Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

Atmospheric Chemistry and Physics Discussions

Assessing London CO_2, CH_4 and CO emissions using aircraft measurements and dispersion modelling

Joseph Pitt[1], Grant Allen[1], Stéphane Bauguitte[2], Martin Gallagher[1], James Lee[3], Will Drysdale[3], Bethany Nelson[3], Alistair Manning[4], Paul Palmer[5]

[1]School of Earth and Environmental Sciences, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[2]Facility for Airborne Atmospheric Measurements (FAAM), Cranfield University, Cranfield, MK43 0AL, UK
[3]Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, Heslington, York YO10 5DD, UK
[4]Met Office, FitzRoy Road, Exeter, EX1 3PB, UK
[5]School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK

Correspondence to: Joseph Pitt (joseph.pitt@manchester.ac.uk)

Abstract. We present a new modelling approach for assessing atmospheric emissions from a city, using an aircraft measurement sampling strategy similar to that employed by previous mass balance studies. Unlike conventional mass balance methods, our approach does not assume that city-scale emissions are confined to a well-defined urban area and that peri-urban emissions are negligible. We apply our new approach to a case study conducted in March 2016, investigating CO, CH_4 and CO_2 emissions from London using aircraft sampling of the downwind plume. For each species, we simulate the flux per unit area that would be observed at the aircraft sampling locations based on emissions from the UK national inventory, transported using a Lagrangian dispersion model. To reconcile this simulation with the measured flux per unit area, assuming the transport model is not biased, we require that inventory values of CO, CH_4, and CO_2 are scaled by 1.00, 0.70, and 1.57, respectively. However, our result for CO_2 must be treated with strong caution as we do not account for the influence of the land and ocean biosphere in this work.

For comparison, we also calculate fluxes using a conventional mass balance approach and compare these to the emissions inventory aggregated over the Greater London area. Using this method we derive much higher inventory scale factors for all three gases, as a direct consequence of neglecting emissions outside the Greater London boundary. That substantially different conclusions are drawn using the conventional mass balance method demonstrates the danger of using this technique for cities whose emissions cannot be separated from significant surrounding sources.

1. Introduction

Over half the people in the world (54%) live in urban areas. This proportion is projected to increase to 66% by 2050 (United Nations, 2014). Consequently, cities are responsible for a large proportion of anthropogenic greenhouse gas (GHG) emissions. The 2015 UNFCCC Paris Agreement requires signatory states to not only report national GHG emissions, but also to establish and improve independent methods for verifying these reported emissions (UNFCCC, 2015). Therefore the

development of top-down methods to measure city-scale emissions has become an increasingly important area of study, aiding both scientific and policy needs.

In the UK, spatially and sectorally disaggregated emissions calculated using a bottom-up methodology are given in the National Atmospheric Emissions Inventory (NAEI; Brown et al., 2017). For Greater London, nearly all sources of anthropogenic CO_2 and CO emissions are associated with fuel combustion. For CO_2, the main sources are domestic and commercial combustion and road transport. Emissions of CO are comprised of a range of combustion sources, with road transport emissions constituting the largest category. For CH_4 the principal sources are waste treatment and disposal, and leakage during natural gas distribution. Providing a top-down constraint on these London emissions is important, not only because London represents a large emission source in its own right, but also because it can help inform the assumptions that go into calculating bottom-up emission estimates for these sectors at a national level.

Aircraft mass balance techniques have previously been employed to measure trace gas fluxes from several cities (e.g. Mays et al., 2009; Turnbull et al., 2011; Cambaliza et al., 2014) including London (O'Shea et al., 2014). These techniques represent a direct method of flux measurement, and unlike atmospheric inversion approaches (e.g. Manning et al., 2011; Bergamaschi et al., 2014; Ganesan et al., 2015) or eddy covariance techniques (e.g. Grimmond et al., 2002; Vogt et al., 2006; Helfter et al., 2011) they can derive a bulk area flux, with a definable uncertainty, using only a few hours of sampling. Typically, horizontal transects are conducted downwind of a city at several altitudes to sample its emitted plume for various trace gas species. These vertically-stacked transects define a 2D plane of sampling downwind of the city. A background mole fraction can be determined by sampling upwind of the city or from downwind measurements outside of the plume. The mass flux of the plume through the 2D plane of sampling is then calculated from the measured mole fraction enhancements (above background) and wind speed.

While it is relatively simple to determine the flux for a given species through this 2D sampling plane downwind of a city, this value only becomes meaningful if it can be compared to the total city emissions given in the bottom-up inventory. To do so it is necessary to determine a boundary within which inventory emissions can be aggregated to calculate a bulk flux for the city. The choice of this boundary is unproblematic for isolated cities whose surrounding areas can be considered comparatively negligible emission sources. However, for cities such as London, which is surrounded by other non-negligible emission sources for CO_2, CH_4 and CO, it is difficult to determine unambiguously the area over which bottom-up emissions should be aggregated to enable comparison with top-down fluxes calculated using the mass balance method.

In this study we have developed an alternative method for assessing bottom-up inventory fluxes using the same aircraft sampling technique described above. In this new approach we have used the UK Met Office's Lagrangian dispersion model, NAME (Numerical Atmospheric dispersion Modelling Environment), to simulate the transport of the inventory fluxes to the location of the measurements. This approach enables direct comparison between measured and simulated fluxes at the aircraft sample locations, avoiding the need to aggregate inventory fluxes over some arbitrarily defined area on the ground. We have also applied the conventional mass balance technique to the same data, and compared the top-down fluxes derived to inventory emissions aggregated over the Greater London administrative area. The two approaches reach

significantly different conclusions regarding the accuracy of the emissions inventory; we suggest that this is a consequence of sources outside Greater London that nonetheless contribute to the plume measured downwind of the conurbation.

2. Case study details

2.1 Aircraft measurements and calibration

We recorded measurements on board the UK's Facility for Airborne Atmospheric Measurement (FAAM) BAe-146 atmospheric research aircraft (henceforth referred to as the FAAM aircraft). For full details of the aircraft payload see Palmer et al. (2018). Here we describe only those measurements that are relevant to this case study.

Mole fractions of CO_2 and CH_4 were measured using a Fast Greenhouse Gas Analyser (FGGA; Los Gatos Research, USA). Paul et al. (2001) describe the operating principle of the instrument, and O'Shea et al. (2013) provide details on the instrument operational practice and performance on the FAAM aircraft across several campaigns. The FGGA was calibrated hourly in flight, using two calibration gas standards traceable to the WMO-X2007 scale (Tans et al., 2009) and WMO-X2004A scale (an extension of the scale described by Dlugokencky et al., 2005) for CO_2 and CH_4, respectively. For this case study, the certified standards (369.54 ppm CO_2, 1853.8 ppb CH_4 and 456.97 CO_2, 2566.0 ppb CH_4 respectively) spanned the range of measured ambient mole fractions.

We measured a target cylinder containing intermediate mole fraction values approximately half way between these hourly calibrations to quantify any instrument non-linearity or drift. This flight formed part of the wider GAUGE (Greenhouse gAs UK and Global Emissions) campaign, during which we derived average target cylinder measurement offsets of 0.036 ppm for CO_2 and 0.07 ppb for CH_4, relative to the WMO-traceable values, with standard deviations of 0.398 ppm and 2.40 ppb, respectively, for 1 Hz sampling. Each individual target cylinder measurement consisted of a 20 s sample after allowing time for the measurements to reach equilibrium. The standard deviation of these 20 s averaged values was 0.245 ppm for CO_2 and 1.42 ppb for CH_4.

