*Journal of Geophysical Research - Atmosphere*

Supporting Information for

# Fluxes of Atmospheric Greenhouse-Gases in Maryland (FLAGG-MD): Emissions of Carbon dioxide in the Baltimore-Washington area

D. Y. Ahn[1], J. R. Hansford[2], S. T. Howe[3⊥], X. R. Ren[3,4,5], R. J. Salawitch[1,3,4], N. Zeng[3,4], M. D. Cohen[5], B. Stunder[5], O. E. Salmon[6*], P. B. Shepson[6,7], K. R. Gurney[8], T. Oda[9,10], I. Lopez-Coto[11], J. Whetstone[12], R. R. Dickerson[3]

[1]Department of Chemistry and Biochemistry, University of Maryland College Park, Maryland, USA, [2]Department of Computer Science, University of Maryland College Park, MD, USA, [3]Department of Atmospheric and Oceanic Science, University of Maryland College Park, MD, USA, [4]Earth System Science Interdisciplinary Center, University of Maryland College Park, MD, USA, [5]National Oceanic and Atmospheric Administration Air Resource Laboratory, College Park, MD, USA, [6]Department of Chemistry, Purdue University, West Lafayette, IN, USA, [7]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA, [8]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA, [9]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA, [10]Goddard Earth Sciences Research and Technology, Universities Space Research Association, Columbia, MD, USA, [11]Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA, [12]Special Programs Office, National Institute of Standards and Technology, Gaithersburg, MD, USA, ⊥Now at Areté, Arlington, VA, USA, *Now at Lake Michigan Air Directors Consortium, Madison, WI, USA

**Contents of this file**

Text S1 to S6
Figures S1 to S11
Tables S1

**Text S1. Wind bias detection**

The existence of the heading-dependent bias in the wind speed measured by the Garmin G600 system onboard the UMD Cessna aircraft was first identified by colleagues at the Pennsylvania State University (Ren et al., 2019). To address this issue, a series of calibration flights were conducted in October 2017 with the same UMD Cessna aircraft used for the flights in February 2015. For these calibration flights, the Cessna aircraft was equipped with both the original Garmin system and a newly installed differential GPS (DGPS) system, which measures aircraft true heading precisely with an accuracy of 0.05º. Figure S7a shows that the aircraft heading measured by the original Garmin system has a cosine-shaped systematic bias with respect to the aircraft heading measured by the DGPS system. The cosine-shaped bias in the Garmin heading measurement implies the existence of a hard-iron effect during the October 2017 flights: i.e., the permanent magnetic field that exists in the aircraft vessel interferes with the magnetometer's reading of the Earth's magnetic field.

For the February 2015 flights, neither DGPS data nor other kinds of records exist that could be used to directly quantify the magnitude of the hard-iron effect on the Garmin heading. However, the difference between measured wind speed and output of the NAM4 model as a function of aircraft heading can be analyzed to qualitatively show the existence of the hard iron effect during the February 2015 flights (Figure S7b). The 'W' shaped pattern in Figure S7b, where the smallest differences of the wind speed were found near 90º and 270º and the largest differences were found near 0º, 180º, and 360º, demonstrates the existence of a hard iron effect during the mass balance flights conducted in February 2015.

**Text S2. Wind bias correction**

The original wind data measured by the Garmin system during February 2015, which we call hereafter version 1 (v1) wind, include an error induced by the systematic bias in the aircraft heading reported by the Garmin G600 system (Figure S7). From the original v1 wind data, the v2 wind field (speed and direction) was derived by correcting the systematic heading bias. Then, the v3 wind field in which the accuracy is further improved was derived utilizing NAM4. Table S1 summarizes the differences in the wind speed perpendicular to the aircraft heading for these three versions of the wind field. The following paragraphs describe the method used to derive the v2 and v3 wind fields.

First, the systematic heading bias for the February 2015 flights data was corrected using the fourth degree polynomial function given at the top of Figure S7a, which was obtained from the calibration flights conducted in October 2017. Following the correction of the heading bias, calibration coefficients of +0.8º and +1.3% were applied to the headings and the true air speeds (TAS), respectively (see Supplement of Ren et al., 2019). Then, the v2 wind speed and wind direction were calculated based upon the bias-corrected/calibrated headings and TAS measurements, along with the original records of ground speed (GS) and true track angle (TTA).

Even after the bias correction and the calibration of the heading and TAS measurements, a systematic bias could still be present in the v2 wind if the magnitude and direction of the hard-iron effect in February 2015 was significantly different from that in October 2017. The same aircraft had been used for both flight months; there is no record of how internal aircraft electronics and support structures may have changed. To address this issue and further improve the accuracy of the v2 wind, NAM4 model wind was used to calculate v3 wind data in the following manner. For the downwind transects measurements for each flight, 10 second running means of the perpendicular wind speed were calculated from v2 wind and from NAM4 wind, respectively ($U^{V2}_{\perp,x,z}$ and $U^{NAM4}_{\perp,x,z}$), as shown in Figure S9. From the two sets of perpendicular wind speed, the mean difference ($\overline{U^{NAM4}_{\perp} - U^{V2}_{\perp}}$) was calculated. Then, v3 perpendicular wind speed was calculated by adding the mean difference to the v2 perpendicular wind speed during the downwind transects: i.e., $U^{V3}_{\perp,x,z} = U^{V2}_{\perp x,z} + \overline{U^{NAM4}_{\perp} - U^{V2}_{\perp}}$

**Text S3. Wind evaluation**

For the evaluation of the series of the aircraft wind correction procedures described in Text S2, three analyses were conducted. First, wind profiler data was used to evaluate the accuracy and the precision of the NAM4 wind data that is a factor in the derivation of the v3 aircraft wind data. Second, the NAM4 wind was used to assess the variations among the three versions of the aircraft winds. Finally, the Continuous Emissions Monitoring System (CEMS) measurement of $CO_2$ emissions from power plants was utilized to evaluate the accuracy of the three versions of the aircraft wind field.

Figure S8 shows a comparison of four variables (wind speed, wind direction, U and V components of the horizontal wind) between the NAM4 and the wind profiler observations at the Beltsville, Maryland site on 8 days in February 2015. An excellent correlation is found between the NAM4 and profiler data for each of these four wind components, without any noticeable systematic bias. The mean difference of wind speed between the NAM4 and the profiler was found to be 0.2 m/s, which translates into a 2.6 % uncertainty in the $CO_2$ flux estimation.

Figures S9 shows a comparison of the perpendicular wind speed derived from the NAM4 versus that derived from the three versions of the aircraft wind field, for flight MD RF4 conduction on 19 February 2015. The original v1 perpendicular wind speed was found to be consistently faster than the value from NAM4. The v2 wind field (i.e., correct for the heading bias) caused the shape of $U^{V2}_{\perp,x,z}$versus time to change, because the aircraft heading varied as a function of time. The shape of the v2 wind as a function of time agrees more closely with the shape of the NAM4 perpendicular wind field. However, the v2 perpendicular wind speed was consistently slower compared to NAM4. The v3 perpendicular wind speed (found as described in Text S2) shows excellent agreement with the NAM4 wind speed, retaining the same shape versus time as the v2 wind. Table S1 documents the root mean square error (RMSE) between the NAM4 perpendicular wind field and the three versions of the aircraft wind field. The v3 wind field displays either the smallest (all flights except UMD RF6) or nearly the smallest (UMD RF6) value of RMSE relative to the NAM4 perpendicular wind.

Figure S10 shows a comparison between the emission rate of $CO_2$ for two local power plants, Chalk Point (CP) and Morgantown (MT), from the CEMS record (see Figure 5 of the main paper for a detailed description) versus the emission rate of $CO_2$ derived from the v1, v2, and v3 wind fields. Three quantitative metrics for the comparison; i.e., MPE (mean percentage error), MAPE (mean absolute percentage error), and a linear regression of our computed $CO_2$ emission versus the CEMS value all indicate that the v3 wind field provides the most accurate estimate of the emission rate of $CO_2$ for the two local power plants. Figure S10c is similar to Figure 6 in the main text, except Figure 6 in the main text also includes data from Purdue RF3 and Purdue RF4.

