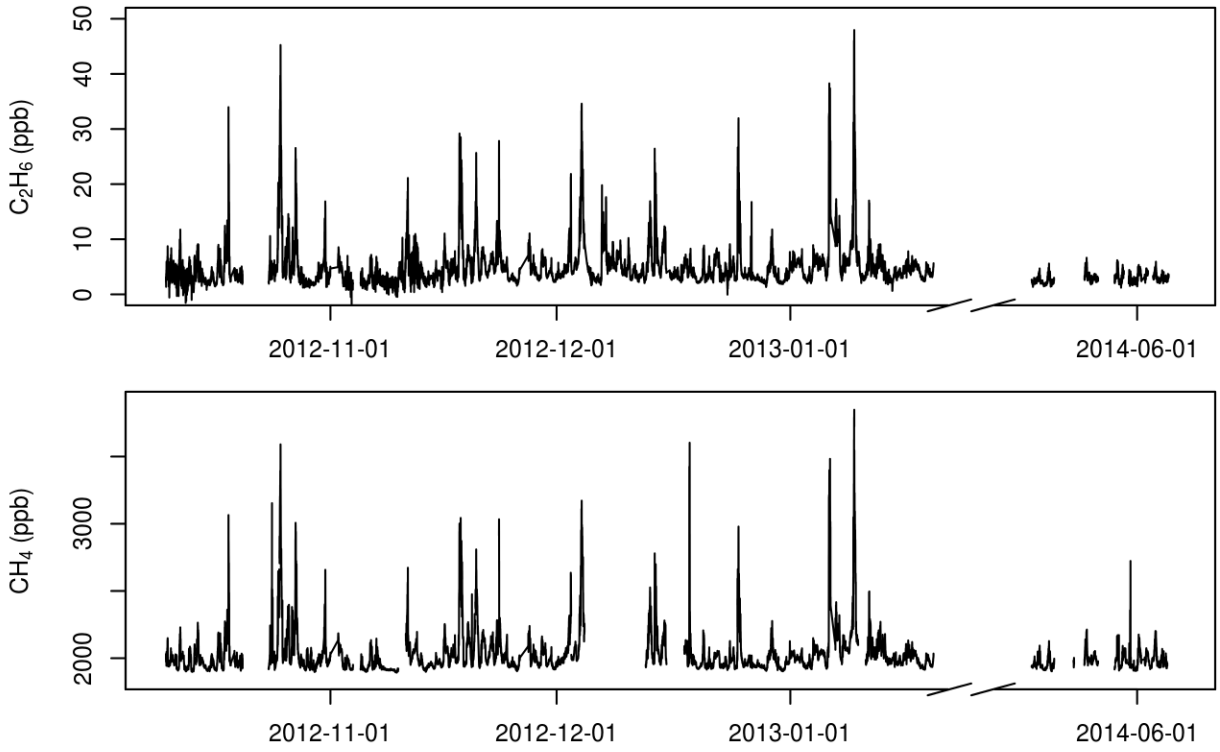

Fig. S6. Five-minute median C_2H_6 and CH_4 measurements from BU.

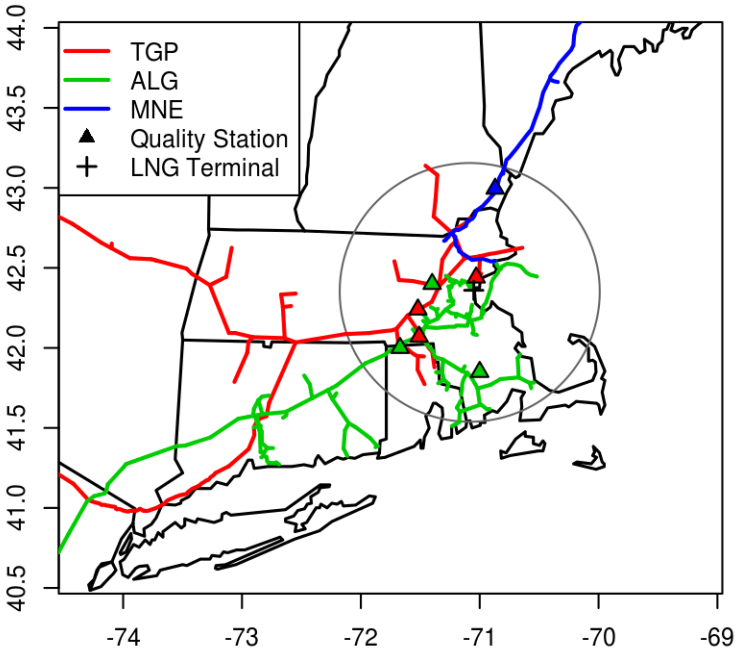

Fig. S7. Approximate locations of the three interstate gas pipelines (Tennessee – TGP, Algonquin – ALG, and Maritimes and Northeast – MNE, 42) serving Boston and the surrounding area, the gas quality measurement stations used in this study, and the LNG import terminal. The gray circle is the study area boundary.

