Supplemental for

Seasonally resolved excess urban methane emissions from the Baltimore/Washington, DC metropolitan region

Yaoxian Huang[*,†,‡], Eric A. Kort[†], Sharon Gourdji[†,§], Anna Karion[§], Kimberly Mueller[†,§], and John Ware[†]

[†]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan 48109, United States

[‡]Department of Civil and Environmental Engineering, Wayne State University, Detroit, Michigan 48202, United States

[§]National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States

[*]E-mail: yaoxian.huang@wayne.edu. Phone: +1-313-577-7829

Number of pages: 16

Number of figures: 9

Number of tables: 3

Table of Contents

1. Re-allocation of wetland methane emissions over the D03 domain for June Page S2

2. Descriptions of trajectory-based background method Page S2

3. Validation of geostatistical inversions Page S3

1. Re-allocation of wetland methane emissions over the D03 domain for June. We use the high-resolution (30m×30m) National Land Cover Database[1] for the year 2011 (NLCD2011) to better allocate wetland CH_4 emissions over the D03 domain and prevent coarse gridding from falsely attributing significant wetland emissions to the D03 domain. The wetland fractions in NLCD2011 were up-gridded from 30m Latitude by 30m Longitude to 0.01° Latitude by 0.01° Longitude. We also downscale the original wetland CH_4 emissions at the resolution of 0.5° Latitude by 0.5° Longitude to 0.01° Latitude by 0.01° Longitude to be consistent with model resolution. We choose a domain (denoted as D02 domain, lat: 33°N-45°N, lon: 80°W-70°W) covering most of eastern coast of the US from our D01 domain, where wetland fractions in each grid box were relatively high, shown in Figure S2. The regridded wetland CH_4 emissions were multiplied by the wetland fractions from NLCD2011 over the D02 domain. A scaling factor is computed as the ratio of original wetland CH_4 emissions over the D02 domain before re-allocation to the total values after applying the wetland fractions from NLCD2011. In order to conserve the mass of original total wetland CH_4 emissions, we applied this scaling factor to the updated wetland CH_4 emissions over the D02 domain, with final spatial distributions of wetland CH_4 emissions shown in Figure S3.

2. Descriptions of trajectory-based background method. Due to the limited background observational site available over D01 domain, we exclusively rely on LEF site (Park Falls, Wisconsin, United States, 45.945°N, 90.273°W, sampling height 396 m above the ground level) for our background values for the trajectory-based method. With the LEF longitude as our boundary, we firstly calculate the ensemble mean hours for particles from all four observed sites (ARL, HAL, NDC and BUC) over NEC-B/W arriving at LEF longitudinal

boundary for each hourly observation, using STILT output. It usually takes 2-4 days for particles arriving at LEF, depending on different seasons and individual days. Then we trace back the time to determine the corresponding background value at LEF. Since there is substantial diurnal variability of CH_4 mole fractions measured at LEF, we only use in situ afternoon (11am – 4pm LST) CH_4 mole fractions. For this reason, we approximate the ensemble mean time for particles arriving at LEF longitudinal boundary (rounded up to days) to calculate the hourly-varying background values.

3. Validation of geostatistical inversions. To verify the inversion setup and validate the inversion results, we have followed Michalak et al[2] to include a χ_R^2 statistic (Equation S1) from each ensemble member for Feb. 2016 to assess the improvement of fit statistically.

$$\chi_R^2 = \frac{1}{n}(\boldsymbol{z} - \boldsymbol{Hs}_{ci})^T R^{-1}(z - \boldsymbol{Hs}_{ci}) \quad \text{(S1)}$$

where n is the number of observations (**z**); $\boldsymbol{H}$ is the footprint matrix and $\boldsymbol{s}_{ci}$ is the conditional realizations; $\boldsymbol{R}$ is model-data mismatch covariance matrix.

Our results show that the ensemble mean of χ_R^2 from each individual ensemble member is close to 1, suggesting that our inversions are statistically consistent[2,3].

Table S1. Observational site information over NEC-B/W.