Another source of measurement uncertainty was the impact of water vapour in the sampled air on the retrieved CO_2 and CH_4 mole fractions. This was principally a consequence of mole fraction dilution (i.e. an increase in the total number of molecules per unit volume relative to dry air) and pressure broadening of the spectral absorption lines. The method used to correct for this is described by O'Shea et al. (2013). Using that technique we have derived maximal uncertainties due water vapour of 0.156 ppm for CO_2 and 1.05 ppb for CH_4. Finally, there are also uncertainties associated with the certification of the target cylinder of 0.075 ppm and 0.76 ppb, respectively. Combining all of these uncertainties with the target measurement standard deviations yields nominal total uncertainties for CO_2 and CH_4 of 0.434 ppm and 2.73 ppb at 1 Hz, and 0.300 ppm and 1.93 ppb when averaged over 20 s.

We measured CO mole fractions using vacuum ultraviolet florescence spectroscopy (AL5002, Aerolaser GmbH, Germany). The principle of this system is described by Gerbig et al. (1999), who also evaluate its performance on board an aircraft. Calibration was performed using in-flight measurements of a single gas standard and the background signal at zero

CO mole fraction. Gerbig et al. (1999) derive a 1 Hz repeatability of 1.5 ppb (at 100 ppb), and an accuracy of 1.3 ppb ± 2.4% for the 1 Hz measurements.

Details of the meteorological instrumentation on board the FAAM aircraft are provided by Petersen and Renfrew (2009). In summary, we measured temperature with a Rosemount 102AL sensor, with an overall measurement uncertainty of 0.3 K at 95% confidence; we took static pressure measurements from the air data computer, based on measurements from pitot tubes around the fuselage, with an estimated absolute accuracy of 0.5 hPa; we made 3D wind measurements using the a 5-hole probe system described by Brown et al. (1983), with an estimated uncertainty in horizontal wind measurements of < 0.5 m s^{-1}.

2.2 Sampling strategy

On 04 March 2016 we conducted a targeted case study flight to measure CO_2, CH_4 and CO mole fraction enhancements downwind of London, so as to assess the representativeness of the bottom-up emissions inventory. The flow over the region was consistently westerly, bringing background air from the northern Atlantic with an average travel time over the British Isles of less than 24 hours (as determined using NAME). As the influence of land-based sources on the recent history of this air mass (upwind of the British Isles) can be assumed to be negligible, we expect that air arriving at the British Isles had relatively homogeneous and well-mixed trace gas composition throughout the boundary layer prior to the influence of local fluxes. Figure 1 shows the flight track from an aerial perspective; between points A and B we flew repeated horizontal transects at various altitudes through a plume of enhanced mole fractions downwind of London emission sources. At the southernmost end of these transects, the constraints of UK airspace forced us to deviate from our desired course perpendicular to the prevailing wind. However, as we sampled the overwhelming majority of the London plume north of this imposed turning point, such that measurements during the deviation to point B represented background (out-of-plume) sampling, we do not expect this deviation to impact on the derived fluxes.

During an initial transect at 1550 m altitude we measured typical uniform free tropospheric background mole fractions for all three gases (CO_2, CH_4 and CO). Following this we descended to 120 m, and subsequently flew 6 transects of increasing altitude, breaking the final transect short to profile up to 1550 m within the observed plume. Figures 2a, 2b and 2c show these transects, coloured by mole fraction for CO_2, CH_4 and CO respectively, projected onto a altitude-latitude plane.

2.3 Dispersion model configuration

To determine the air history corresponding to the continuous aircraft sampling we ran the NAME dispersion model in backwards mode, releasing 100 tracer particles at each 1 Hz aircraft measurement location and tracking their motion back in time. NAME was driven by meteorological data from the UK Met Office's UKV model (Tang et al., 2013), which provides hourly data on 70 vertical levels at 1.5 km horizontal resolution over the British Isles. NAME determines particle motion based on the mean wind field (which is determined by interpolating the met data spatially and temporally to the particle location for each time-step) and a parameterisation of unresolved turbulent and mesoscale motions (for details see Jones et

al., 2007, and references therein). In this study we used a NAME model time-step of 1 minute. By way of guidance, it is worth noting that although this NAME setup is more computationally intensive than is typically employed, because the release duration was less than 3 hours and the maximum particle travel time before leaving the domain was 37 hours (and less than 24 hours for the majority of particles) the run completion time remained on the order of hours rather than days using the JASMIN scientific computing facility.

To assess the influence of surface fluxes on the sampled air, we recorded the total time each tracer particle spent in the lowest 100 m above ground level on a 1 km x 1 km horizontal grid (UK National Grid), matching the spatial resolution of the NAEI emissions inventory. These individual particle residence times were then averaged over all particles released during each minute of aircraft sampling, giving the average time the sampled air spent in each grid box for every minute of the flight (henceforth referred to as a release period). A time series of simulated mole fraction enhancements (X) at the aircraft sample locations due to NAEI inventory emissions (F_{ij}) was then calculated according to:

$$X(t) = \sum_{i,j} D_{ij}(t) \times F_{ij} \quad (1)$$

where the air history matrix D_{ij} represents the mole fraction enhancement at the sample locations due to a unit flux in each grid box, calculated for each release period according to:

$$D_{ij}(t) = \frac{T_{ij}(t)}{d \times n(t)} \quad (2)$$

The indices i and j in Eqs. (1) and (2) represent the northing and easting components of the horizontal grid. $T_{ij}(t)$ is the average particle residence time in each grid box for every release period, d is the height of the grid boxes (here 100 m) and $n(t)$ is the average molar density of air (mol_{air} m^{-3}) at the aircraft sampling locations based on UKV data. The emissions F_{ij} are given here in units of mol_X m^{-2} s^{-1}, where X represents either CO_2, CH_4 or CO. Figures 2d, 2e and 2f show the equivalent data to Figs. 2a, 2b and 2c, coloured by simulated mole fraction enhancement rather than measured mole fraction.

3. Inventory flux comparisons

In this section we present two approaches to assess the accuracy of the NAEI inventory emission values relative to the measured mole fractions during this case study. The first is a new approach, referred to hereafter as the flux-dispersion method, using the simulated mole fraction enhancements from Sect. 2.3 to derive simulated fluxes through the downwind sampling plane based on inventory emissions, thus enabling comparison with corresponding measured fluxes. The results from this method represent our best assessment of inventory fluxes for this case study.

In Sect. 3.2 we then employ a conventional mass balance method to derive top-down fluxes which are compared to an aggregated NAEI value. We discuss the outcomes of both approaches in Sect. 3.3 and explain how the conventional mass balance approach can lead to spurious conclusions in cases such as this.

The most recent gridded emissions available in the NAEI at the time of writing were for the year 2015; therefore we have used these 2015 emissions to represent the 2016 values in both approaches. The UK totals (not spatially disaggregated)

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

for 2016 have been released, allowing us to compare these to the 2015 totals. For CO there was a 9.4% reduction in total reported emissions between 2015 and 2016, while for CO_2 there was a reduction of 5.8% and for CH_4 there was an increase of 0.01%.

3.1 Flux-dispersion method

3.1.1 Methodology

To make a comparison between the measured and simulated datasets described in Sect. 2 it is first necessary to calculate a background mole fraction for both, so that the mole fraction enhancement due to the London plume can be determined. To determine periods of sampling that were not significantly influenced by the London plume, and therefore can be considered to represent background mole fractions, we again utilised the air history information given by the NAME dispersion modelling. From the gridded air histories described in Sect. 2.3, we calculated the fraction of $D_{ij}(t)$ that was within the Greater London administrative region for each release period, and defined all release periods where this fraction was less than 0.05% as background periods. This Greater London fraction is shown in Fig. 3, with the background periods coloured red.