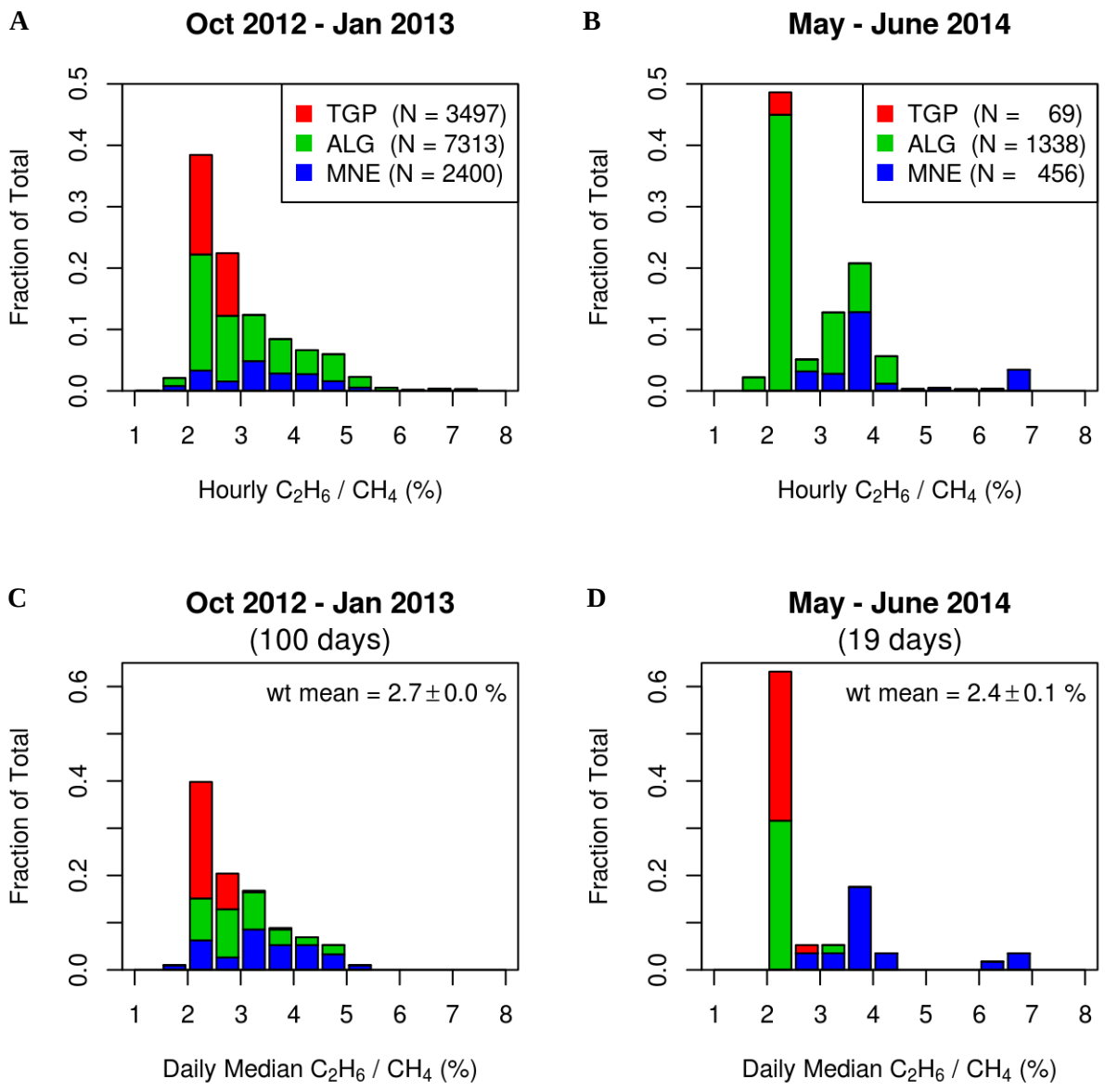

Fig. S8. Stacked histograms of (A,B) hourly and (C,D) daily median ratios of C_2H_6 and CH_4 in the pipeline gas during the two time periods of the atmospheric C_2H_6 measurements (Fig. S5). (A,B) Hourly data were obtained from the three pipelines (Tennessee – TGP, Algonquin – ALG, and Maritimes and Northeast – MNE, 10-11) for the stations shown in Fig. S6. (C,D) Daily median ratios for each pipeline were used to estimate a mean ratio for each of the two time periods, weighted by the volumes delivered by each pipeline to Massachusetts. Each of the three pipelines is equally represented in the daily median plots, but not the hourly plots because hourly data coverage varied between stations.

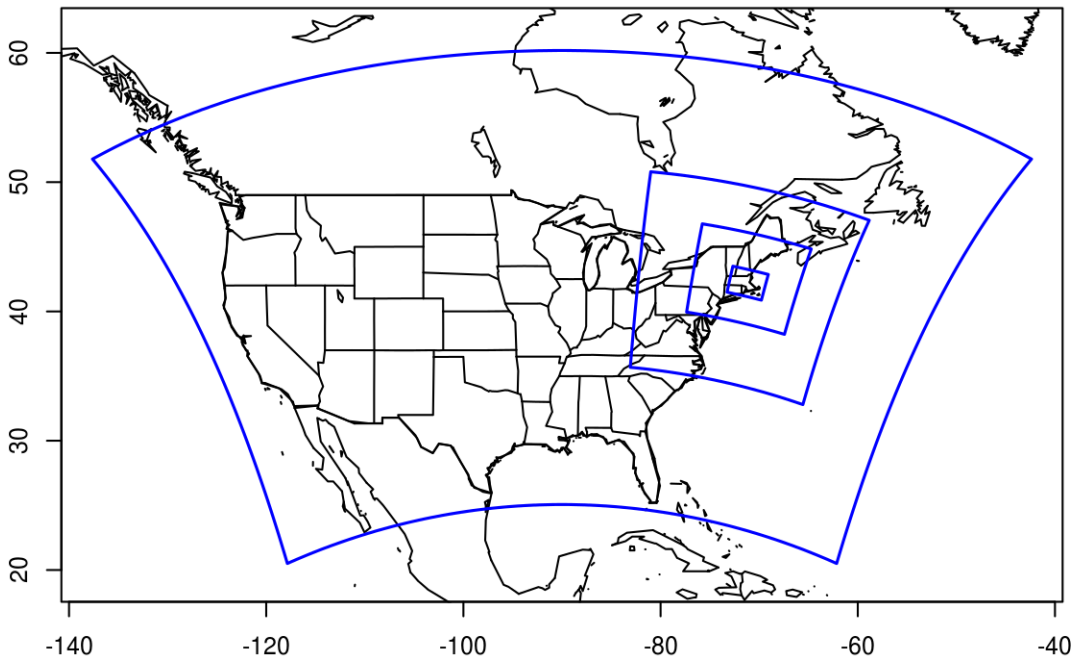

Fig. S9. Location of the four nested WRF domains. The horizontal resolutions of the largest to smallest domains are 27, 9, 3, and 1 km.