Site full name	Site code	Latitude (°N)	Longitude (°W)	Sampling height (m)*
Arlington, VA	ARL	38.892	77.132	92
Halethorpe, MD	HAL	39.255	76.675	58
Washington, DC	NDC	38.950	77.08	91
Bucktown, MD	BUC	38.460	76.043	75

*The elevation is above ground level.

Table S2. Configurations of various simulations. Each ensemble is calculated from the mean of the 18 ensemble members (3 emission inventories, 3 met products and 2 background methods).

Simulation name	Ensemble a priori or optimized	Simulation month
Ensemble_apriori_201602	ensemble a priori	Feb.
Ensemble_optimized_201602	ensemble optimized	Feb.
Ensemble_apriori_201604	ensemble a priori	Apr.
Ensemble_optimized_201604	ensemble optimized	Apr.
Ensemble_apriori_201606	ensemble a priori	June
Ensemble_optimized_201606	ensemble optimized	June
Ensemble_apriori_201611	ensemble a priori	Nov.
Ensemble_optimized_201611	ensemble optimized	Nov.

Table S3. Regression slopes (*k*), correlation coefficients (*r*), mean errors and root mean square errors between observations and ensemble model simulations for the months of Feb., Apr., June and Nov. 2016. Mean errors are calculated as the mean of observations minus modeled values. Configurations of simulations are shown in Table S1.

Site name	Simulation name	*k*	*r*	*mean errors* (ppb)	*root mean square errors* (ppb)
ARL	Ensemble_apriori_201602	0.41	0.72	38.8	56.6
	Ensemble_optimized_201602	0.77	0.93	2.37	22.6
	Ensemble_apriori_201604	0.47	0.48	10.6	29.1
	Ensemble_optimized_201604	0.69	0.83	1.09	14.8
	Ensemble_apriori_201606	0.71	0.59	-9.77	36.0
	Ensemble_optimized_201606	0.75	0.87	1.52	16.7
	Ensemble_apriori_201611	0.21	0.50	28.6	70.2
	Ensemble_optimized_201611	0.72	0.87	-6.54	37.2
HAL	Ensemble_apriori_201602	0.39	0.56	35.9	74.0
	Ensemble_optimized_201602	0.60	0.83	7.82	43.0
	Ensemble_apriori_201604	0.38	0.52	10.1	32.0
	Ensemble_optimized_201604	0.65	0.87	2.62	17.9
	Ensemble_apriori_201606	0.89	0.76	-10.1	30.1
	Ensemble_optimized_201606	0.80	0.89	4.27	17.6
	Ensemble_apriori_201611	0.13	0.34	48.2	124.5
	Ensemble_optimized_201611	0.47	0.79	2.48	78.5
NDC	Ensemble_apriori_201602	0.55	0.73	33.3	47.9
	Ensemble_optimized_201602	1.01	0.84	12.8	35.3
	Ensemble_apriori_201604	0.42	0.39	6.11	27.6
	Ensemble_optimized_201604	0.63	0.78	3.02	14.9
	Ensemble_apriori_201606	0.64	0.59	-7.04	34.8
	Ensemble_optimized_201606	0.73	0.88	4.06	17.5
	Ensemble_apriori_201611	0.25	0.47	28.5	72.8
	Ensemble_optimized_201611	0.70	0.91	1.50	33.8
BUC	Ensemble_apriori_201602	0.44	0.79	23.0	32.1
	Ensemble_optimized_201602	0.91	0.98	-0.39	7.07
	Ensemble_apriori_201604	0.92	0.78	5.98	14.8
	Ensemble_optimized_201604	0.95	0.95	-0.38	5.67
	Ensemble_apriori_201606	0.93	0.47	-19.8	43.1
	Ensemble_optimized_201606	0.89	0.91	0.22	9.13
	Ensemble_apriori_201611	0.32	0.49	10.4	40.5
	Ensemble_optimized_201611	0.81	0.95	-6.55	15.8