## Text S4. Uncertainty in the emission rate of $CO_2$ from CEMS records

The uncertainty for CEMS $CO_2$ emissions ($\sigma_{CEMS}$) in Figure 6b was determined by combing three independent sources of uncertainty in a root mean sum of error fashion: 1) uncertainty in CEMS records based on the RATA performance specification ($\sigma_{CEMS,\ RATA}$), 2) the difference of CEMS records against fuel-consumption based EIA datasets ($\sigma_{CEMS,\ EIA}$), and 3) the uncertainty in the air transport time (between the power plant and aircraft) estimated using HYSPLIT back trajectories ($\sigma_{CEMS,\ Transport}$).

First, $\sigma_{CEMS,\ RATA}$ is determined based upon the main performance specification values described in the Relative Accuracy Test Audit (RATA). The RATA is the periodical comparison test of CEMS against the concurrent measurements made by the EPA reference method (U.S. Environmental Protection Agency, 2009). The value of $\sigma_{CEMS,\ RATA}$ was found by propagating the relative accuracy of 10% for concentration and volumetric flow rate measurements into the $CO_2$ mass emission rate calculation equation shown in Table 6 of USEPA (U.S. Environmental Protection Agency, 2009).

Second, $\sigma_{CEMS,\ EIA}$ was considered because Gurney et al. (2016) found that monthly $CO_2$ emissions in facility CEMS records (stack measurements based estimates) differ by more than ±13% compared to those in EIA datasets (fuel consumption based estimates) for about one-fifth of U.S. power plants. Quick & Maryland (2019) identified and corrected systematic errors in either the U.S. EPA CAMD (Clean Air Markets Division) or the U.S. EIA (Energy Information Administration) datasets (i.e., unreported unit emissions in the CAMD dataset and emission factor error in the EIA dataset). We confirmed from Quick & Marland (2019) that the CAMD dataset for the CP and MT power plants are not affected by unreported unit emissions. Further, we compared CO2 emissions for CP and MT from the CAMD datase against corresponding EIA data for February 2015. For the CP power plant, the monthly $CO_2$ emission for Feb 2015 in CAMD is 4% greater than in EIA. For the MT power plant, the emission for Feb 2015 given by CAMD is 8% lower compared than that provided by EIA. While such differences could be caused by errors in either the CAMD or EIA estimate, we used our computed difference values of – 4% and 8% as $\sigma_{CEMS,\ EIA}$ for the CP and MT power plants, respectively.

Finally, the value of $\sigma_{CEMS,\ Transport}$ was determined as the standard deviation of the CEMS hourly $CO_2$ emissions within ± 1 hour (i.e., 3 hours span) from our baseline estimate of the transport time from the power plant stack to the aircraft. The baseline plume transport time was estimated using HYSPLIT back trajectories run with NAM12 meteorology.

## Text S5. Emissions of $CO_2$ from human and pet respiration and NFA-$CO_2$ sources and uncertainty propagation

To estimate emissions of $CO_2$ from respiration by humans and pets, we adopted a similar approach to Gurney et al. (2017). A value of 254 gC/person/day was used as the average $CO_2$ release rate by human respiration (Prairie & Duarte, 2007). The population of the Balt-

Wash area for 2015 was estimated as 8,153,000 based on GPWv4 (Gridded Population of the World) data, as described in the main text. Emissions of $CO_2$ from dog and cat respiration were also estimated assuming that the study area follows the average U.S. per capita ownership of 0.22 dogs/person and 0.24 cats/person, and a dog/cat release rate of $CO_2$ of 25% of the human release rate (American Veterinary Medical Association, 2012).

Once the human/pet respiration estimate for the emission of $CO_2$ (~2,000 mol/s) was subtracted from the mass balance estimate for each flight, 4.7% of the remaining $CO_2$ mass balance emission estimate was apportioned to anthropogenic sources other than the combustion of fossil fuel (i.e., Non-Fossil fuel Anthropogenic $CO_2$, or NFA-$CO_2$). According to the MDE GHG inventory, 4.7% of the total in-state emissions of $CO_2$ are from the following sectors: 1) industrial processes (cement manufacture, limestone and dolomite, soda ash, ammonia and urea production), 2) agriculture (urea fertilizer usage), 3) waste management (waste combustion, landfills, and residential open burning) (MDE, 2016). The MDE estimates are based on annual emissions for 2014; the 4.7% value was adopted, unchanged, for February 2015.

The uncertainty range of the FLAGG-MD monthly total FF$CO_2$ estimate was determined by propagating four independent sources of uncertainty: 1) uncertainty in the mass balance estimate ($\sigma_{\text{mass-balance}}$), 2) uncertainty in the human/pet respiration estimate ($\sigma_{\text{human/pet-respiration}}$), 3) uncertainty in the ratio of NFA-$CO_2$ to total $CO_2$ ($\sigma_{\text{NFA-CO2}}$), and 4) uncertainty in the temporal scaling factor used to relate our seven mass balance estimates to the monthly total emission of $CO_2$ ($\sigma_{\text{temporal-scaling}}$). First, $\sigma_{\text{mass-balance}}$ was determined from a Monte Carlo simulation by propagating the uncertainties of five parameters that enter the mass balance equation (See Table 2). Second, $\sigma_{\text{human/pet-respiration}}$ was specified to be ±30%, based on a conservative estimate in how local pet ownership might vary relative to the national averaged. Given the preponderance of dogs and cats in the Balt-Wash region and the lack of large-scale animal feedstock, emissions of CO2 from animals other than human, dog, and cat should be well covered by this ±30% value. Third, $\sigma_{\text{NFA-CO2}}$ was determined to be ±1.5%, based upon as the standard deviation of three NFA-$CO_2$ ratios derived from MDE GHG inventory for year 2006, 2011, and 2014. Finally, $\sigma_{\text{temporal-scaling}}$ was determined to be 0.4%, based upon the standard deviation of three temporal scaling factors from FFDASv2.2, TIMES, and ACESv1 (see section 3.5.3).

## Text S6. Bottom-up gridded emissions products: Discrepancies and harmonizing efforts

FFDASv2.2 consists of the downscaled IEA estimate of fossil fuel combustion emissions and the EDGAR (Emissions Database for Global Atmospheric Research) version 4.3.2 estimate of aviation and shipping emissions. FFDASv2.2 data files did not provide any sector specific emissions. In Figure 9, the FF$CO_2$ value from FFDASv2.2 was directly derived from hourly NetCDF data files available at http://ffdas.rc.nau.edu.

EDGARv432 monthly data for year 2010 consists of source sectors specified by IPCC, as detailed in Table S4 of Janssens-Maenhout et al., (Janssens-Maenhout et al., 2017). In Figure 9, the FF$CO_2$ value of EDGARv432 consists of the following sectors: Power Industry, Energy for Buildings, Combustion for Manufacturing Industry, Road Transportation, Aviation (landing & take off, climbing & descending, and cruise), Shipping and Railways, Pipelines, and Off-Road Transport. The FF$CO_2$ value was calculated solely from the long cycle C (file name: "CO2_excl_short-cycle_org_C") to be consistent with our other estimates of $FFCO_2$. In Figure 9, the "ELEC" label of the EDGARv432 indicates emissions from the Power Industry sector. The "RCI" label consists of the Energy for Buildings and the Combustion for Manufacturing Industry sectors. The "Onroad" label indicates the Road Transportation sector, and the "Nonroad" label consists of emissions from the Aviation, Shipping, and Off-Road Transport sectors.