In practice there is no sharp distinction between in-plume and background sampling, so any criteria used to separate sampling into these two categories inherently involves some level of human judgement. However, it is important to understand that changing this threshold, and therefore selecting different background periods, essentially changes the interpretation of the results as well as the flux values themselves. This is better illustrated by considering Fig. 4, which shows the air history ($D_{ij}(t)$) aggregated over both the background (Fig 4a) and in-plume periods (Fig 4b), which clearly illustrate that the background criteria used here avoids air histories that have passed over the London conurbation. The comparison between measured and simulated flux discussed later in this section is a comparison between the flux enhancement from the areas sampled in Fig. 4b relative to any flux enhancement from the air histories sampled in Fig. 4a. The justification for choosing the threshold of 0.05% here is essentially that Fig. 4b is dominated by Greater London air histories and Fig. 4a is not, enabling us to specifically assess accuracy of the NAEI for the London conurbation.

For each constant-altitude aircraft transect we calculated average background mole fractions to the north and south of the plume separately, for both measured and simulated datasets. We then calculated the mole fraction enhancement, $\Delta X_{London}(t)$, for both datasets using Eq. (3):

$$\Delta X_{London}(t) = X(t) - \frac{\left(\overline{X_{bgd\,N}}(z) + \overline{X_{bgd\,S}}(z)\right)}{2} \quad (3)$$

Here $X(t)$ is the mole fraction time series and $\overline{X_{bgd\,N}}(z)$ and $\overline{X_{bgd\,S}}(z)$ are the average background mole fractions to the north and south of the plume for each transect. The motivation for treating the background in this way is to capture potential latitudinal and vertical gradients in mole fraction. Vertical gradients are accounted for by calculating a separate background value for each transect, while latitudinal gradients are accounted for by the separate calculation of the north and south backgrounds. If a straight average over all background periods was used for each transect, this would be subject to

potential bias in cases where there was more background sampling on one side of the plume than the other. Calculating north and south backgrounds separately as in Eq. (3) mitigates this issue.

The time series of measured and simulated mole fraction enhancements calculated using Eq 3 are directly comparable quantities. However, to reduce the sensitivity of our results to any difference between modelled and measured wind speed, we convert these mole fraction enhancements into fluxes per unit area in the mean wind direction (i.e. through the downwind sampling plane) before making a comparison between them. To define a representative wind direction, we took the average of the mean UKV model wind direction and the mean measured wind direction during the sampling period. A time series of flux per unit area in this average wind direction, hereafter referred to as the flux density, was then calculated for both measured and simulated datasets using Eq. (4):

$$FD(t) = \Delta X_{London}(t) \times n_{air}(t) \times U_{\perp}(t) \quad (4)$$

The mole fraction enhancement (ΔX_{London}), molar density of air (n_{air}) and wind speed in the mean wind direction ($U_{\perp}$) were calculated independently for the measured and simulated datasets, producing flux densities (FD) in mol m^{-2} s^{-1} that are directly comparable.

3.1.2 Flux-dispersion results

Figure 5 shows a comparison between these measured and simulated flux densities as a function of latitude for each plume transect. The top two transects from Figs. 2 and 3 have been excluded here because they are entirely above the average boundary layer height of 759 m used by the NAME dispersion model for this simulation (which is taken directly from the UKV met data). Above this height the parameters used by NAME to describe the turbulent motion of the particles are set to fixed values resulting in poorer representation of particle dispersion. Within the boundary layer these parameters are calculated from the friction velocity and characteristic convective velocity. Therefore the flux densities calculated for transects within the model boundary layer are more accurate than those above it. The lowest transect (~120 m) has also been excluded because no value for $\overline{X_{bgd\ N}}$ was obtained – this was the first transect conducted before the position of the plume had been fully established so its northern extent was not sampled.

Having calculated time series of flux density for the measured and simulated datasets, we then calculated average flux densities for each transect altitude. We also calculated flux densities as an overall average using data from all three transects. These values are given in Table 1 for CO_2, CH_4 and CO, along with the ratios between measured and simulated flux densities. These ratios represent the factors by which the NAEI inventory needs to be multiplied in order to reproduce the measured flux densities (assuming there is no bias induced by the NAME transport). Therefore for this case study we conclude that the NAEI emissions for London require scaling by the overall values of 1.00 for CO, 0.70 for CH_4 and 1.57 for CO_2.

While there are small uncertainties associated with the measured mole fractions (as discussed in Sect. 2.1), the uncertainty in these overall inventory scale factors is expected to be dominated by NAME transport uncertainty. This transport uncertainty is fundamentally unknown and a probability distribution of transport error for this case study cannot be

determined. Instead we take the range of scale factors across the different transects of 0.90-1.13 for CO, 0.61-0.79 for CH_4 and 1.40-1.77 for CO_2 to give an indication of the total scale factor uncertainty for each species.

Considering both the central estimate and uncertainty range for CO, we can conclude that the measurements taken during this case study are consistent with the emissions given in the NAEI inventory. For CH_4, however, the NAEI emissions yield an overestimate of flux density enhancement for each transect, and overall require downscaling by a factor of 0.70 to agree with observations. This is qualitatively consistent with studies suggesting national NAEI CH_4 emission totals have been too high in previous years (e.g. Manning et al., 2011; Ganesan et al., 2015) but it differs from the most recent top-down verification report (O'Doherty et al., 2017), which finds good agreement between the NAEI CH_4 emission totals and continuous ground-based measurements in recent years.

There are multiple possible explanations for the difference between our results for CH_4 emissions and those in the verification report. Our study specifically measures the magnitude of emissions from the London conurbation relative to emissions from surrounding areas. It is possible that although NAEI emission totals agree with long term observations, the spatial distribution of these emissions is not well represented in the inventory, such that the proportion of emissions ascribed to urban areas is too large. Alternatively, as our study represents a single snapshot of emissions taken during a single flight in March 2016, this difference could reflect temporal variability in CH_4 emissions which is not represented in the NAEI. Helfter et al. (2016) used an eddy-covariance technique to determine the diurnal variability of London CH_4 emissions, and found that emissions increased by a factor of 1.9 (maximum-to-minimum ratio) between the early and late morning. However, in addition to true temporal variability (e.g. rush hour emissions). such techniques are susceptible to the complex nature of urban boundary layer evolution during these transition periods (Halios and Barlow, 2018), which can result in overestimation of diurnal flux variability. Ultimately we cannot distinguish between spatial and temporal patterns in emissions using only a single case study. However, future flights could be used in conjunction with the method presented here to build a statistically representative sample of both diurnal and intra-annual flux variability, thus also allowing the spatial distribution of NAEI emissions to be better assessed on an annual timescale.