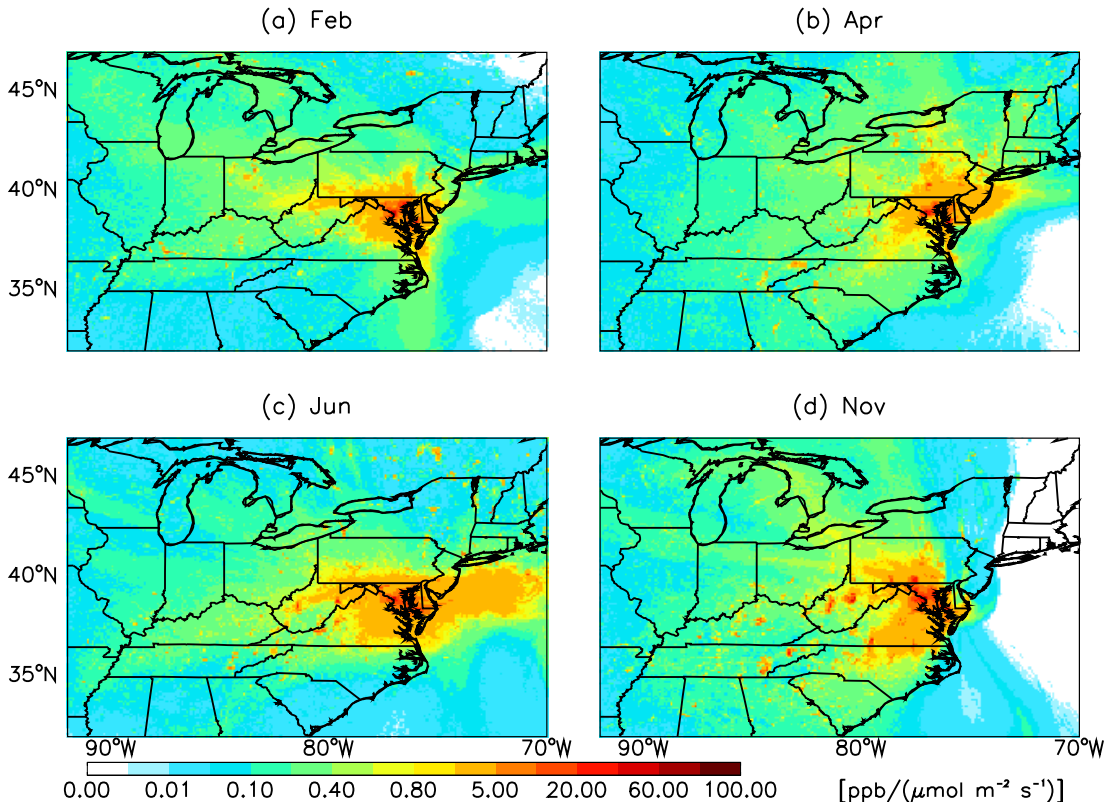

Figure S1. Ensemble mean of footprints from different meteorological products (HRRR, NARR, and GDAS) for the four in-situ tall towers (ARL, HAL, NDC and BUC) over D01 domain for (a) Feb., (b) Apr., (c) June and (d) Nov. 2016, respectively. Please refer to Figure 1 for the location of the four tall towers.