The ACESv1 data for year 2014 consist of emissions from the following sectors: Residential, On-Road Transportation, Oil and Gas Production, Off-Road Vehicles/Marine/Rail, Non-Electricity Generating Facilities, Electricity Generating Facilities, Airport, and Industrial and Commercial. In Figure 9, the $FFCO_2$ value of ACESv1 consists of all of the sectors listed above. The "ELEC" label for ACESv1 denotes emissions from the Electricity Generating Facilities sector. The "RCI" label consists of the Residential, Industrial and Commercial, and Non-Electricity Generating Facilities sectors. The "Onroad" label indicates emissions from On-Road Transportation, whereas the "Nonroad" label combines emissions from the Airport and the Off-Road Vehicles/Marine/Rail sectors. The total emissions of CO2 for 2014 from ACESv1 are held constant to that for their year 2011 analysis, but re-distributed based on variations in meteorology, fuel consumption, and traffic patterns between these two years (Gately & Hutyra, 2018).

ODIAC2018 data consists of two emission categories: emissions over land (variable name: "land") and emissions from international aviation and marine bunkers (variable name: "bunker"). The land sector consists of emissions from fossil-fuel combustion, cement manufacturing, and gas flaring. The bunker sector was only available on a 1×1° lat/lon grid provided via NetCDF data files, while the land sector was available on both 1×1 km spatial grid via GeoTIFF files and the 1×1° grid via NetCDF files. In Figure 9, the $FFCO_2$ value from ODIAC2018 consists both land and bunker sectors. The land emissions were obtained from the 1×1 km data file. For bunker emissions, the ratio of bunker to land emissions for our study domain was calculated using data from both 1×1° files, and the ratio was multiplied by the land emissions computed using data from the 1×1 km file. In Figure 9, the "Nonroad" label for ODIAC2018 indicates emissions from the bunker sector. Note that the $FFCO_2$ values marked by the "Nonroad" label for ACESv1 and EDGARv432 consist of not only aviation and bunker emissions, but also the off-road vehicle and rail sectors.

The MDE GHG inventory for year 2014 Microsoft Excel data file consists of various sources sectors (including imported electricity) and sinks of GHG. The state-wide annual total $FFCO_2$ was calculated as the sum of emissions from following sectors: In-state Energy Production (coal, natural gas, and oil), Residential/Commercial/Industrial Fuel Use (coal, natural gas & LPG, petroleum), Transportation (on-road gasoline & diesel, nonroad gasoline & diesel, rail, marine vessels, lubricants & natural gas & LPG, and jet fuel & aviation gasoline), and Fossil-Fuel Industry (natural gas industry). Emissions from the following sectors were summed to calculate NFA-$CO_2$ (Non-Fossil fuel Anthropogenic $CO_2$): industrial processes (cement manufacture, limestone & dolomite, soda ash, and ammonia & urea production), agriculture (urea fertilizer usage), and waste management (waste combustion, landfills, and residential open burning).

Several sector mismatches exist for $FFCO_2$ derived from the five bottom-up inventory datasets. First, FFDASv2.2 does not cover the cement manufacturing and gas flaring sectors (CM&GF). The EDGARv432 and MDE inventories cover CM&GF, but we excluded these sectors when calculating $FFCO_2$. The ACESv1 and ODIAC2018 datasets cover CM&GF, but these two sectors could not be isolated from other $FFCO_2$ sectors in the data files provided by these two groups. Therefore, emissions of CO2 from the CM&GF sectors remain the bottom-up inventories from ACESv1 and ODIAC2018. According to the MDE inventory, the CM&GF sectors emitted 0.4 MtC during year 2014, which is about 2% of the state-wide annual total $FFCO_2$ estimate.

Second, EDGARv432, FFDASv2.2, and ODIAC2018 cover both the aircraft landing & takeoff sector as well as the airborne aircraft emissions sector, while ACESv1 only covers the aircraft landing & takeoff sector. Note that the aircraft emissions sector of FFDASv2.2 was directly adopted from EDGAR. The MDE inventory estimate of aviation emissions was based on aviation fuel consumption statistics, and thus does not necessarily indicate emission within the geographical boundary of the state. According to EDGARv432, airborne aircraft emissions ("TNR_Aviation_CDS/CRS") emitted 0.05 MtC during February 2010, which is again about 2% of the monthly total $FFCO_2$ estimate. According to the MDE inventory, emissions from the jet fuel & aviation gasoline usage constitute about 1% of the state-wide annual total $FFCO_2$ emission inventory.

Finally, FFDASv2.2 does not cover emissions from the oil and natural gas refining and transformation sectors. Emissions provide by ODIAC2018 and ACESv1 do cover these sectors. Emissions of $CO_2$ from oil and natural gas refining and transformation could not be isolated from emissions of $CO_2$ from the more dominant combustion sectors for ODIAC2018, whereas according to ACESv1 there was no $CO_2$ emitted from these oil and gas sectors in our study domain.

EDGARv432's oil refineries and transformation industry sector (file name: "REF_TRF") and fuel exploitation sector (file name: "PRO") denote emissions from these oil and gas sector; non-combustion emissions of $CO_2$ are also provided in these files. Since FFFASv2.2 does not cover the oil and gas sector emissions provided by these "REF_TRF" and "PRO" files of EDGAR, these emissions were excluded from $FFCO_2$ of EDGARv432 shown in Figure 9. The MDE inventory does include emissions from pipeline fuel combustion within the natural gas industry sector; these emissions are included in the calculation of $FFCO_2$ from MDE discussed in section 3.5.3. According to MDE, only 0.0001 MtC of $CO_2$ was emitted from the oil and gas sector in year 2014 (including pipeline fuel combustion), which is less than 0.001 % of the total annual value of $FFCO_2$.