For CO_2 we find that the NAEI would require upscaling in order to be consistent with observations. However, this conclusion must be treated with caution as the NAEI does not include biospheric fluxes which are significant for CO_2. These biospheric fluxes include uptake due to photosynthesis (gross primary production; GPP), emission from autotrophic respiration and emission from heterotrophic respiration. Net primary production (NPP) is calculated as the difference between photosynthetic uptake and autotrophic respiration. Hardiman et al. (2017) investigated biospheric CO_2 fluxes in Massachusetts and found higher NPP values outside the Boston conurbation. Combined with higher heterotrophic respiration in more populated areas, these higher rural NPP values result in a positive net biospheric flux from urban centres relative to their surrounding areas. As we have not accounted for this net biospheric flux in our simulated flux densities we can expect them to underestimate the measured values, even if the NAEI emissions are entirely accurate. Quantifying the biospheric impact on the derived scale factor would require the use of an ecosystem model, and is beyond the scope of this study. However the influence of the biosphere must be considered as a key source of uncertainty on our derived CO_2 scale factor.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

Atmospheric Chemistry and Physics Discussions

3.2 Conventional mass balance method

3.2.1 Methodology

Detailed descriptions of the mass balance technique in the context of measuring urban GHG emissions are provided by many sources. In general, in the context of bulk area flux measurement, these sources can be categorized into two basic approaches; either the emissions are assumed to be well mixed up to a given height at which they are capped by a temperature inversion (Turnbull et al., 2011; Karion et al., 2013; Smith et al., 2015), or the vertically varying shape of the plume is derived by interpolation between transects flown at multiple altitudes (Mays et al., 2009; Cambaliza et al., 2014; O'Shea et al., 2014), often using a kriging approach. Figures 2a, 2b and 2c clearly show that the assumptions of the first of these approaches (i.e. well mixed plumes up to a capping height) are not met in this case. We therefore adopt the latter of these approaches, and use kriging to represent the full structure of the plume. This approach necessarily assumes temporal invariance of the plume over the period of sampling: in this case ~2.5 hours.

Following the work of (Mays et al., 2009; Cambaliza et al., 2014; O'Shea et al., 2014) we derive fluxes using Eq. (5):

$$F = \int_0^{z_{max}} \int_A^B \left(X_{ij} - X_0\right) n_{air}(z)\, U_{\perp ij}\, \mathrm{dx}\, \mathrm{dz} \quad (5)$$

Here F (mol s^{-1}) is the bulk flux for the emission source, X_{ij} is the kriged mole fraction for a given species, X_0 is the background mole fraction, $n_{air}(z)$ is the molar air density (here derived as a linear function of altitude based on measured values) and $U_{\perp ij}$ is the kriged wind speed perpendicular to the vertical sample plane across which the integral is taken.

Kriging is an interpolation method based on a stochastic Gaussian model, and is described in detail by Kitanidis (1997). It converts samples with sparse spatial coverage into a 2D grid of estimated values, with an associated grid of standard errors for these values. Here we use a modified version of the EasyKrig software (© Dezhang Chu and Woods Hole Ocean Institution) to perform the kriging; again more detail regarding the application of this software with regards to aircraft mass balance flux calculations is given by Mays et al. (2009). More detail regarding the kriging parameters used is included in the supplementary material.

The results from the kriging were output on a 20 x 29 cell grid, with a vertical resolution of 50 m and a horizontal resolution of 5 km respectively, as shown in Fig. 6. As the lowest transect was conducted at ~120 m altitude, the structure of the plume below this level was not constrained well by our sampling. Therefore the mole fractions for the lowest 100 m above ground level were taken to be the same as the kriged output for the layer at 100 – 150 m. We determined the background mole fraction X_0 for each trace gas by taking the average mole fraction over all cells within 15 km of the horizontal boundaries. This approach follows Mays et al. (2009) in determining the background from measurements in the downwind plane outside of the influence of the plume, and contrasts with the approach used by O'Shea et al. (2014), who instead used measurements upwind of London to determine the background. Using an upwind background approach such as in O'Shea et al. (2014), one is forced to assume that the boundary layer height remains constant, and that there is no

exchange of air between the boundary layer and the free troposphere. In comparison, the downwind background method used here only assumes that the boundary layer height and free tropospheric exchange is the same throughout the downwind plane, which represents a less stringent assumption that is likely to hold better in practice.

3.2.2 Mass balance results

The fluxes calculated using Eq. (5) are given in Table 2, along with 1σ uncertainties derived by combining the kriging standard errors with the uncertainty in background mole fraction, taken to be the standard deviation for all background cells used. Also given are the aggregated NAEI emissions for the Greater London administrative area for each species. We have derived inventory scale factors, in principle analogous to those in Sect. 3.1.2, by taking the ratio of these aggregated NAEI emissions to the flux calculated using Eq. (5) for each species. Using the conventional mass balance method we calculate that the NAEI requires rescaling by factors of 2.27 for CO, 1.22 for CH_4 and 3.08 for CO_2. The differences between these values and those derived in Sect. 3.1.2 are discussed in Sect. 3.3 below.

3.3 Comparing the flux-dispersion and mass balance methods

In Sects. 3.1 and 3.2 two different methods were applied to the same dataset to derive scale factors for the NAEI inventory such that it agrees with aircraft observations. However, the scale factors derived using the flux-dispersion method are significantly lower than those derived using the conventional mass balance method. This is because one of the key elements of the mass balance method, the assumption that a city acts as an isolated emission source surrounded by areas with negligible emissions, is clearly violated in this case. By choosing to aggregate only NAEI emissions within the Greater London administrative area we have ignored any surrounding sources, but in the case of London there are significant sources of all three gases outside this administrative area. Leaving these sources out of the inventory total biases the derived inventory scale factors high for the mass balance method. The influence of these surrounding emission sources could also explain the large inventory upscaling factors derived by O'Shea et al. (2014), who used a mass balance approach to estimate London emissions for a previous case study in 2012.

Clearly one could choose a different region over which to aggregate NAEI emissions, but for a non-isolated source such a choice is inherently arbitrary and yet has a direct impact on the derived inventory scale factor. Just as ignoring sources which contribute to the plume resulted here in high-biased scale factors, aggregating over too wide an area would introduce a low bias, as sources which actually contribute predominantly to the measured background are instead assumed to contribute to the plume. The impact of the arbitrary choice of inventory aggregation area on the conclusions drawn using the mass balance method demonstrates the inappropriateness of the mass balance technique for non-isolated emission sources. Instead, the flux-dispersion method provides a good alternative in these cases, as it is not subject to such biases.

5. Conclusions

Aircraft mass balance techniques are an effective way of determining emissions from isolated sources, but they require surrounding areas to be negligible emission sources in order to yield robust results. This is a well-known assumption associated with these methods. However in the absence of alternative techniques using the same sample dataset against which the mass balance results can be compared, one is forced either to simply state this assumption as a caveat or to abandon the effort entirely.

In this study we have developed an alternative technique using a Lagrangian dispersion model to quantify the transport of inventory emissions to the aircraft sample locations, so that a direct comparison of flux per unit area can be made at the measurement locations. In contrast to the conventional mass balance technique, this method does not require cities to be isolated from surrounding emission sources, rendering it more appropriate in many cases. We have demonstrated this new technique by applying it to a single-flight case study measuring London emissions, which yielded inventory scale factors of 1.00 (0.90-1.13) for CO, 0.70 (0.61-0.79) for CH_4 and 1.57 (1.40-1.77) for CO_2. Using a mass balance approach we derived significantly higher values (2.27, 1.22 and 3.08 respectively), which we conclude are biased as a consequence of choosing an inventory aggregation area that ignores significant sources outside the Greater London administrative region. The magnitude of this bias demonstrates how employing a mass balance method for a non-isolated source can lead to highly misleading conclusions regarding the accuracy of the emissions inventory under study.

It is important to emphasise that the inventory scale factors derived here represent the results from a single case study and therefore are not necessarily representative of the annual timescale of the NAEI emissions. In order to better validate the inventory on this timescale repeated flights following a similar sampling strategy are required. Using the flux-dispersion method developed here in combination with representative aircraft sampling on an annual timescale could provide a robust assessment of inventory fluxes at the city scale in the case of non-isolated sources for which the mass balance technique is not appropriate.

Author contributions

Conceptualization, J. P., G. A., J. L.; Data curation; J. P., S. B.; Formal analysis; J. P.; Funding acquisition; G. A., M. G., J. L., A. M., P. P.; Investigation; J. P., G. A., S. B., J. L., W. D., B. N.; Methodology; J. P., G. A.; Software, A. M.; Supervision, G. A., M. G., J. L., P. P.; Writing – original draft, J. P.; Writing – review & editing, J. P. G. A., P. P.