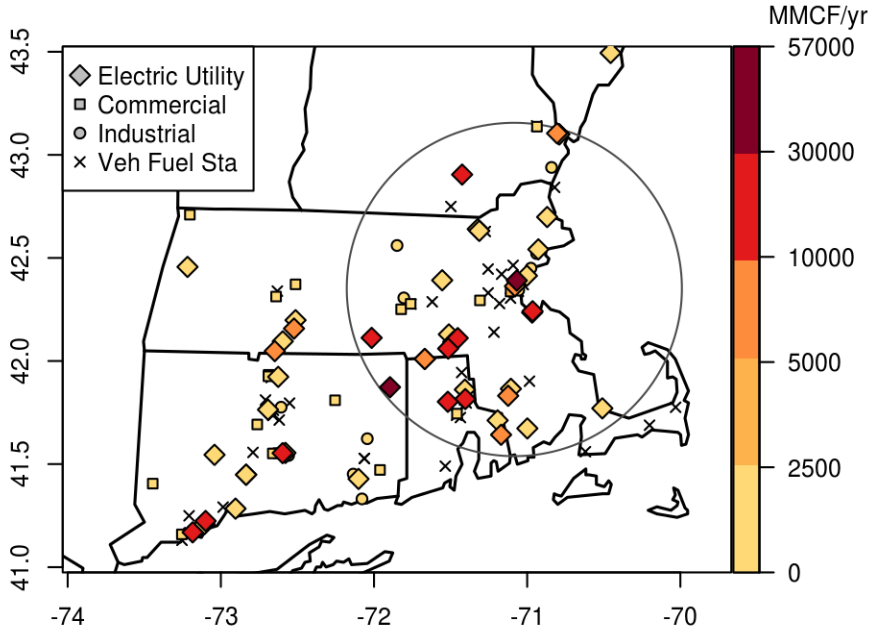

Fig. S10. Location and gas consumption (units: 10^6 scf yr^{-1}) during the study year of each NG-fueled power plant (21) and CNG or LNG vehicle fueling station (23) in the four states included in the study area, delineated by the gray circle.

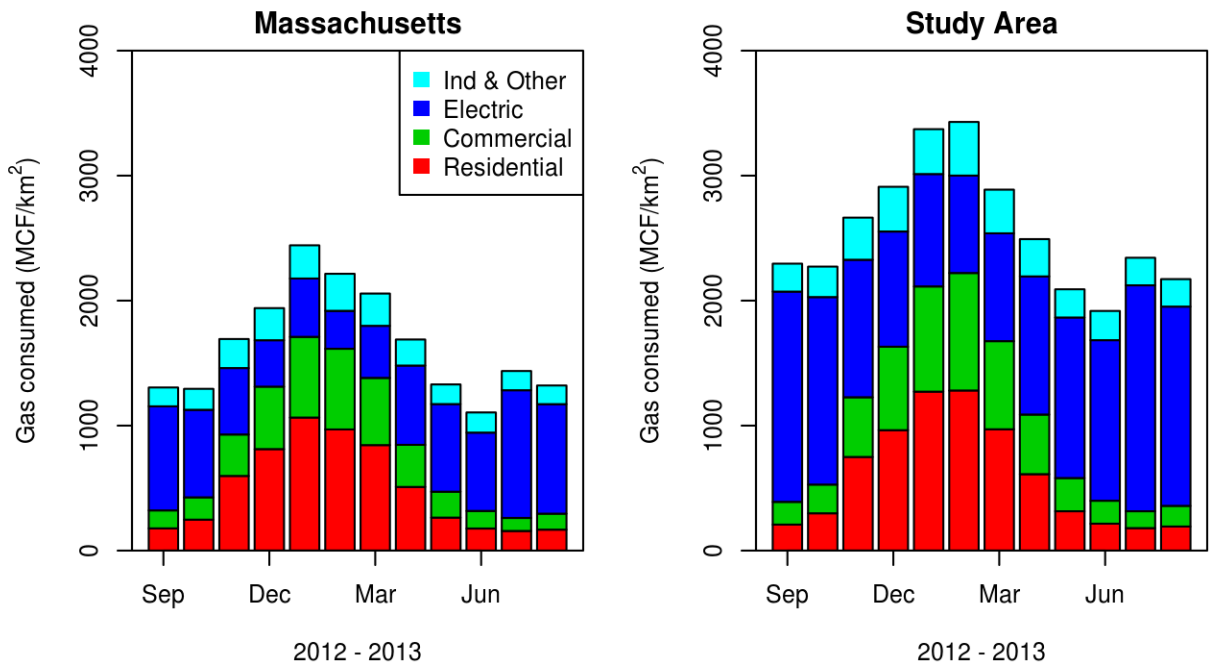

Fig. S11. Average monthly NG consumption by sector in Massachusetts (20) and the study area, spatially allocated as described in section 3.2.1. The "Ind & Other" category includes industrial, vehicle fuel, and pipeline and distribution use.