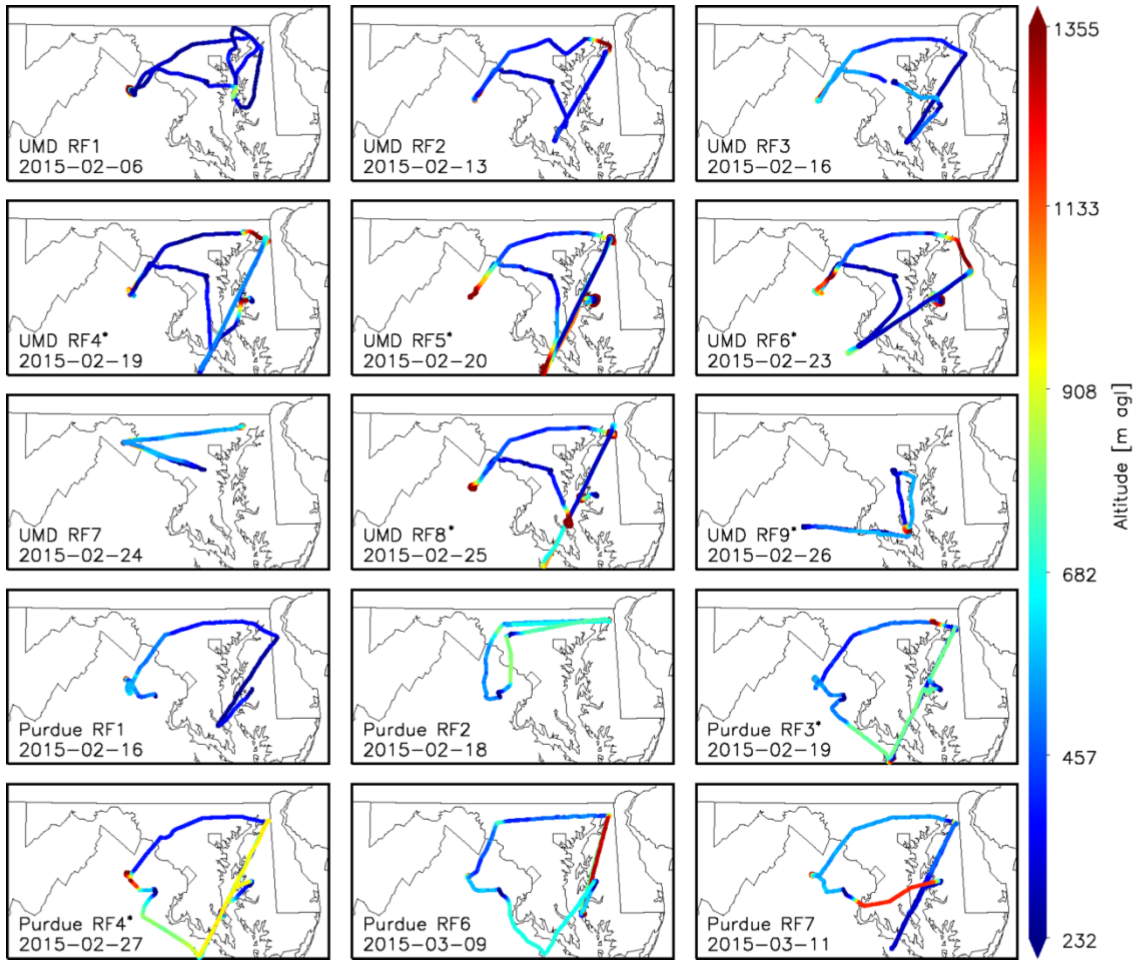


**Figure S1.** Flight tracks of the 15 research flights conducted during the winter 2015 FLAGG-MD campaign. A total of nine flights were conducted by the UMD aircraft and six flights were conducted by the Purdue aircraft. The date of each research flight is shown at the bottom left of each panel, in a year-month-day format. The asterisk (*) symbol next to each RF number indicates that in-situ data of that flight was used for the mass balance estimate of the emission of $CO_2$ from the Balt-Wash area within our study.

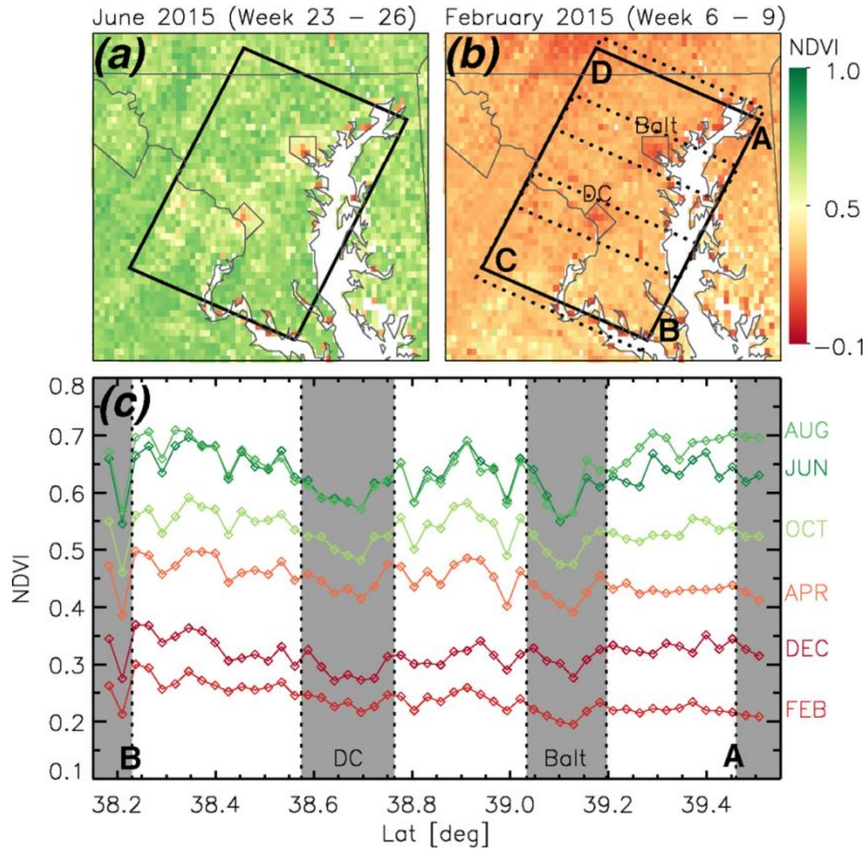


**Figure S2.** Normalized Difference Vegetation Index (NDVI) for (a) June 2015 and (b) February 2015. The v1r12 weekly NDVI data on a 4 km × 4 km grid from the Visible Infrared Imaging Radiometer Suite (VIIRS) is available from the following link: https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/index.php. Only data for February 2015 are used in the analysis; measurements for June 2015 are shown to illustrate that the VIIRS determination of NDVI is more sensitive to the rural/urban setting during summer than winter. Points A, B, C, and D as well as the rectangular box denoting our study area are the same as used in Figure 1 of the main text. (c) Averages of NDVI along a series of diagonal boxes that extend from just south point B and just north of point A on panel (b), plotted as a function of the middle latitude of each box along line AB (called as "horizontal transect" in the main text). The most southerly box and the most northerly box correspond to "edge areas" used to define background $CO_2$ for six of the seven mass balance flights. The latitudinal span of these boxes, as well as the latitudinal span of Washington, D.C. (DC) and the city of Baltimore (Balt), are shown by the grey shaded regions. Results for NDVI are shown for six months in 2015, as indicated. The slight decline in NDVI for DC and Balt for Feb 2015 is used to scale the results of the biogenic emission of $CO_2$ computed by the VEGAS model (see main text).