Acknowledgements

The authors would like to thank everyone at FAAM, Airtask and Avalon Aero who helped with the collection and processing of the aircraft data used here. We acknowledge use of the NAME atmospheric dispersion model and associated NWP meteorological data sets made available to us by the Met Office. We acknowledge the significant storage resources and

analysis facilities made available to us on JASMIN by STFC CEDA along with the corresponding support teams. J. R. Pitt received a NERC CASE studentship in partnership with FAAM, grant number NE/L501/591/1, supervised by G. Allen. This work was supported by the GAUGE (Greenhouse gAs UK and Global Emissions) NERC project, grant numbers NE/K002449/1 and NE/K00221X/1.

References

Bergamaschi, P., Corazza, M., Karstens, U., Athanassiadou, M., Thompson, R. L., Pison, I., Manning, A. J., Bousquet, P., Segers, A., Vermeulen, A. T., Janssens-Maenhout, G., Schmidt, M., Ramonet, M., Meinhardt, F., Aalto, T., Haszpra, L., Moncrieff, J., Popa, M. E., Lowry, D., Steinbacher, M., Jordan, A., O'Doherty, S., Piacentino, S. and Dlugokencky, E.: Top-down estimates of European CH_4 and N_2O emissions based on four different inverse models, Atmos. Chem. Phys., 15, 715–736, doi:10.5194/acp-15-715-2015, 2015.

Brown, E. N., Friehe, C. A. and Lenschow, D. H.: The Use of Pressure Fluctuations on the Nose of an Aircraft for Measuring Air Motion, J. Clim. Appl. Meteorol., 22, 171–180, doi:10.1175/1520-0450(1983)022<0171:TUOPFO>2.0.CO;2, 1983.

Brown, P., Broomfield, M., Cardenas, L., Choudrie, S., Kilroy, E., Jones, L., Passant, N., Thomson, A., Wakeling, D., Buys, G., Forden, S., Gilhespy, S., Glendining, M., Gluckman, R., Henshall, P., Hobson, M., MacCarthy, J., Malcolm, H., Manning, A., Matthews, R., Milne, A., Misselbrook, T., Moxley, J., Murrells, T., Salisbury, E., Sussams, J., Thistlethwaite, G., Walker, C. and Webb, N.: UK Greenhouse Gas Inventory, 1990 to 2015, Ricardo Energy & Environment, Harwell, UK, 2017.

Cambaliza, M. O. L., Shepson, P. B., Caulton, D. R., Stirm, B., Samarov, D., Gurney, K. R., Turnbull, J., Davis, K. J., Possolo, A., Karion, A., Sweeney, C., Moser, B., Hendricks, A., Lauvaux, T., Mays, K., Whetstone, J., Huang, J., Razlivanov, I., Miles, N. L. and Richardson, S. J.: Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions, Atmos. Chem. Phys., 14, 9029–9050, doi:10.5194/acp-14-9029-2014, 2014.

Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W., Hall, B. D., Elkins, J. W. and Steele, L. P.: Conversion of NOAA atmospheric dry air CH4 mole fractions to a gravimetrically prepared standard scale, J. Geophys. Res., 110, D18306, doi:10.1029/2005JD006035, 2005.

Ganesan, A. L., Manning, A. J., Grant, A., Young, D., Oram, D. . E., Sturges, W. T., Moncrieff, J. B. and O'Doherty, S.: Quantifying methane and nitrous oxide emissions from the UK and Ireland using a national-scale monitoring network, Atmos. Chem. Phys., 15, 6393–6406, doi:10.5194/acp-15-6393-2015, 2015.

Gerbig, C., Schmitgen, S., Kley, D., Volz-Thomas, A., Dewey, K. and Haaks, D.: An improved fast-response vacuum-UV resonance fluorescence CO instrument, J. Geophys. Res.-Atmos., 104, 1699–1704, doi:10.1029/1998JD100031, 1999.

Grimmond, C. S. ., King, T. ., Cropley, F. ., Nowak, D. . and Souch, C.: Local-scale fluxes of carbon dioxide in urban environments: methodological challenges and results from Chicago, Environ. Pollut., 116, S243–S254, doi:10.1016/S0269-

7491(01)00256-1, 2002.

Halios, C. H. and Barlow, J. F.: Observations of the Morning Development of the Urban Boundary Layer Over London, UK, Taken During the ACTUAL Project, Bound.-Lay. Meteorol., 166, 395–422, doi:10.1007/s10546-017-0300-z, 2018.

Hardiman, B. S., Wang, J. A., Hutyra, L. R., Gately, C. K., Getson, J. M. and Friedl, M. A.: Accounting for urban biogenic fluxes in regional carbon budgets, Sci. Total Environ., 592, 366–372, doi:10.1016/j.scitotenv.2017.03.028, 2017.

Helfter, C., Famulari, D., Phillips, G. J., Barlow, J. F., Wood, C. R., Grimmond, C. S. B. and Nemitz, E.: Controls of carbon dioxide concentrations and fluxes above central London, Atmos. Chem. Phys., 11, 1913–1928, doi:10.5194/acp-11-1913-2011, 2011.

Helfter, C., Tremper, A. H., Halios, C. H., Kotthaus, S., Bjorkegren, A., Grimmond, C. S. B., Barlow, J. F. and Nemitz, E.: Spatial and temporal variability of urban fluxes of methane, carbon monoxide and carbon dioxide above London, UK, Atmos. Chem. Phys., 16, 10543–10557, doi:10.5194/acp-16-10543-2016, 2016.

Jones, A., Thomson, D., Matthew Hort and Devenish, B.: The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III, in: Air Pollution Modeling and Its Application XVII, edited by C. Borrego and A.-L. Norman, 580–589, Springer US, New York, NY, USA., 2007.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Hardesty, R. M., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R. and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophys. Res. Lett., 40, 4393–4397, doi:10.1002/grl.50811, 2013.

Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, Cambridge, U.K., 1997.

Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G. and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, J. Geophys. Res., 116, D02305, doi:10.1029/2010JD014763, 2011.

Mays, K. L., Shepson, P. B., Stirm, B. H., Karion, A., Sweeney, C. and Gurney, K. R.: Aircraft-based measurements of the carbon footprint of Indianapolis., Environ. Sci. Technol., 43, 7816–7823, doi:10.1021/es901326b, 2009.

O'Doherty, S., Stanley, K., Rigby, M., Stavert, A., Manning, A., Redington, A., Simmonds, P., Young, D. and Sturges, B.: Long-term atmospheric measurement and interpretation (of radiatively active trace gases), Report to BEIS (Department for Business, Energy and Industrial Strategy), Contract number: GA0201, 2017.

O'Shea, S. J., Bauguitte, S. J.-B., Gallagher, M. W., Lowry, D. and Percival, C. J.: Development of a cavity-enhanced absorption spectrometer for airborne measurements of CH4 and CO2, Atmos. Meas. Tech., 6, 1095–1109, doi:10.5194/amt-6-1095-2013, 2013.

O'Shea, S. J., Allen, G., Fleming, Z. L., Bauguitte, S. J.-B., Percival, C. J., Gallagher, M. W., Lee, J., Helfter, C. and Nemitz, E.: 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.-Atmos., 119, 4940–4952, doi:10.1002/2013JD021269, 2014.