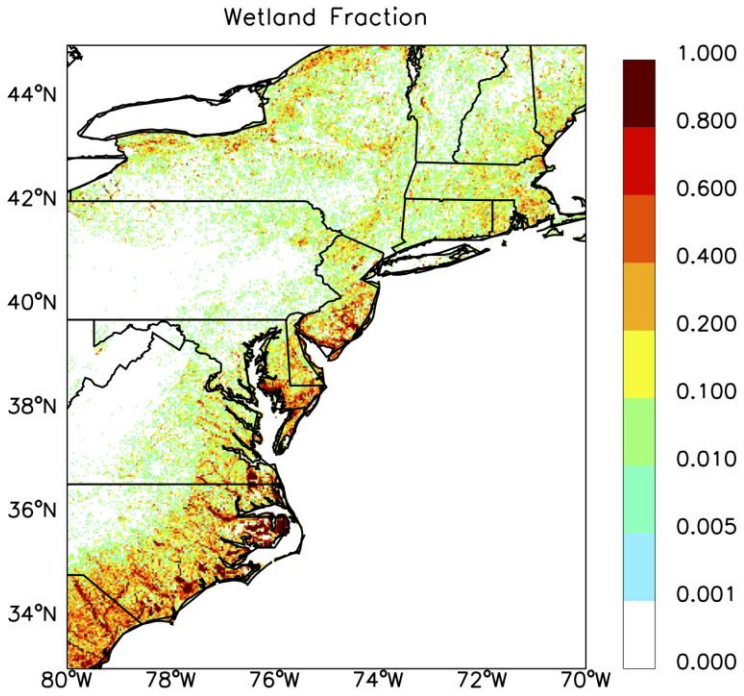

Figure S2. Wetland fractions from National Land Cover Database for the year 2011 over the D02 domain.

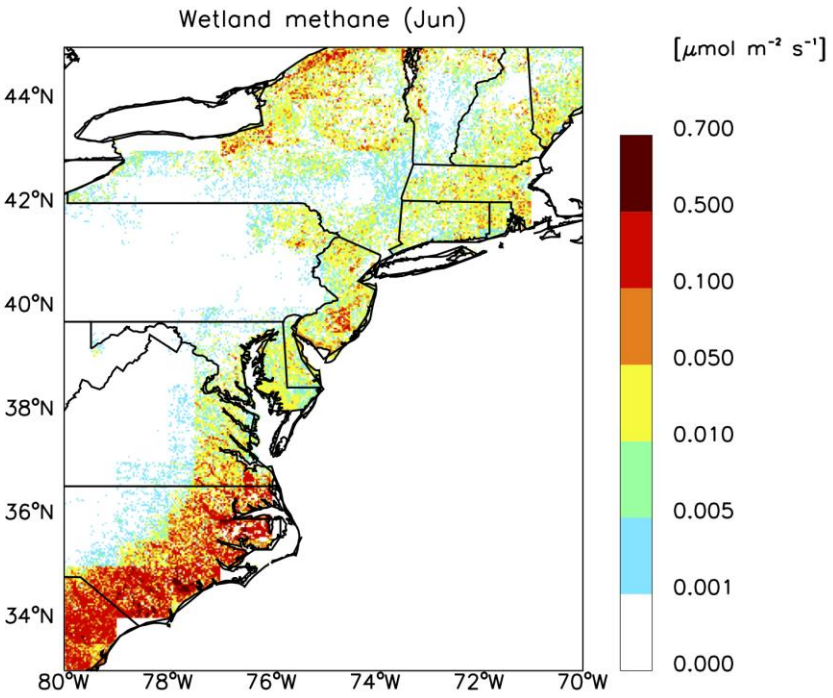

Figure S3. Final re-gridded wetland CH_4 emissions over the D02 domain.