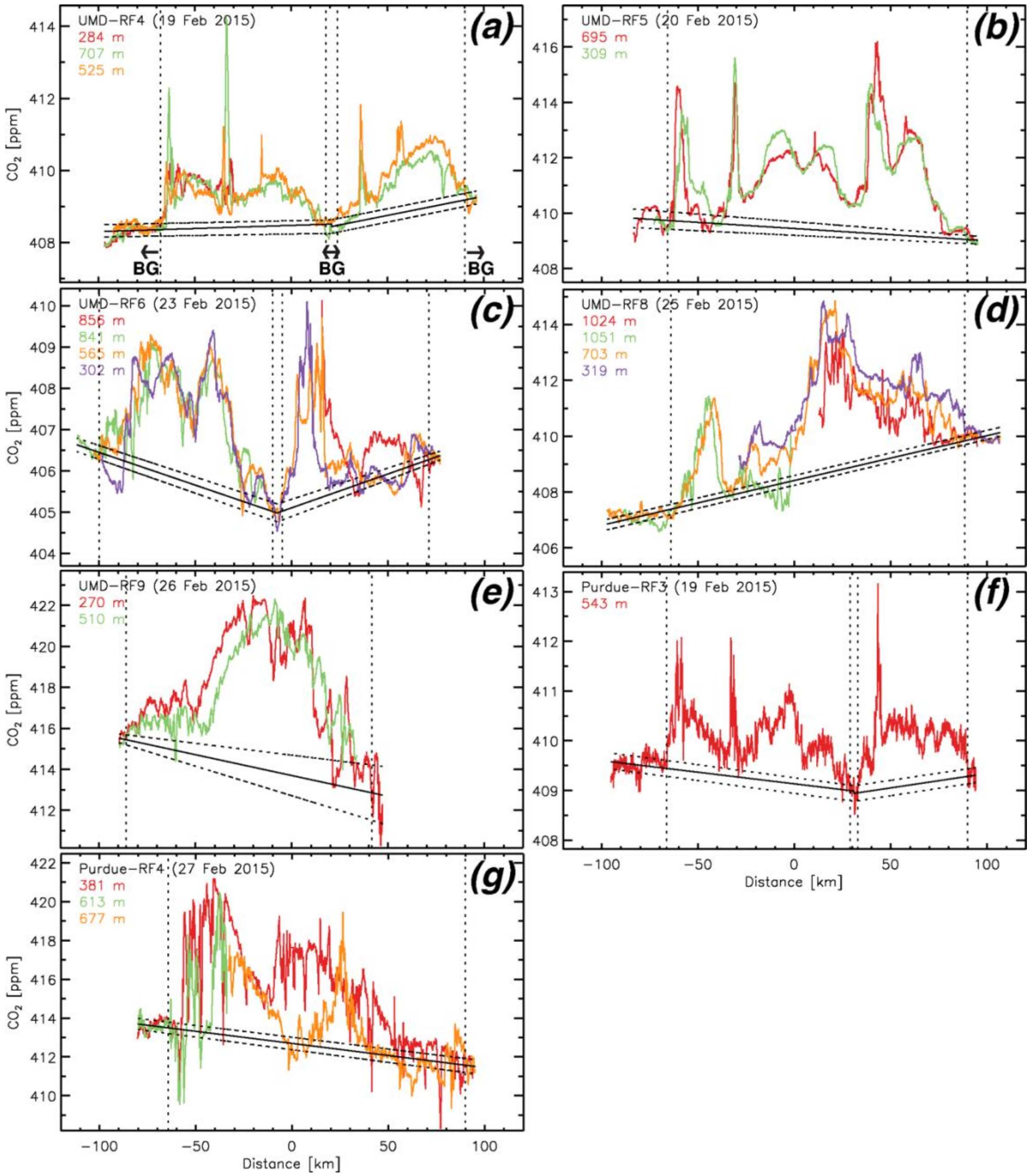


**Figure S3.** (a-g) Mole fraction of $CO_2$ measured downwind of the Balt-Wash area (colored) and the background $CO_2$ (black, solid) for the seven mass balance flights. Each colored line indicates downwind horizontal transects at different altitudes. The flight date and the mean altitude of each horizontal transect is shown at the left-top of each panel. The black solid lines indicate background $CO_2$ used to estimate the emission rate of $CO_2$; the black dotted lines indicate the $\pm 1\sigma$ bound of background $CO_2$ used for the sensitivity analysis. Dotted vertical lines indicate the boundaries of flight segments used to define the values of background $CO_2$.

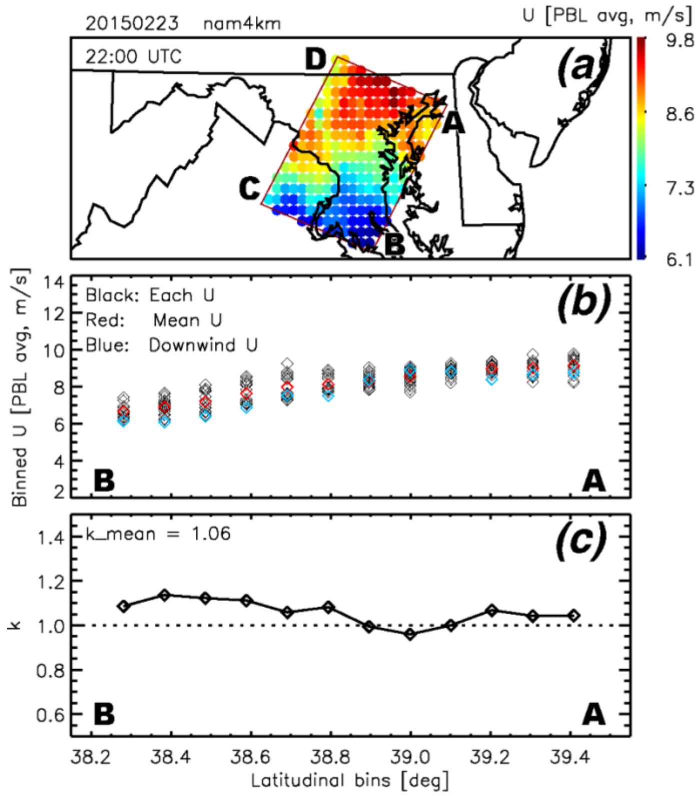


**Figure S4.** (a) Map showing the average value of U within the PBL ($\overline{U_{PBL}}$) derived from NAM4 for every cell on the 0.1°× 0.1° lat/lon grid. (b) Same $\overline{U_{PBL}}$ data shown in (a) but binned in 0.1° diagonal latitudinal bins (see Figure 7). For each diagonal latitudinal bin, the black diamonds indicate $\overline{U_{PBL}}$ from each grid 0.1°× 0.1° NAM4 grid point that lies within the bin. The red diamond indicates the mean value of $\overline{U_{PBL}}$ within the diagonal bin (i.e., the average of the black diamonds. The blue diamond indicates $\overline{U_{PBL}}$ for the NAM4 grid located closest to the downwind portion of the study area (i.e., line AB in Figure 1). (c) Black diamonds indicate the scaling factors *k* derived for each latitudinal bin, and the black line indicates the linearly interpolated scaling factor applied to wind measurements for the mass balance calculation.

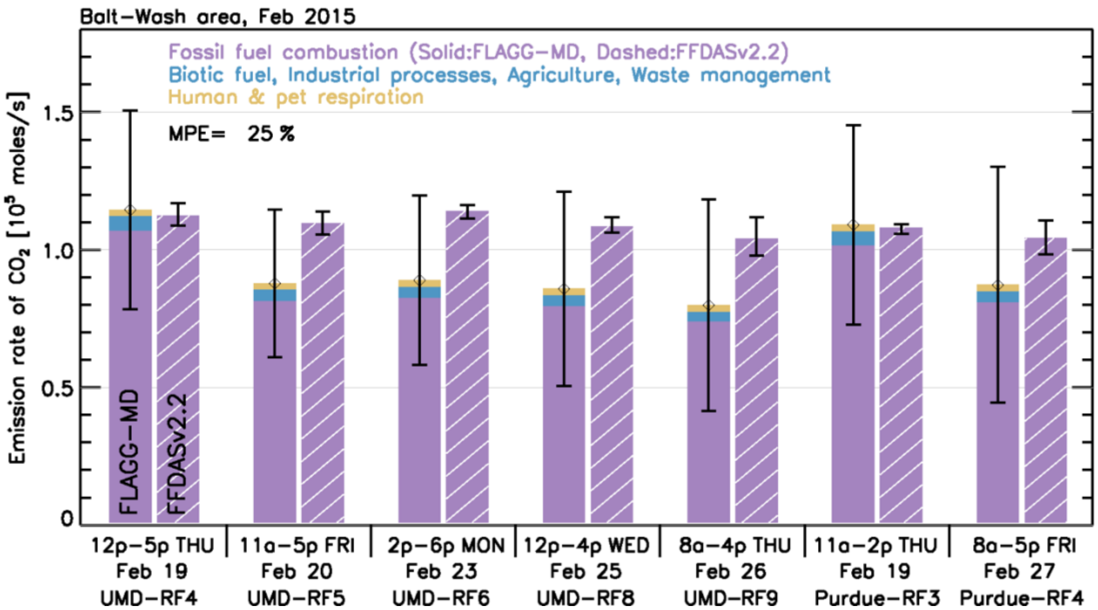


**Figure S5.** The emission rates of $CO_2$ from the Balt-Wash area during the sampling period for seven research flights conducted in February 2015. This figure is identical to Figure 8 of the main paper, except here we have computed the FLAGG-MD mass balance emissions assuming a value of unity for the scaling factor *k* described in section 2.5.2. In other words, here we assume the wind speed perpendicular to the aircraft flight track was steady during the transport time over the Balt-Wash area (i.e., $k = 1$ in Equation (1) of the main text). Overall, the FLAGG-MD fluxes shown here are 5% larger than those shown in Figure 8.