Palmer, P. I., O'Doherty, S., Allen, G., Bower, K., Bösch, H., Chipperfield, M. P., Connors, S., Dhomse, S., Feng, L., Finch, D. P., Gallagher, M. W., Gloor, E., Gonzi, S., Harris, N. R. P., Helfter, C., Humpage, N., Kerridge, B., Knappett, D., Jones, R. L., Le Breton, M., Lunt, M. F., Manning, A. J., Matthiesen, S., Muller, J. B. A., Mullinger, N., Nemitz, E., O'Shea, S., Parker, R. J., Percival, C. J., Pitt, J., Riddick, S. N., Rigby, M., Sembhi, H., Siddans, R., Skelton, R. L., Smith, P., Sonderfeld, H., Stanley, K., Stavert, A. R., Wenger, A., White, E., Wilson, C. and Young, D.: A measurement-based verification framework for UK greenhouse gas emissions: an overview of the Greenhouse gAs Uk and Global Emissions (GAUGE) project, Atmos. Chem. Phys., 18, 11753–11777, doi:10.5194/acp-18-11753-2018, 2018.

Paul, J. B., Lapson, L. and Anderson, J. G.: Ultrasensitive absorption spectroscopy with a high-finesse optical cavity and off-axis alignment., Appl. Optics, 40, 4904–4910, 2001.

Petersen, G. N. and Renfrew, I. A.: Aircraft-based observations of air-sea fluxes over Denmark Strait and the Irminger Sea during high wind speed conditions, Q. J. Roy. Meteor. Soc., 135, 2030–2045, doi:10.1002/qj.355, 2009.

Smith, M. L., Kort, E. A., Karion, A., Sweeney, C., Herndon, S. C. and Yacovitch, T. I.: Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions, Environ. Sci. Technol., 49, 8158–8166, doi:10.1021/acs.est.5b00219, 2015.

Tang, Y., Lean, H. W. and Bornemann, J.: The benefits of the Met Office variable resolution NWP model for forecasting convection, Meteorol. Appl., 20, 417–426, doi:10.1002/met.1300, 2013.

Tans, P., Zhao, C. and Kitzis, D.: The WMO Mole Fraction Scales for CO2 and other greenhouse gases, and uncertainty of the atmospheric measurements, in: 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases, and Related Tracer Measurement Techniques, Jena, Germany., 2009.

Turnbull, J. C., Karion, A., Fischer, M. L., Faloona, I., Guilderson, T., Lehman, S. J., Miller, B. R., Miller, J. B., Montzka, S., Sherwood, T., Saripalli, S., Sweeney, C. and Tans, P. P.: Assessment of fossil fuel carbon dioxide and other anthropogenic trace gas emissions from airborne measurements over Sacramento, California in spring 2009, Atmos. Chem. Phys., 11, 705–721, doi:10.5194/acp-11-705-2011, 2011.

UNFCCC: Adoption of the Paris Agreement, in: Conference of the Parties: Twenty-first session, Paris, France., 2015.

United Nations: World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352), United Nations, Department of Economic and Social Affairs, Population Division, 2014.

Vogt, R., Christen, A., Rotach, M. W., Roth, M. and Satyanarayana, A. N. V: Temporal dynamics of CO2 fluxes and profiles over a Central European city, Theor. Appl. Climatol., 84, 117–126, doi:10.1007/s00704-005-0149-9, 2006.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

Atmospheric Chemistry and Physics Discussions

	CO			CH_4			CO_2		
	Mean flux density (µmol m^{-2} s^{-1})			Mean flux density (µmol m^{-2} s^{-1})			Mean flux density (µmol m^{-2} s^{-1})		
Alt (m)	Measured	Simulated	Ratio	Measured	Simulated	Ratio	Measured	Simulated	Ratio
287	1.80	1.80	1.00	2.73	3.46	0.79	541	348	1.56
460	1.65	1.84	0.90	2.12	3.50	0.61	468	334	1.40
575	1.94	1.71	1.13	1.96	2.85	0.69	556	314	1.77
Overall	1.79	1.79	**1.00**	2.28	3.28	**0.70**	521	332	**1.57**

Table 1: Mean flux densities calculated using the flux-dispersion method, given for each transect and taken over all three transects. The ratios between measured and simulated flux densities are all given.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

	CO		CH_4		CO_2	
	Mean	1σ	Mean	1σ	Mean	1σ
Flux (kmol s^{-1})	0.178	0.006	0.182	0.009	44.7	1.2
NAEI emissions (kmol s^{-1})	0.079	N/A	0.149	N/A	14.5	N/A
Ratio	**2.27**	0.07	**1.22**	0.06	**3.08**	0.08

Table 2: Bulk fluxes calculated using a conventional mass balance technique and corresponding NAEI emissions, aggregated over the Greater London administrative region. The ratio of mass balance flux to NAEI emission is also given.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

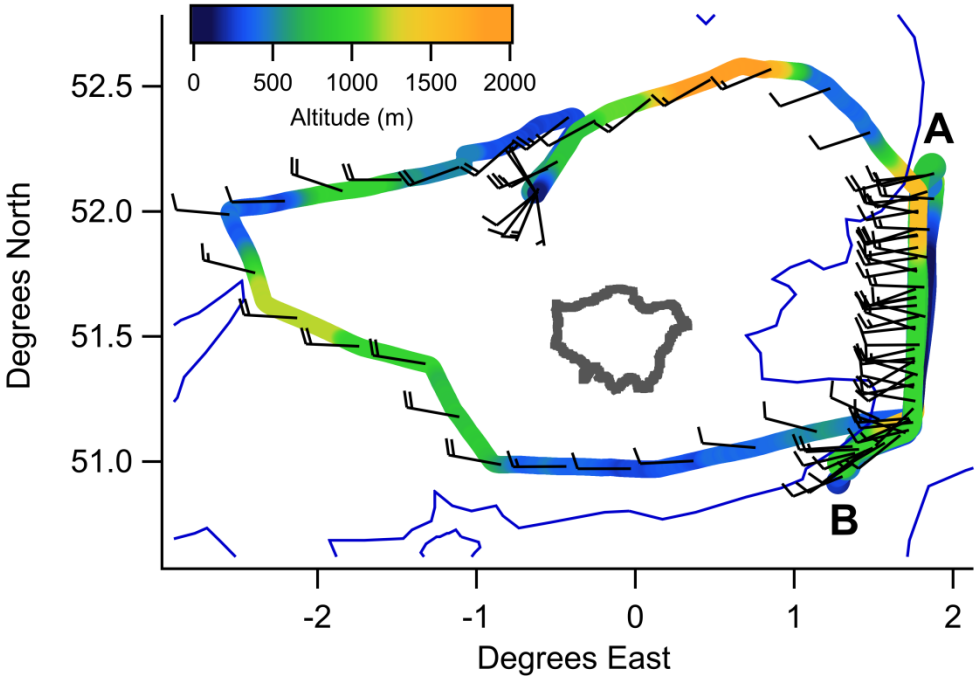

Figure 1: Aircraft flight track on 04 March 2016, coloured by altitude. Wind barbs are used to represent wind speed and direction, averaged over 5 minutes, using the convention where each full wind barb represents a wind speed of 10 knots. The border of the Greater London administrative region is shown in grey for reference.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