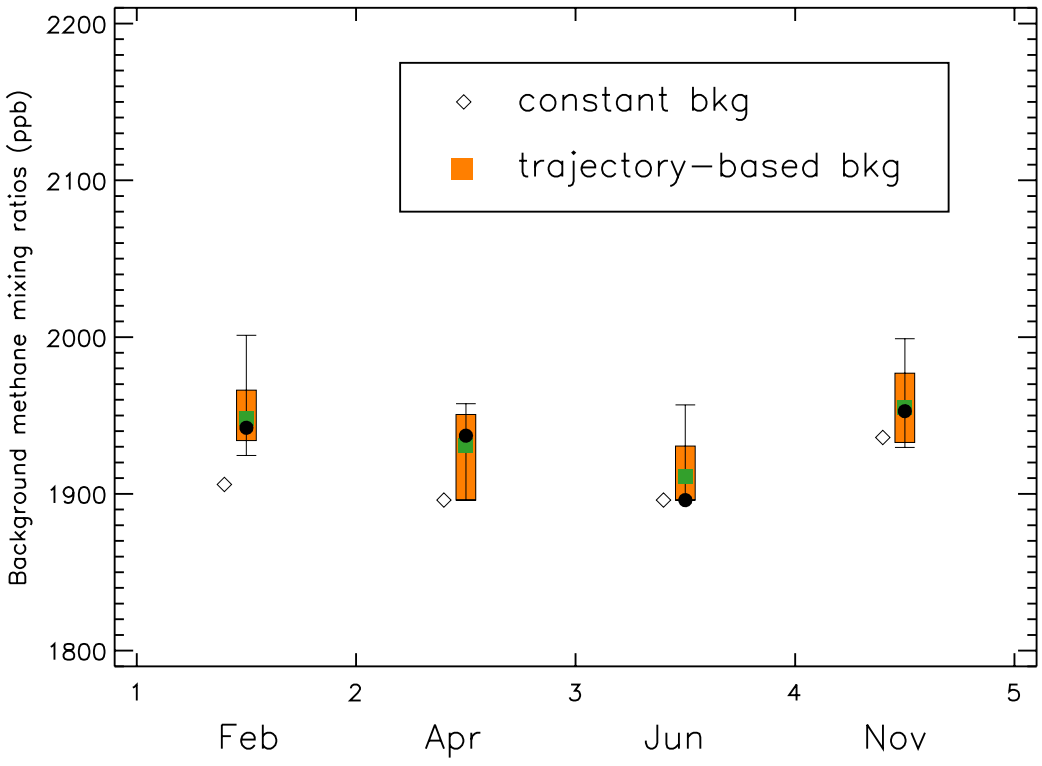

Figure S4. Box plots of trajectory-based background CH_4 concentrations for Feb., Apr., June and Nov. 2016, with thick (thin) bars representing 67% (95%) percentiles of the data for each month. Median and mean values of the trajectory-based background CH_4 concentrations in each month are shown in filled black circles and green rectangles respectively. Open diamonds next to the box plots denote the constant background CH_4 concentrations in each month. Note that we use the hybrid trajectory-based method for the months of Apr. and June.

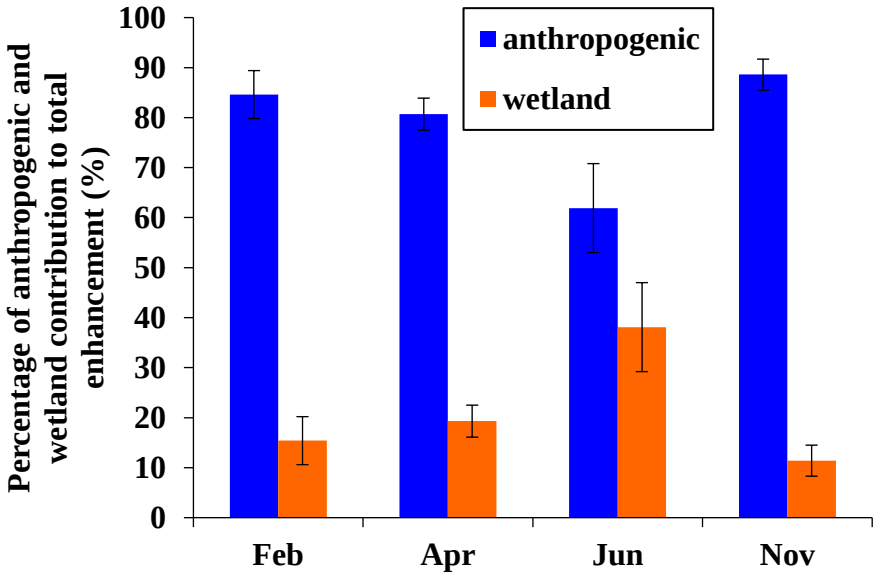

Figure S5. Ensemble fractional contributions of anthropogenic (blue) and wetland (orange) emissions to total enhancements from D01 fluxes in each month. Anthropogenic emission inventory here is based on EPA for the year 2012. Ensemble mean in each month is calculated as the average across each site and each met product. Error bars represent 1 standard deviation from all ensemble members in each month.