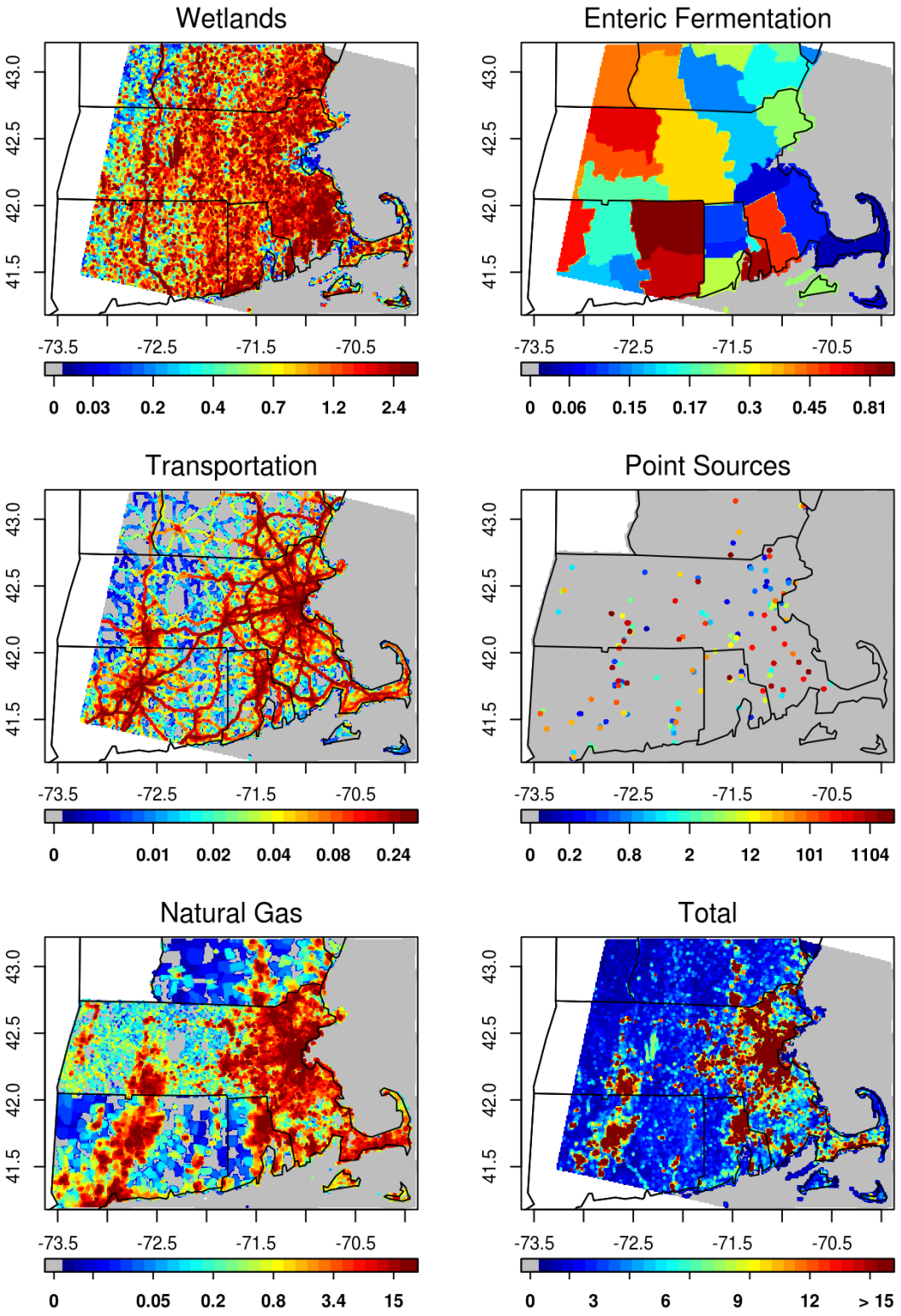

Fig. S12. Maps of prior Emissions (g CH_4 m^{-2} yr^{-1}) by source type and in total on a 1 km^2 grid from the inventory constructed for this study. Scale bars for individual sectors are not linear and were set to have an equal sample size in each bin in order to better render spatial patterns. The scale bar for total emissions is linear. The methods used to construct these maps are described in section 3.2.2.

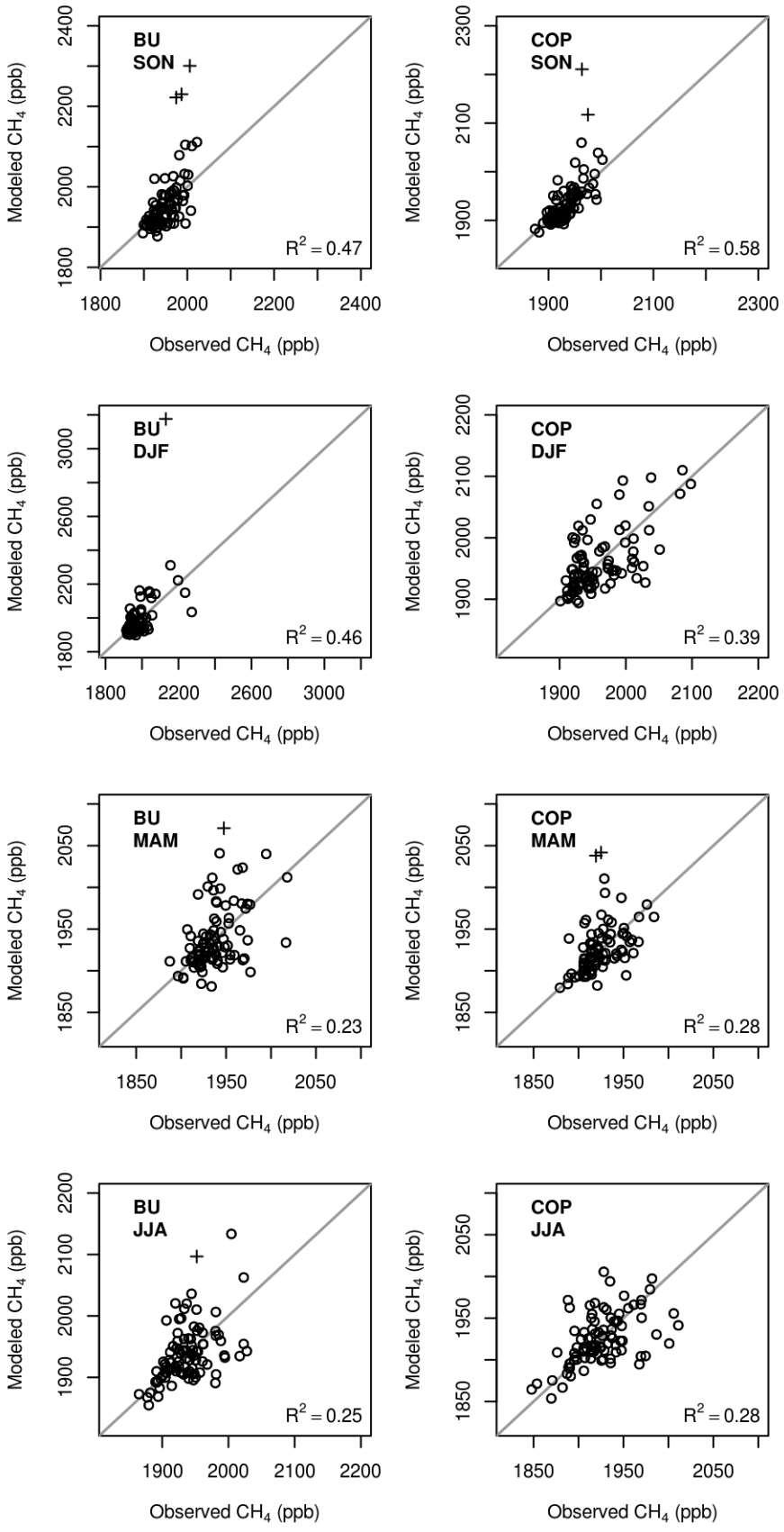

Fig. S13. Optimized modeled daily versus observed average afternoon CH_4 concentrations for the two sites and four seasons. The gray line is the one-to-one line. Outlying points marked by crosses were excluded from model optimization.