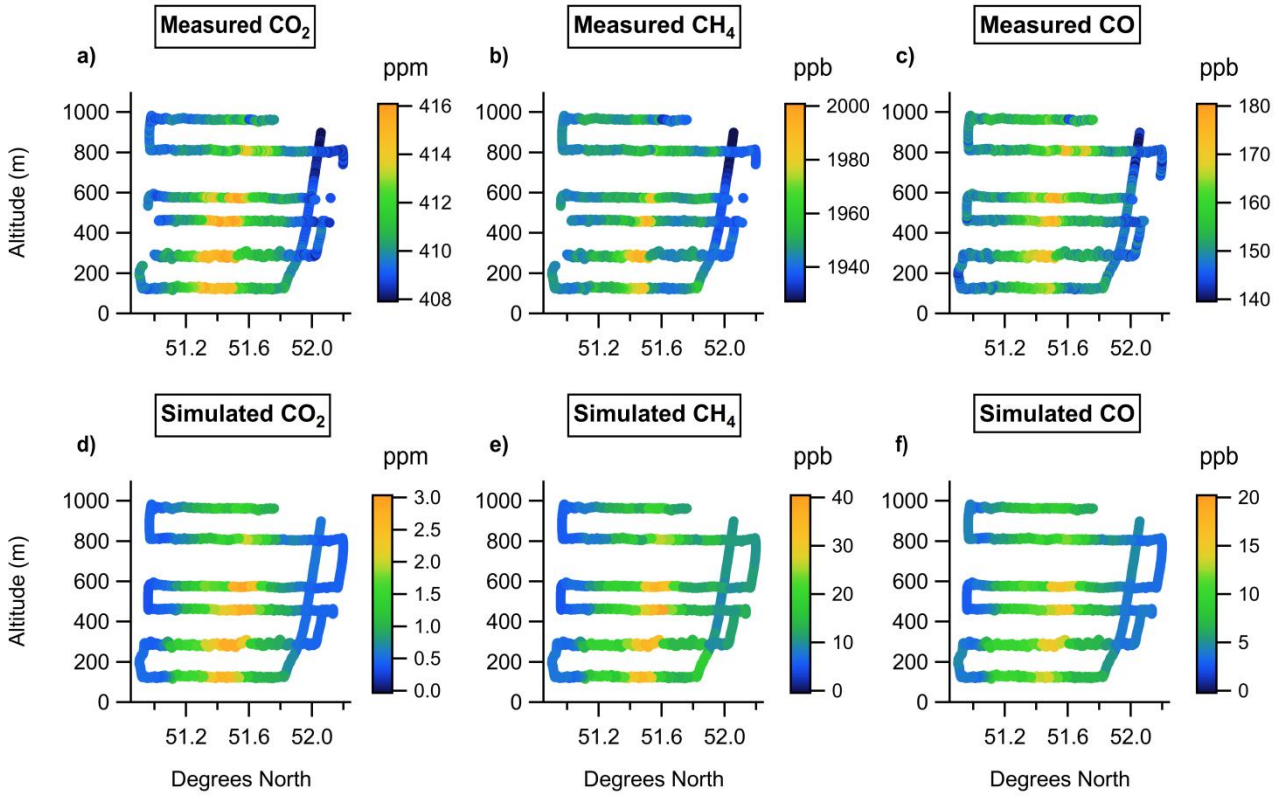

Figure 2: Altitude-latitude projections of measured mole fraction (a – c) and simulated (d – f) mole fraction enhancement downwind of London for each species.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

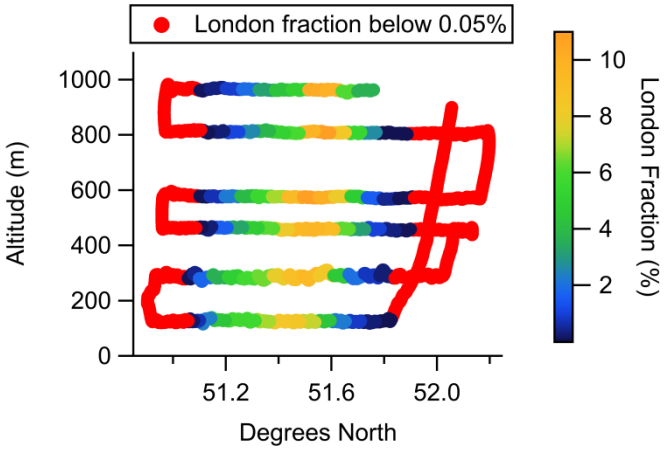

Figure 3: Altitude-latitude projection showing the influence of London on the downwind sampling, as determined from the NAME air histories. The colour scale represents the fraction of aggregated NAME air history D_{ij} within the Greater London administrative region for each NAME release period. Background periods, where the London fraction is less than 0.05%, are shown in red.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

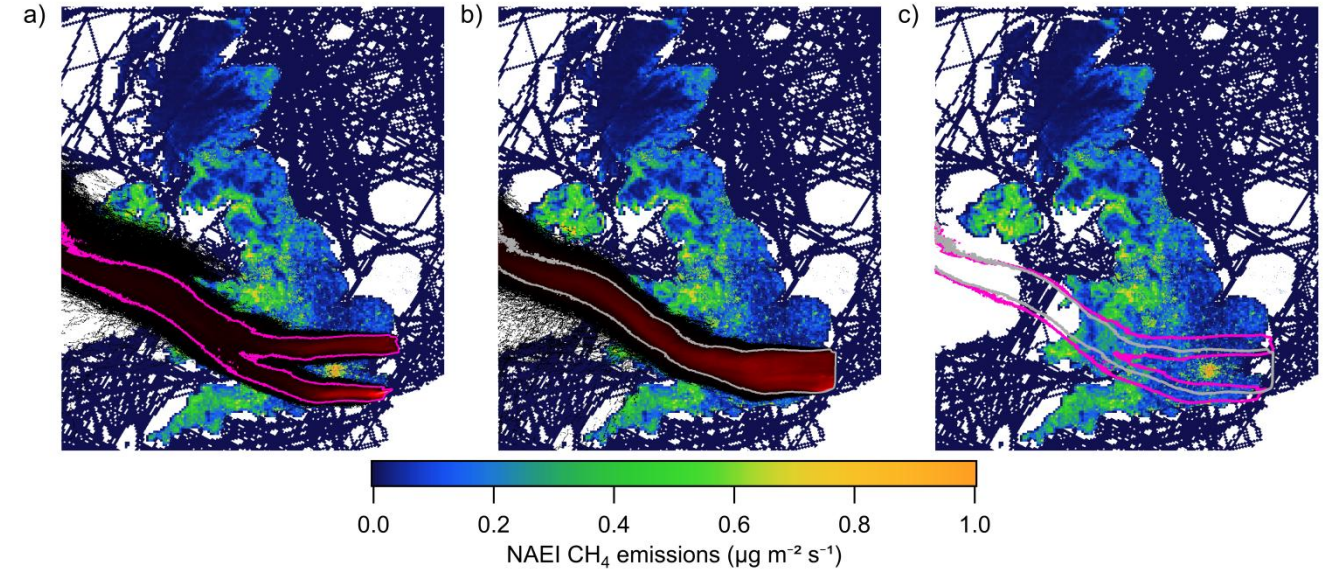

Figure 4: NAME air histories aggregated over (a) background sampling periods, and b) in-plume sampling periods, overlaid on an NAEI emissions inventory map for CH_4 (shown using a saturated colour scale). Both air histories have been normalised such that they sum to 1, with grey and pink contours shown in each plot surrounding the vast majority (99.9995%) of sample influence. These contours are included in panel c) to provide a better visual comparison between the two aggregate air histories in the context of the inventory emissions.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