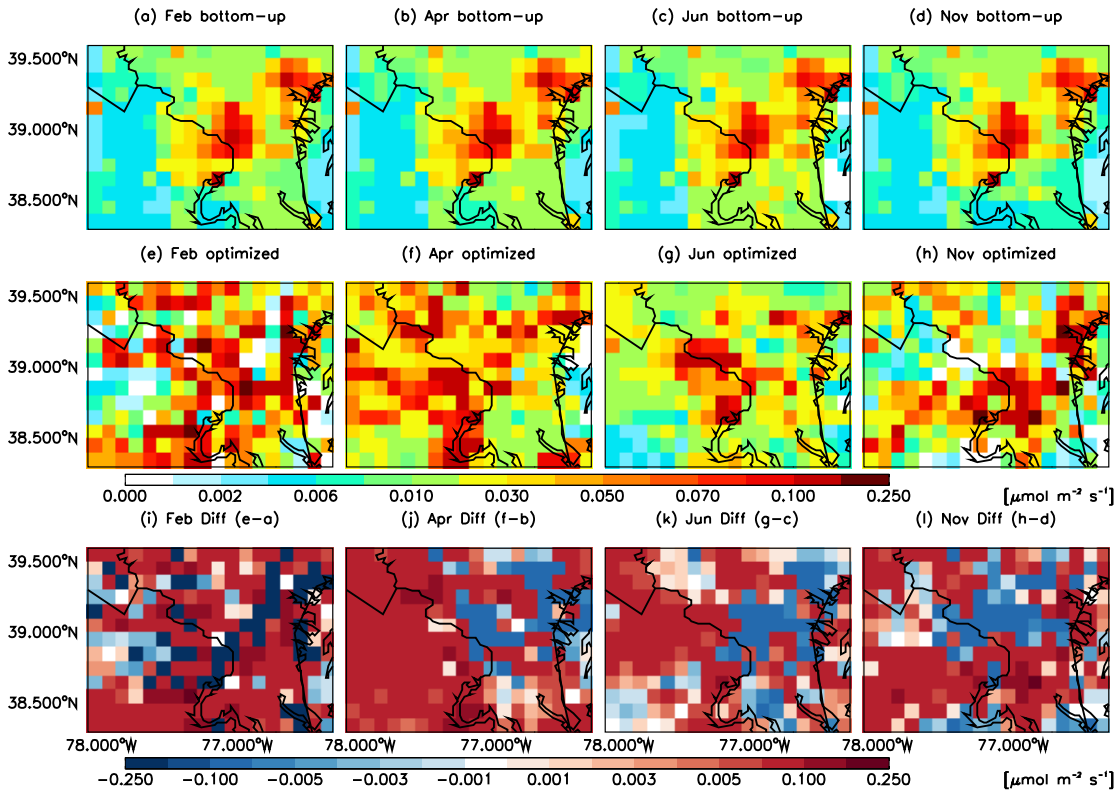

Figure S6. Spatial distributions of ensemble bottom-up (first row), optimized (second row) and the differences (optimized minus bottom-up) between ensemble optimized and bottom-up (last row) CH_4 emissions over the D03 domain for Feb. (a, e, i), Apr. (b, f, j), June (c, g, k) and Nov. (d, h, l) 2016 respectively.