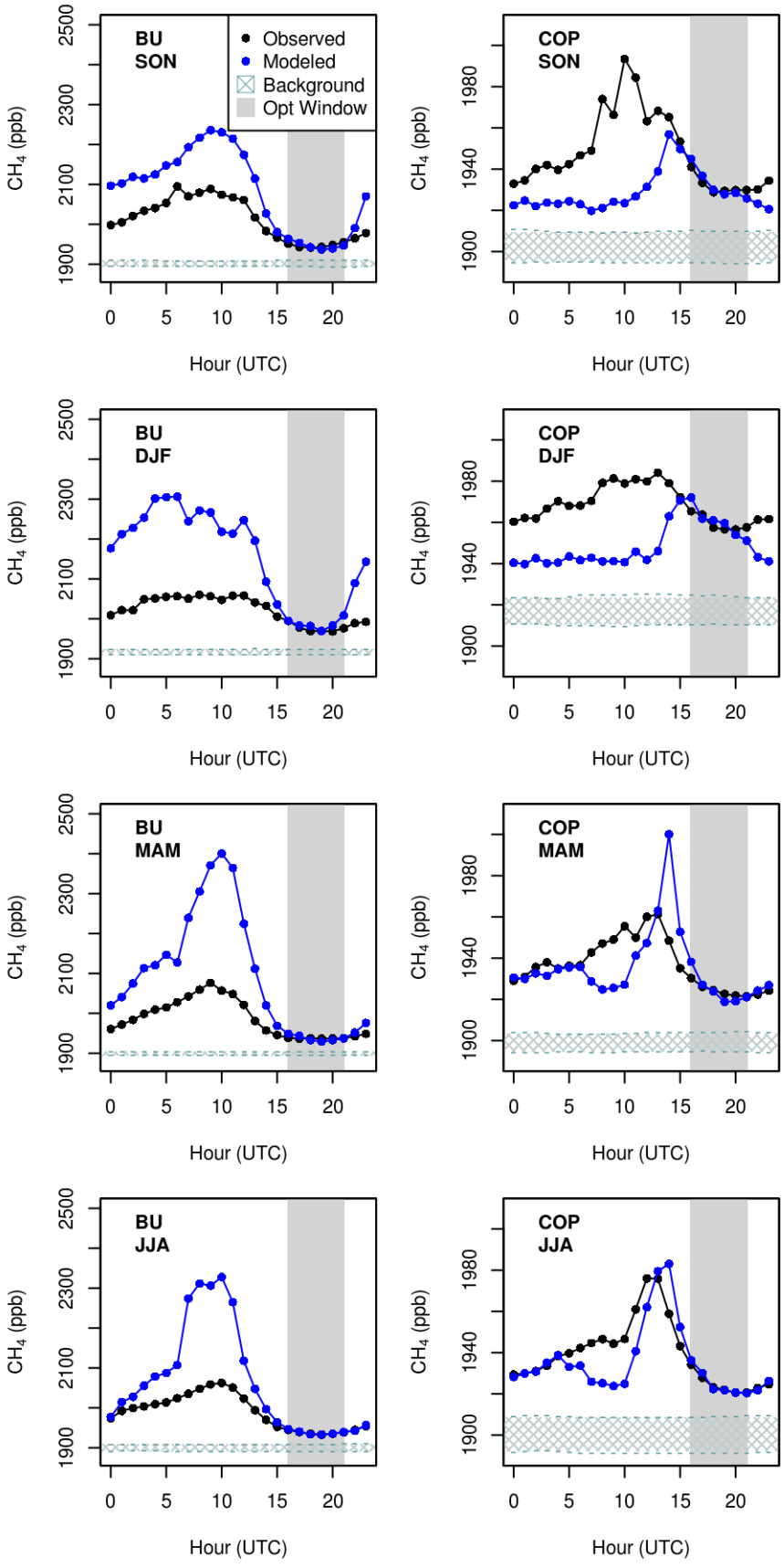

Fig. S14. Observed, optimized modeled, and background CH_4 concentrations averaged by hour of the day for the two sites and four seasons. The horizontal hatched area shows the average range of sampled background concentrations, derived from 5-35th percentiles of the background station data. The gray vertical shaded area indicates the afternoon model optimization period, 11-16 h EST (16-21 h UTC). Although both high and low biases exist in the nighttime model data, modeled and observed data from the afternoon optimization window are in good agreement.