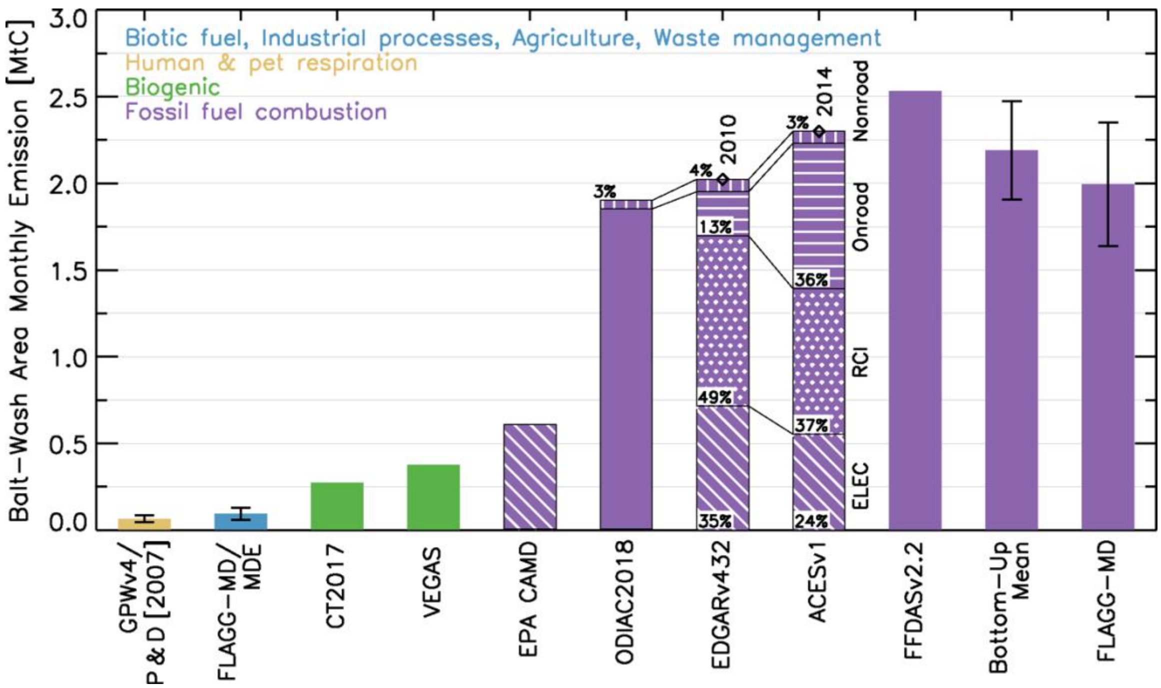


**Figure S6.** Emissions of $CO_2$ from the Balt-Wash area during February 2015. This figure is identical to Figure 9 of the main paper, except here we have again computed the FLAGG-MD emissions assuming $k = 1$ in Equation (1) (i.e., steady perpendicular winds across the study area). The FLAGG-MD monthly emission shown here (last vertical bar) is 5% larger than that shown in Figure 9.

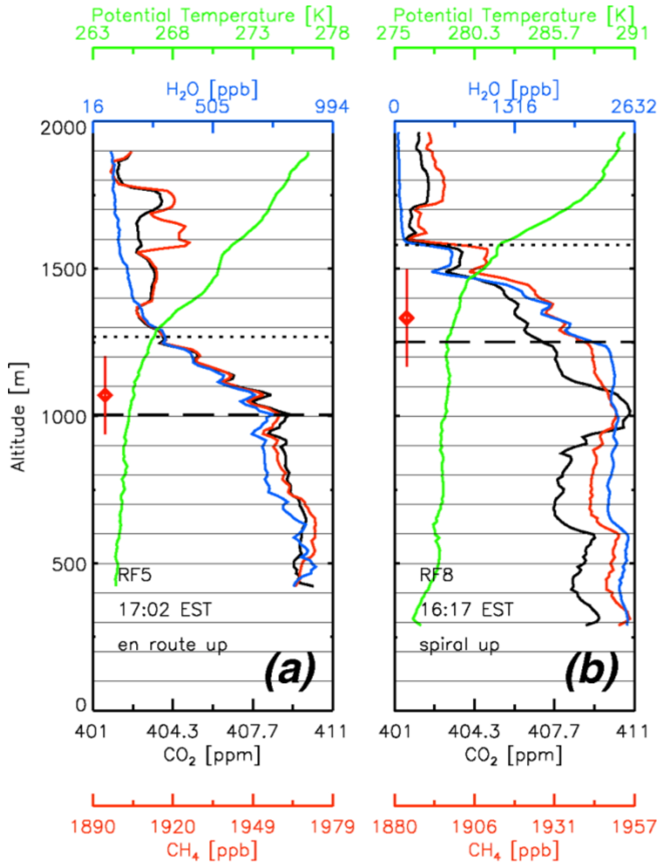


**Figure S7.** Vertical profiles of $CO_2$, $CH_4$, $H_2O$ and potential temperature downwind of the Balt-Wash area on (a) 20 February 2015 (UMD-RF5) and (b) 25 February 2015 (UMD-RF8). The locations of these vertical profiles are indicated as VP3 and VP5 in Figure 1. The dashed line represents the top of the well-mixed PBL. The dotted line represents the entrainment height. The red diamond and vertical error bar indicate the adjusted mixing height and its $\pm 1\sigma$ uncertainty range, used for the flux estimation in Equation (1).

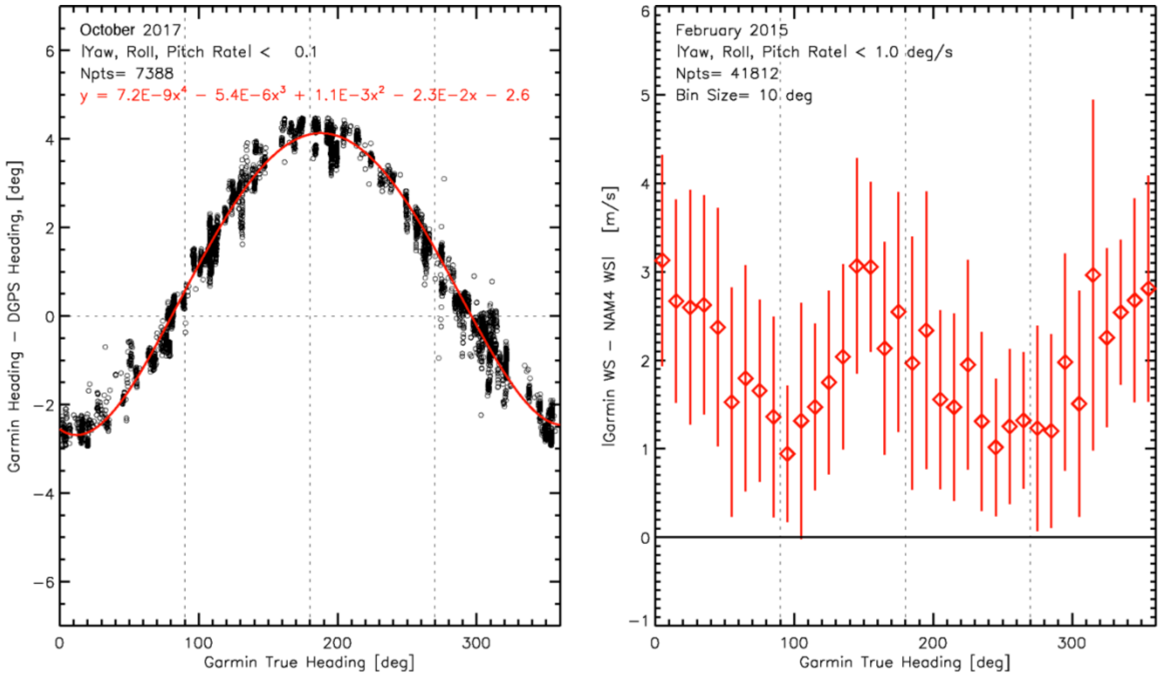


**Figure S8.** (a) The difference of true heading measurements obtained by the Garmin system and the Differential GPS (DGPS) during four calibration research flights conducted in October 2017. (b) The difference of v1 wind speed derived from the Garmin output and NAM4 sampled along the flight track as a function of the Garmin true heading, during eight UMD research flights conducted in February 2015.