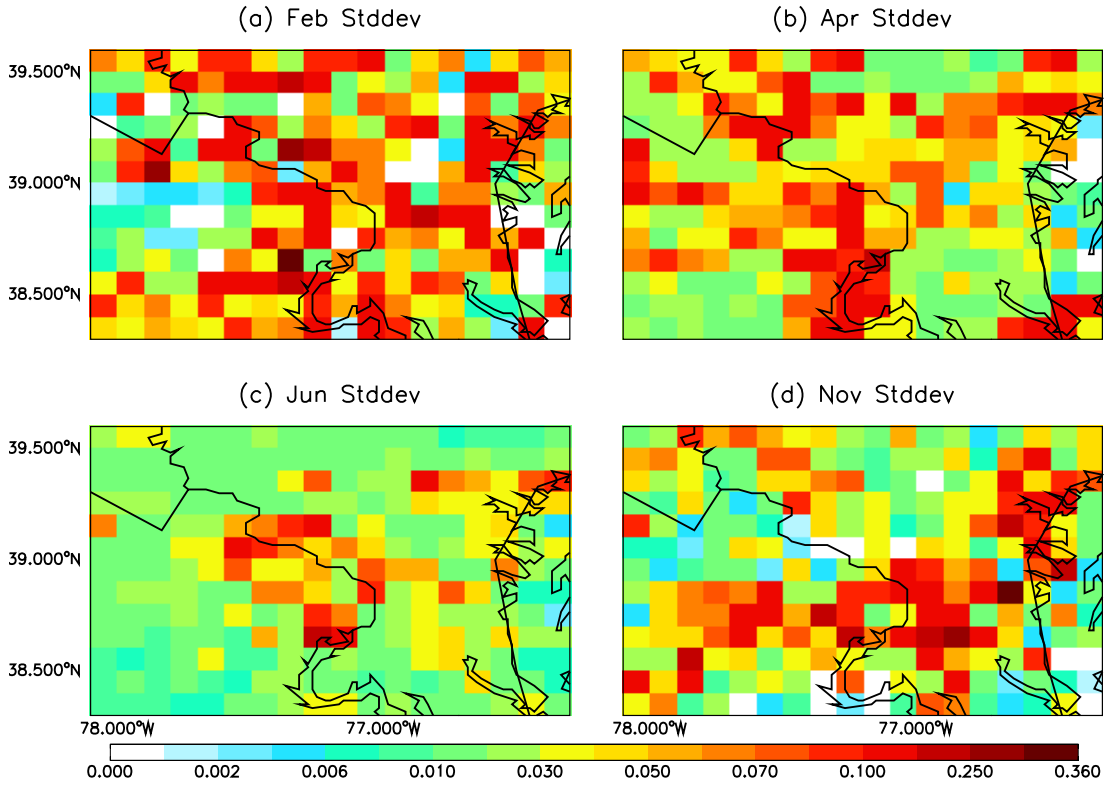

Figure S7. Standard deviations of ensemble spatial posterior CH_4 emission fluxes (μmol m^{-2} s^{-1}) over D03 domain for Feb. (a), Apr. (b), June (c) and Nov. (d) 2016, respectively.

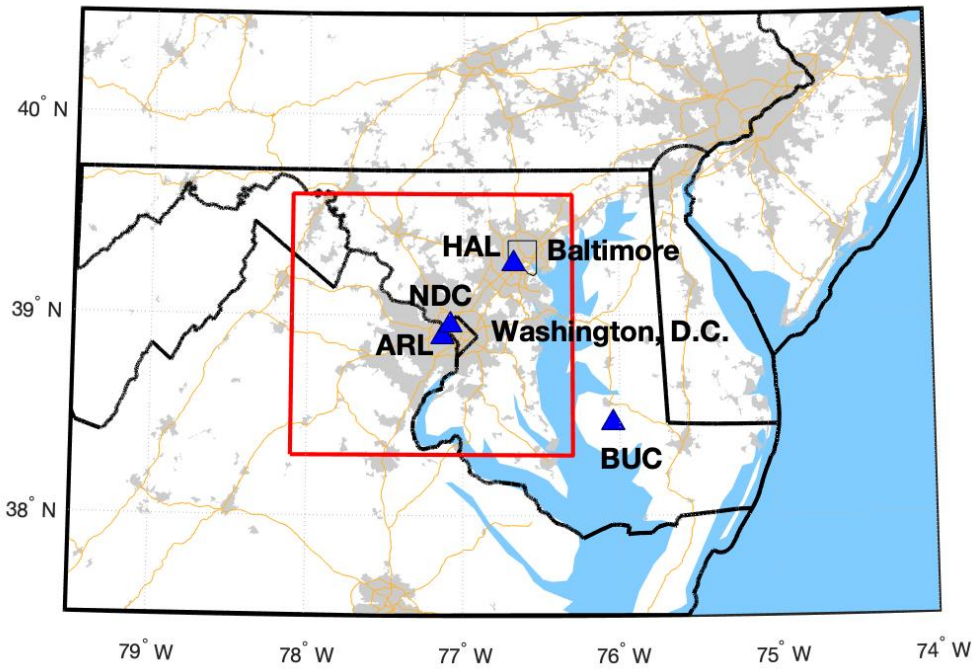

Figure S8. The definition of D03 domain (red box) in our study, with urban regions and highways shown in grey and orange respectively. Cities of Baltimore and Washing, DC and four in-situ towers (blue triangles; HAL, NDC, ARL, and BUC) are inserted in the plot.