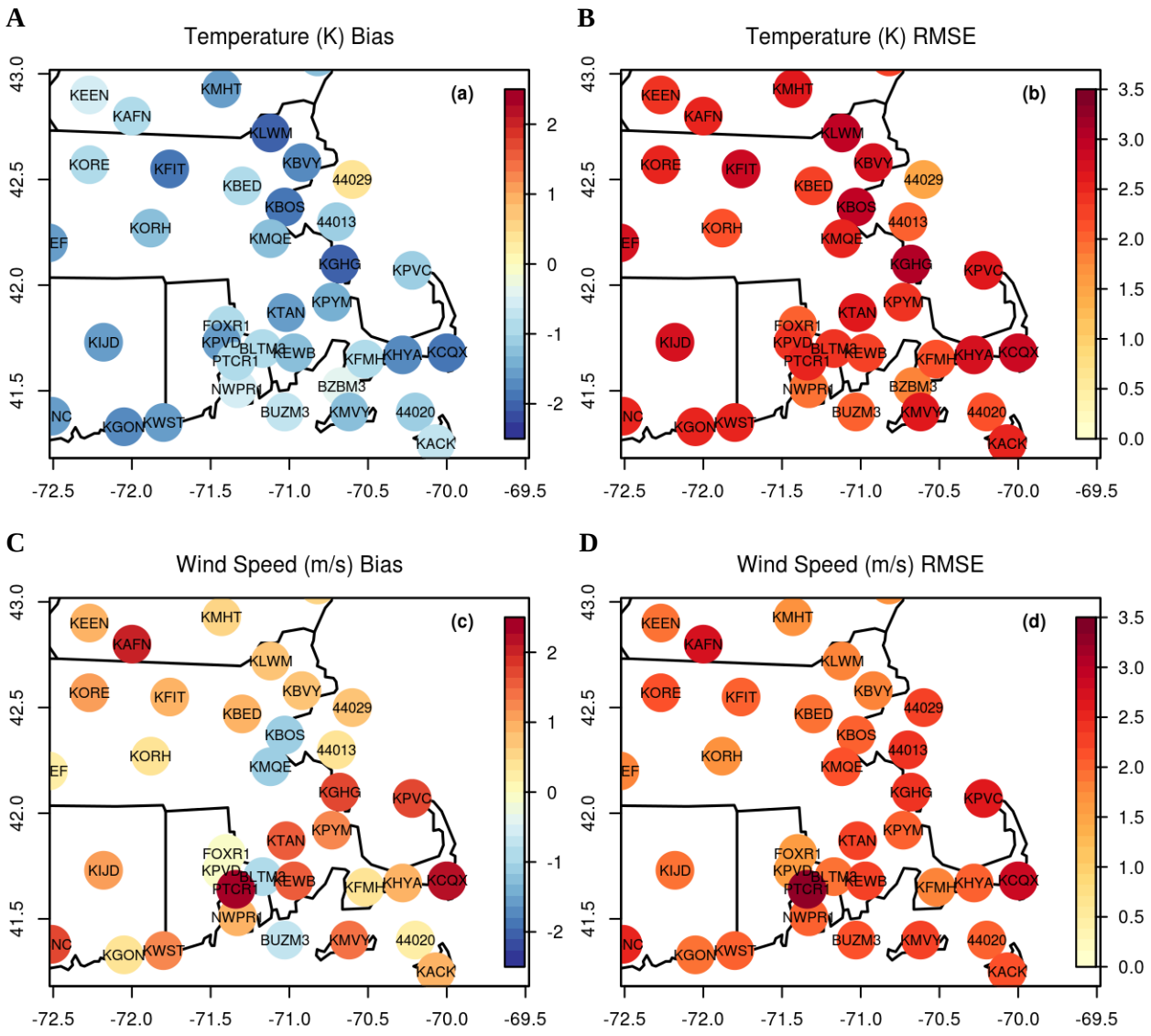

Fig. S15. Average (A,C) bias and (B,D) root mean square error (RMSE) of WRF-simulated (A,B) near-surface temperature and (C,D) wind speed for National Weather Service stations (37) in the innermost WRF domain.

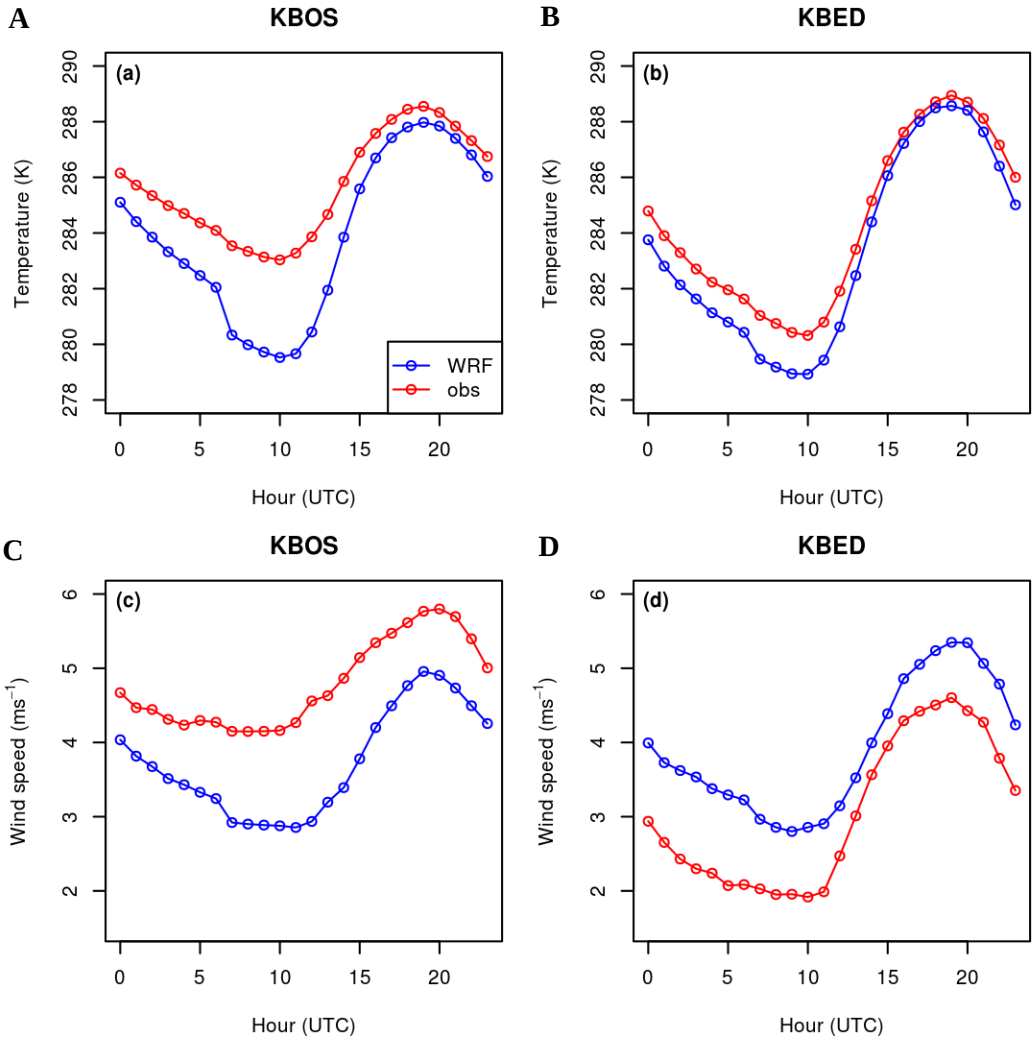

Fig. S16. Observed (red) and WRF-simulated (blue) near-surface (A,B) temperature (C,D) and wind speed, averaged by hour (UTC) for one year and stations (A,C) KBOS (at Boston Logan airport) and (B,D) KBED (northwest of Boston).