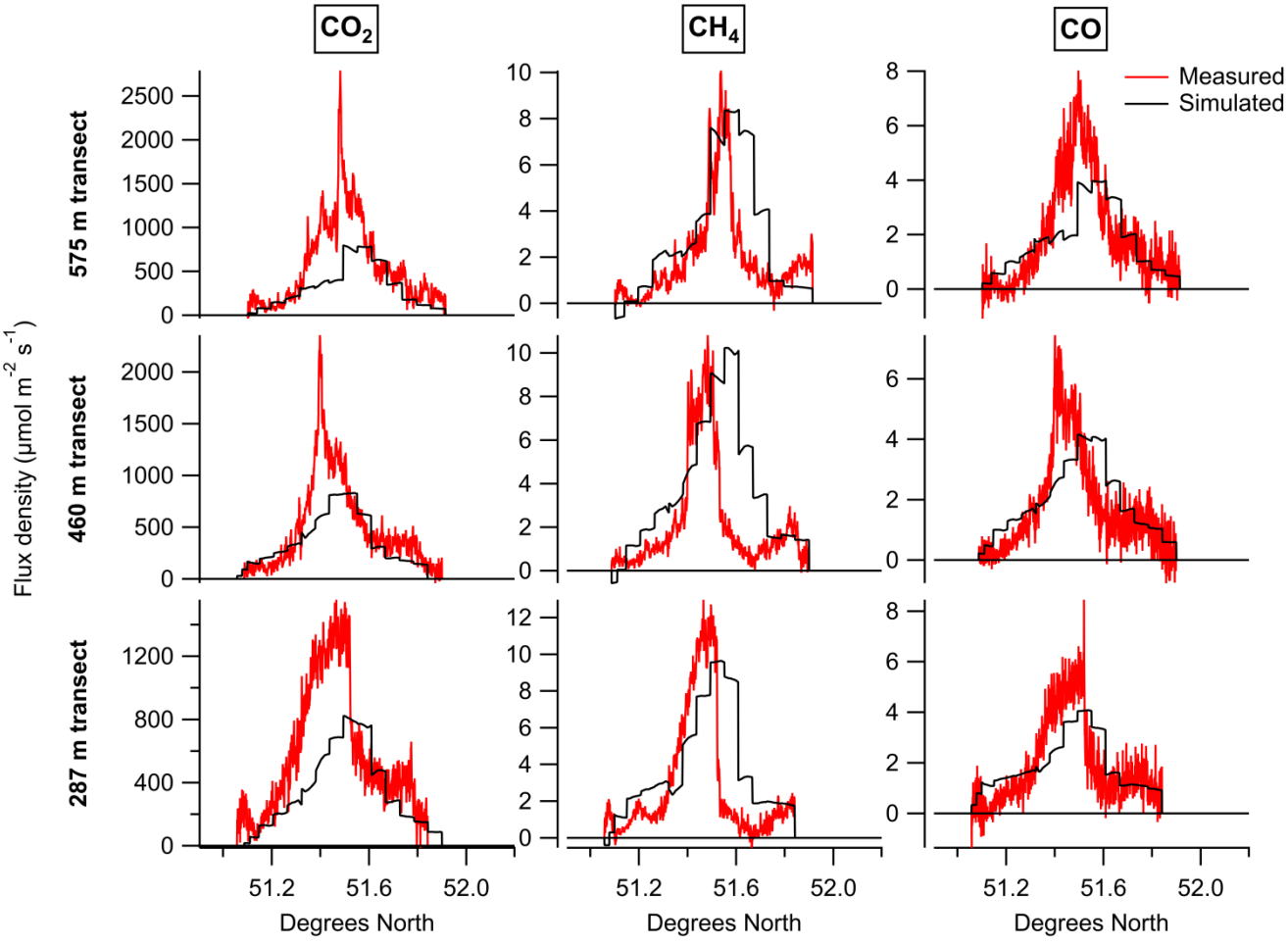

Figure 5: Measured and simulated flux densities for CO_2, CH_4 and CO, given for the three transects (287 m, 460 m and 575 m) used to assess inventory accuracy.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1033
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 October 2018

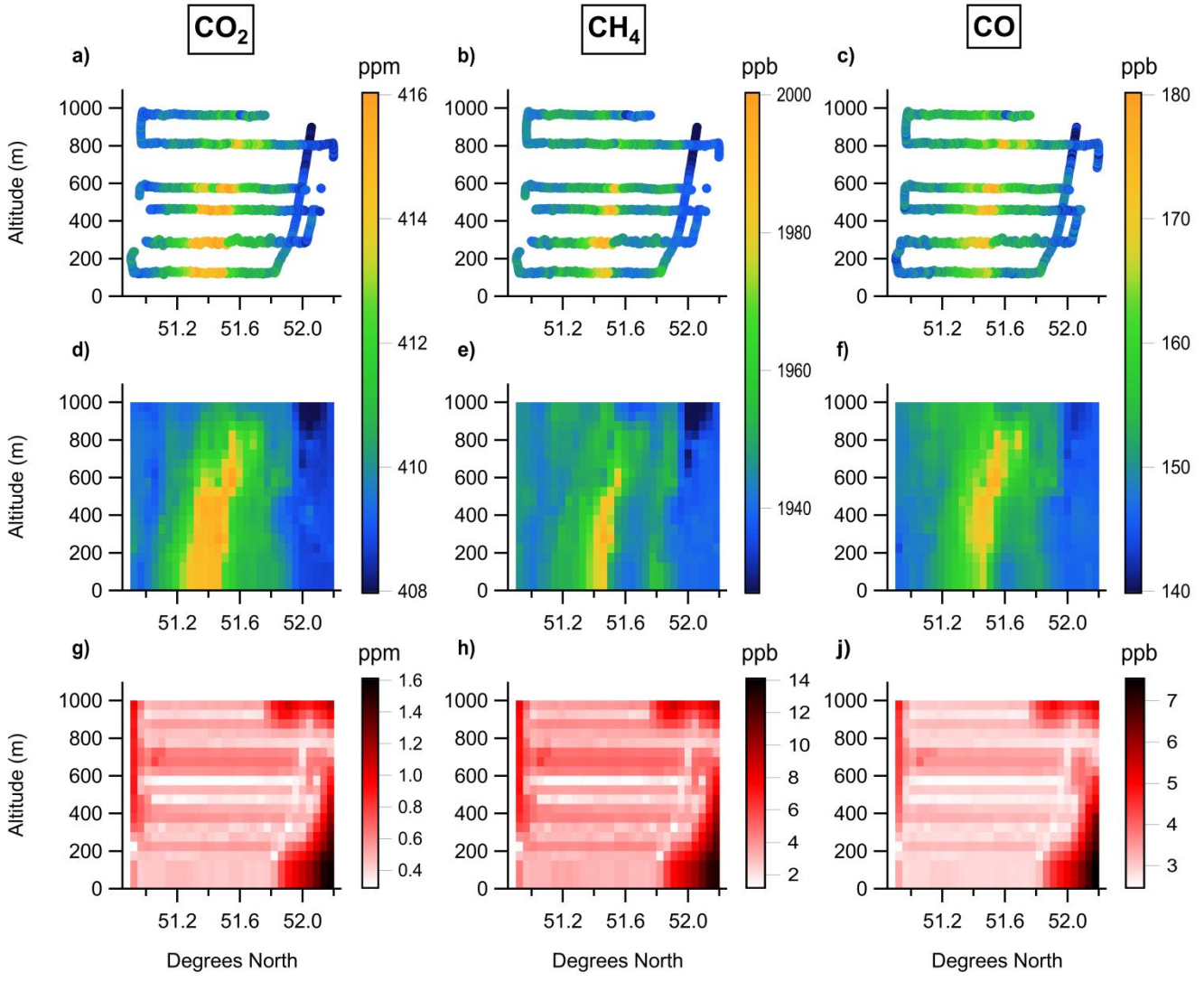

Figure 6: Altitude-latitude projections of: a) – c) measured data, d) – f) kriged data, g) – j) kriging standard error, for CO_2, CH_4 and CO respectively.

5