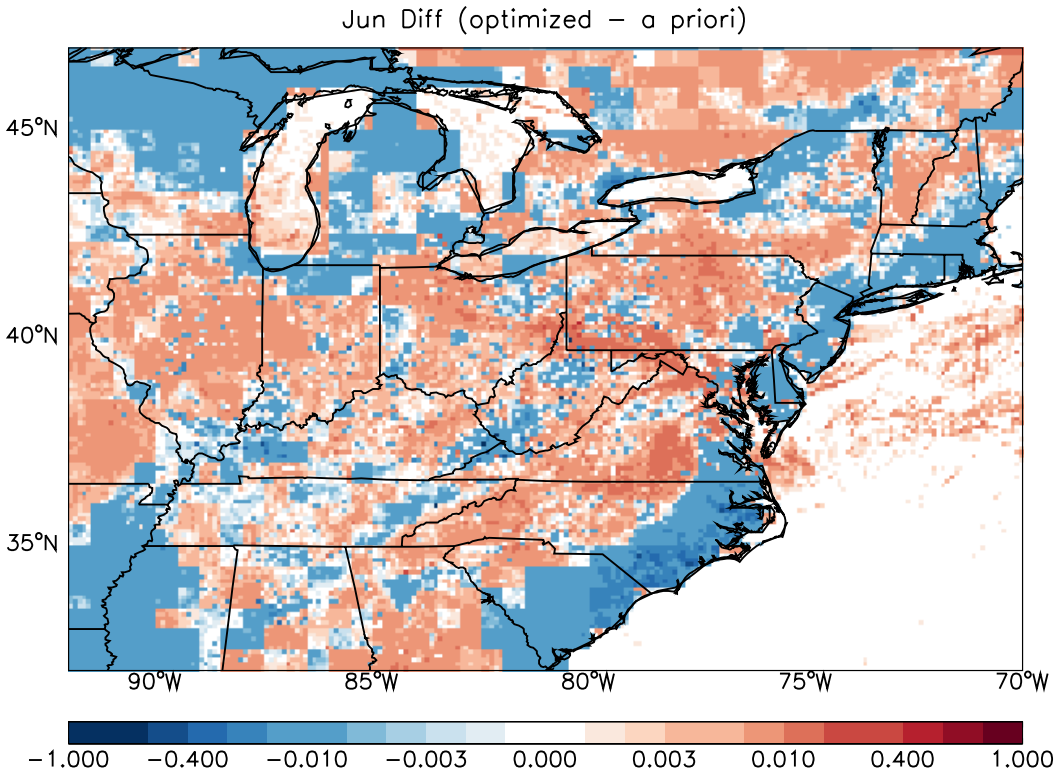

Figure S9. Spatial differences between ensemble optimized and a priori CH_4 emission fluxes over the D01 domain for June 2016. Color bar units: μmol m^{-2} s^{-1}.

References

(1) Wickham, J.; Stehman, S. V.; Homer, C. G. Spatial Patterns of the United States National Land Cover Dataset (NLCD) Land-Cover Change Thematic Accuracy (2001-2011). *Int. J. Remote Sens.* **2018**, *39* (6), 1729–1743.

(2) Michalak, A. M.; Hirsch, A.; Bruhwiler, L.; Gurney, K. R.; Peters, W.; Tans, P. P. Maximum Likelihood Estimation of Covariance Parameters for Bayesian Atmospheric Trace Gas Surface Flux Inversions. *J. Geophys. Res. Atmos.* **2005**, *110* (24), 1–16.

(3) Mueller, K. L.; Gourdji, S. M.; Michalak, A. M. Global Monthly Averaged CO2 Fluxes Recovered Using a Geostatistical Inverse Modeling Approach: 1. Results Using Atmospheric Measurements. *J. Geophys. Res. Atmos.* **2008**, *113* (21), 1–15.