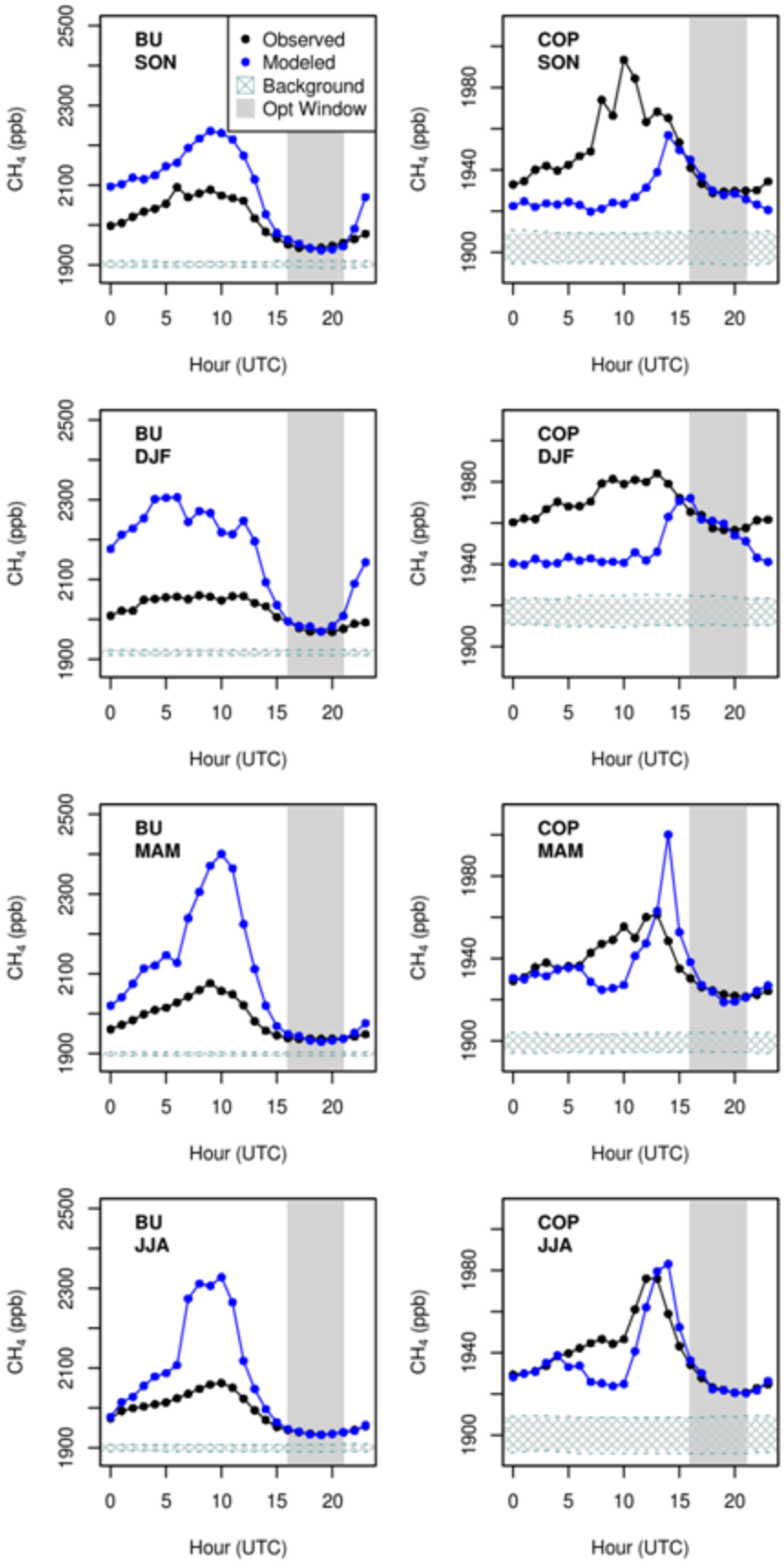

Fig. S16. Observed, optimized modeled, and background CH_4 concentrations averaged by hour of the day for the two sites and four seasons. The horizontal hatched area shows the average range of sampled background concentrations, derived from 5-35th percentiles of the background station data. The gray vertical shaded area indicates the afternoon model optimization period, 11-16 h EST (16-21 h UTC). Although both high and low biases exist in the nighttime model data, modeled and observed data from the afternoon optimization window are in good agreement.

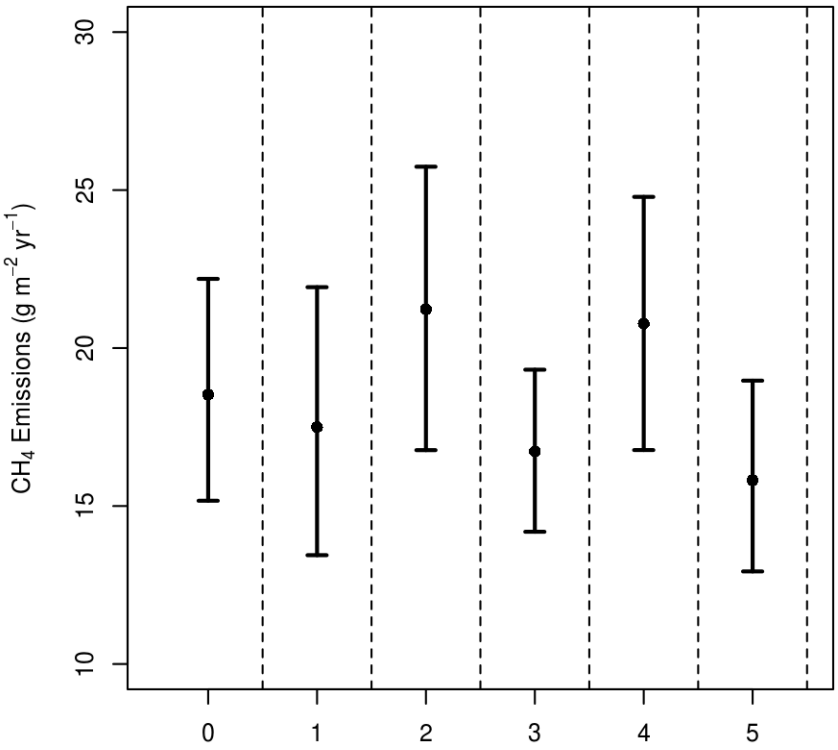

Fig. S17. Mean annual emission results from the main configuration (point 0) and several variants of the modeling and analysis framework (points 1-5). The emission estimate presented in the main text (point 0) resulted from the custom emission inventory at 1 km^2 spatial resolution, data aggregation into daily afternoon (11-16 h EST) points, and removal of extreme outliers. Point 1 shows the emission results when outliers were not removed. Point 2 shows the emission result from aggregating the four lowest hourly observed concentrations and model enhancements in the period of 9-18 h EST each day instead of afternoon hours (Fig. S2). Point 3 shows the emission result from EDGAR instead of the custom emissions model (section 3.2.2). Point 4 shows the result when the custom emissions model was aggregated to a coarser spatial resolution. Point 5 shows the result when the NG emissions layer in the custom prior was scaled to contribute 80% of the total emissions (~2.2 times larger than in the main configuration), to be approximately consistent with the attribution results from ethane data.