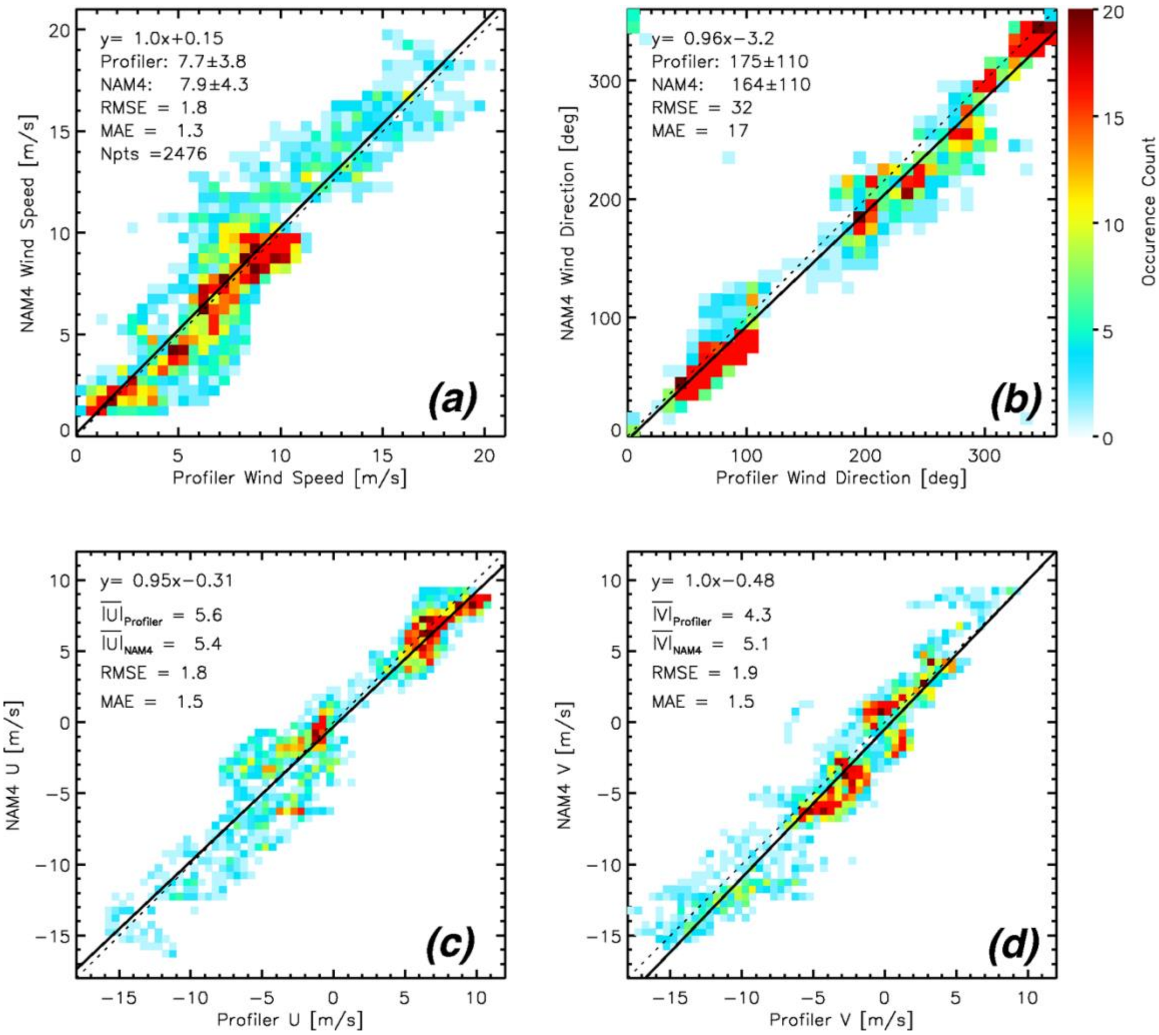


**Figure S9.** Scatter plots comparing the Beltsville site wind profiler measurements and the NAM4 meteorological model for (a) wind speed, (b) wind direction, (c) U component, and (d) V component wind. Dotted line indicates 1 to 1 ratio line and solid line indicates the linear regression fitted to the data. The data plotted were obtained during the eight flight days during the campaign (i.e., 13, 16, 19, 20, 23, 24, 25, and 26 February 2015).

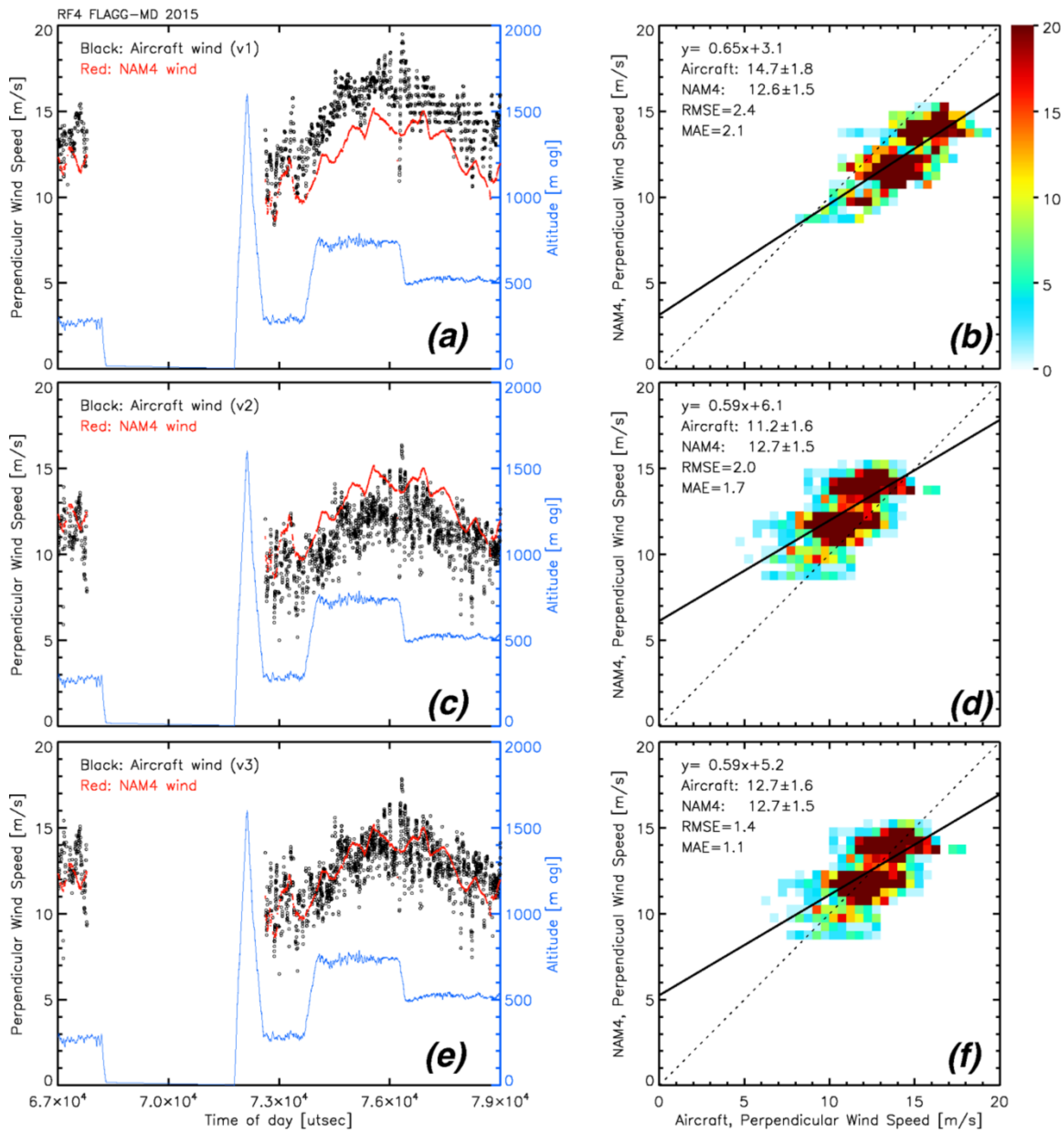


**Figure S10.** Comparisons between three versions of the aircraft wind perpendicular to the aircraft flight track and the perpendicular wind from NAM4. For each row, the left and right plots showing the same data, but as time series and scatter plots, respectively. The first row shows the comparison for the original v1 aircraft perpendicular wind. The second row shows the comparison for the v2 aircraft wind, which incorporates the magnetic heading bias correction and true airspeed calibration described in Text S1 to S3. The third row shows the comparison for the v3 aircraft wind, which is derived by scaling the perpendicular wind speed to the NAM4 data, as described in Text S2.