Table S1. Methane emissions in urban areas from atmosphere-based ("top-down") studies. Only studies that reported emission rates averaged in time and space are listed.

Ref.	Location	Measurement year	Emission Rate (g CH_4 m^{-2} yr^{-1})
43	Nagoya, Japan	1990-91	7
44	Midwest town, USA	1991	55
45	Two towns in East Germany	1992	12, 60
46	North Britain	1994	28 – 56
47	Heidelberg, Germany	1995-97	8 ± 2
48	Krakow, Poland	1996-97	20
49	St. Petersburg, Russia	1996-2000	32 ±9
50	Beijing, China	2000	50
51	Los Angeles County, CA, USA	2007-08	205 ± 6*
52	South Coast Air Basin, CA, USA	2007-08	228 ± 38*
53	Indianapolis, IN, USA	2008	71 ± 50
54	South Coast Air Basin, CA, USA	2010	167 ± 57*
55	South Coast Air Basin, CA, USA	2010	156 ± 14*
56	South Coast Air Basin, CA, USA	2010	127 ± 21*
57†	South Coast Air Basin, CA, USA	2010	160 ± 30*
57‡	South Coast Air Basin, CA, USA	2010	118 ± 30*
58	Florence, Italy	2011	58
59	London, UK	2012	66 ± 10

*Basin-total fluxes reported in the California studies were converted to average area fluxes using areas from the California Air Resources Board (42).

†Aircraft observations

‡Satellite observations

Table S2. Locations and instrument models of the four measurement sites.

Site	**Longitude (˚)**	**Latitude (˚)**	**Height (m above ground)**	**Picarro model**
Boston University (BU)	-71.10	42.35	29	G2101-i
Copley Square (COP)	-71.08	42.35	215	G2401
Harvard Forest (HF)	-72.17	42.54	29	G2301
Nahant (NHT)	-70.91	42.42	16	ESP-1000

Table S3. Methane measurement performance statistics for the four sites and one year. Individual measurement precision is the average standard deviation of the raw 2-6 second surveillance measurements made throughout the year at each site (Fig. S3). Surveillance tank uncertainties derive from both uncertainties in the primary NOAA calibration tanks (4) against which they were calibrated and from the measurement precision of the instrument used to calibrate the tanks. The surveillance tank at BU is still in use and has not yet been returned for a second, post-deployment calibration.

Site	**Long-term Instrument Drift (ppb)**	**Individual Measurement Precision (1σ) (ppb)**	**Surveillance Tank Uncertainty (1σ) (ppb)**	**Surveillance Tank Long-term Drift (ppb)**
BU	10.5	0.9	0.1	--
COP	2.8	0.3	0.2	0.1
HF	6.4	0.3	0.1	0.1
NHT	1.7	0.3	0.1	0.2

Table S4. Average CH_4 emission in Massachusetts and in the 90-km radius study area (Fig. 1) by sector and in total from two prior emission models and the Massachusetts GHG inventory (36). The Massachusetts inventory was not tested in the modeling framework because it is not spatially resolved.

Inventory	**Sector**		**Average Emissions ($g\ CH_4\ m^{-2}\ yr^{-1}$) (% of total)**	
			Massachusetts	**Study Area**
Custom*	Wetlands		0.93 (16%)	1.01 (14%)
	Enteric Fermentation		0.25 (4%)	0.22 (3%)
	Transportation		0.05 (1%)	0.06 (1%)
	Large Point Sources		0.88 (15%)	1.38 (19%)
	Natural Gas		3.72 (64%)	4.68 (64%)
	Total		5.83	7.35
EDGAR (2010)	Energy production (1A1_1A2)		0.17 (2%)	0.24 (2%)
	Non-road Transportation (1A3a_c_d_e)		0.00	0.00
	Road Transportation (1A3b)		0.05 (1%)	0.06 (1%)
	Stationary Combustion (1A4)		0.28 (4%)	0.38 (4%)
	Fugitive from Solid Fuels (1B1)		0.00	0.00
	Oil Production & Refining (1B2a)		0.12 (2%)	0.15 (2%)
	Natural Gas Production & Distribution (1B2b)		3.50 (47%)	4.75 (49%)
	Industrial (2)		0.03 (0%)	0.04
	Enteric Fermentation (4A)		0.15 (2%)	0.14 (1%)
	Manure Management (4B)		0.05 (1%)	0.05 (0%)
	Agricultural Soils (4C_4D)		0.00	0.00
	Agricultural Waste Burning (4F)		0.00	0.00
	Solid Waste Disposal (6A_6C)		2.13 (28%)	2.44 (25%)
	Waste Water (6B)		1.04 (14%)	1.44 (15%)
	Fossil Fuel Fires (7A)		0.00	0.00
	Total		7.50	9.67
Massachusetts State (2011)	Stationary Combustion	Residential	0.19 (3%)	
		Commercial	0.05 (1%)	
		Industrial	0.01 (0%)	
		Electric Power	0.01 (0%)	
	Mobile		0.09 (1%)	
	Natural Gas Transmission & Distribution		3.73 (58%)	
	Enteric Fermentation		0.19 (3%)	
	Manure Management		0.03 (1%)	
	Landfills & Waste Combustion		1.16 (18%)	
	Wastewater		0.96 (15%)	
	Total		6.43	

*The custom prior covered the majority, but not the entire state of Massachusetts (Fig. S12), so the average emissions rate for Massachusetts was calculated from the area of the state that is covered by the inventory.