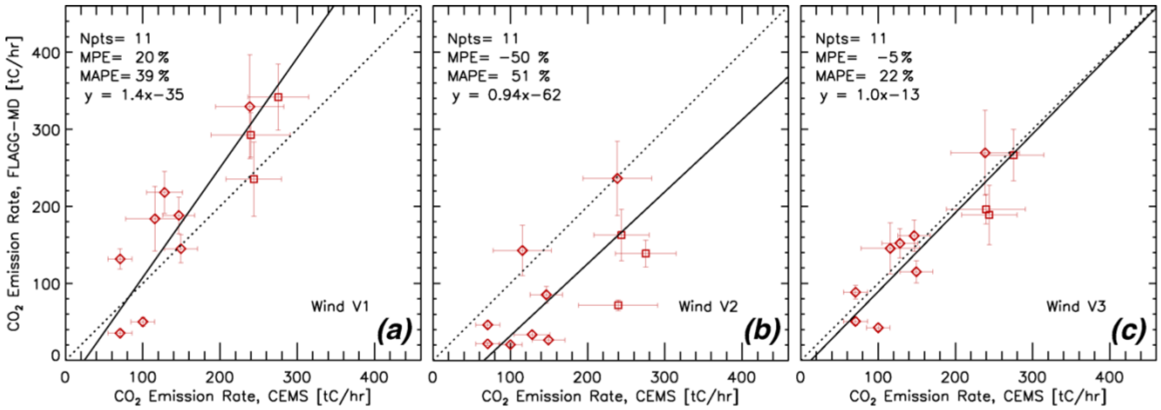


**Figure S11.** Scatter plots of the emission rate of $CO_2$ from the CEMS record of Chalk Point (CP) and Morgantown (MT) power plants versus the emission rate of $CO_2$ estimated using (a) v1 wind, (b) v2 wind, and (c) v3 wind fields of the UMD Cessna aircraft. The data points shown in (c) are identical to the UMD data points shown in Figure 6b. The mean percentage error (MPE) and the mean absolute percentage error (MAPE) of the UMD mass balance versus CEMS emissions are shown at the top left of each panel. The dotted line shows the 1 to 1 ratio and the solid line shows a linear least squares for of the data points, for each version of the wind field.  The close agreement of the linear fir on panel (c) to the 1 to 1 line supports the validity of the v3 wind field.

**Table S1.** The mean and the standard deviation of the three different versions of the aircraft perpendicular wind speed. The root mean square error (RMSE) of the perpendicular wind speed against the corresponding the NAM4 wind data are shown.

| Unit: m/s | | Wind v1 | | Wind v2 | | Wind v3 | |
|---|---|---|---|---|---|---|---|
| | Date | $\overline{\boldsymbol{U_\perp}} \pm \mathbf{1}\boldsymbol{\sigma}$ | RMSE | $\overline{\boldsymbol{U_\perp}} \pm \mathbf{1}\boldsymbol{\sigma}$ | RMSE | $\overline{\boldsymbol{U_\perp}} \pm \mathbf{1}\boldsymbol{\sigma}$ | RMSE |
| UMD-RF1 | Feb 6 2015 | 7.4±3.1 | 3.1 | 5.1±3.2 | 2.8 | 5.3±3.2 | 2.8 |
| UMD-RF3 | Feb 16 2015 | 5.4±1.0 | 2.1 | 2.3±1.0 | 1.3 | 3.4±1.0 | 0.7 |
| UMD-RF4 | Feb 19 2015 | 14.7±1.8 | 2.4 | 11.2±1.6 | 2.0 | 12.7±1.6 | 1.4 |
| UMD-RF5 | Feb 20 2015 | 7.2±1.6 | 1.5 | 3.4±1.5 | 2.9 | 6.1±1.5 | 1.0 |
| UMD-RF6 | Feb 23 2015 | 11.1±1.3 | 1.4 | 8.5±1.4 | 2.6 | 10.6±1.4 | 1.5 |
| UMD-RF8 | Feb 25 2015 | 6.7±2.2 | 2.4 | 2.9±1.9 | 2.8 | 5.1±1.9 | 1.7 |
| UMD-RF9 | Feb 26 2015 | 3.7±1.2 | 1.2 | 3.5±1.1 | 1.1 | 4.1±1.1 | 1.0 |

**References**

American Veterinary Medical Association. (2012). U.S. Pet ownership statistics. Retrieved September 7, 2019, from https://www.avma.org/KB/Resources/Statistics/Pages/Market-research-statistics-US-pet-ownership.aspx

Gately, C. K., & Hutyra, L. R. (2018). CMS: $CO_2$ Emissions from Fossil Fuels Combustion, ACES Inventory for Northeastern USA. ORNL DAAC, Oak Ridge, Tennessee, USA. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ornldaac/1501

Gurney, K., Liang, J., Patarasuk, R., O'Keeffe, D., Huang, J., Hutchins, M., et al. (2017). Reconciling the differences between a bottom-up and inverse-estimated $FFCO_2$ emissions estimate in a large US urban area. *Elem Sci Anth*, *5*. https://doi.org/10.1525/elementa.137

Gurney, K. R., Huang, J., & Coltin, K. (2016). Bias present in US federal agency power plant $CO_2$ emissions data and implications for the US clean power plan. *Environmental Research Letters*, *11*(6), 64005. https://doi.org/10.1088/1748-9326/11/6/064005

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., et al. (2017). EDGAR v4.3.2 Global atlas of the three major greenhouse gas emissions for the period 1970-2012. *Earth Syst. Sci. Data Discuss.*, *2017*, 1–55. https://doi.org/10.5194/essd-2017-79

MDE (Maryland Department of the Environment). (2016). Maryland 2014 Periodic GHG Emissions Inventory. Retrieved August 1, 2018, from https://mde.state.md.us/programs/Air/ClimateChange/Pages/GreenhouseGasInventory.aspx

Prairie, Y. T., & Duarte, C. M. (2007). Direct and indirect metabolic $CO_2$ release by humanity. *Biogeosciences*, *4*(2), 215–217. https://doi.org/10.5194/bg-4-215-2007

Quick, J. C., & Marland, E. (2019). Systematic error and uncertain carbon dioxide emissions from US power plants. *Journal of the Air & Waste Management Association*, 1–13. https://doi.org/10.1080/10962247.2019.1578702

Ren, X., Hall, D. L., Vinciguerra, T., Benish, S. E., Stratton, P. R., Ahn, D., et al. (2019). Methane emissions from the Marcellus shale in southwestern Pennsylvania and northern west Virginia based on airborne measurements. *Journal of Geophysical Research: Atmospheres*. https://doi.org/10.1029/2018JD029690

U.S. Environmental Protection Agency. (2009). Plain English Guide to the Part 75 Rule. Retrieved May 10, 2019, from https://www.epa.gov/sites/production/files/2015-05/documents/plain_english_guide_to_the_part_75_rule.pdf