Answers to reviewers for "Estimating CH4, CO2, and CO emissions from coal mining and industrial activities in the Upper Silesian Coal Basin using an aircraft-based mass balance approach"

We would like to thank the two reviewers for the suggestions to improve the manuscript. Below you find our answers to their comments. The reviewer's comments are written in normal font, our answers in italics.

Review 1 by Zachary Barkley

The authors describe 2 mass balance flights performed during a single day over a coal basin in Poland. Despite only having data for a single day, the analysis done is extremely thorough for this type of study, providing an extra layer of confidence to the overall solution, and serve as a sanity check for bottom-up inventory estimates of CH4 from coal in the region (and helpful guidance for other trace gases). Nothing about the results are particularly remarkable, but it's good, necessary science nonetheless, and well-written at that. In its current state, I have no objective with publishing this paper after some extremely minor revisions are addressed.

Minor comments

Line 110: "*Since the morning is not an ideal time for the in situ mass balance method because of the growing convective planetary boundary layer, we consider the estimate from the afternoon flight to be more reliable. However, we describe the morning flight as well and consider its results as additional information."* The later, afternoon loop tends to be more reliable for a number of reasons (morning blobs everywhere!). When I read this statement, I was bracing myself for crazy signals that cast doubt on the entire study. But such a thorough job regarding capturing the signal and meteorology from the first flight (as well as the period before it) that I would argue you're underselling your loop 1 results with this comment. To me, loop 1 and loop 2 together provide pretty good confidence in your calculated total for the day. I'd consider dropping these sentences entirely.
We have deleted these sentences.

Line 126: It would be helpful to mention here the local time relative to UTC time. – *Done.*

Line 207: The background downwind method also requires the assumption that there are no sources upwind of your area of interest that would create a complex concentration pattern flowing into your domain. With that said, you have an upwind here and it's pretty clean, so it's obviously not an issue here.
We added the following sentences: "The downwind method also requires that there are no sources upwind of the area of interest which would create a complex concentration pattern flowing into the domain. This is shown in our upwind flight transect."

Line 385: I just wanted to say I appreciate you mentioning the overnight winds, because so many mass balances neglect possible accumulation from stagnant winds in the overnight/dawn hours, leading to massive enhancement blobs scattered throughout the observations.
Thank you for this comment!

Minor comments bonus points:
This study does a remarkable job ensuring the validity of the mass balance technique by performing multiple vertical transects and even driving underneath the flight path to capture the signal at the surface layer. Many mass balance studies do not put in this level of effort, and it would be good to know how necessary these extra precautions are with regard to calculating the true emissions. Furthermore, if we're going to use mass balance techniques at any point to verify emissions from a policy enactment standpoint, we'd want to be as efficient as possible with resources. So what I'd be curious about is, if you took the central transect from each loop and calculated the emissions using the simple assumption of a perfectly mixed boundary layer, low different would your solution be compared to your kriging results? Such a simple comparison would be useful to have in your uncertainty analysis and increase the scientific impact of your findings.
We did this analysis and added it as Section 3.5.

3.5 Single transect emission estimates

"This detailed calculation of the emissions can help to understand uncertainties of a mass balance in cases where less information is available. We ensured the validity of the mass balance technique by performing multiple vertical transects and even driving underneath the flight path to capture the signal at the surface layer. Many mass balance studies do not put in this level of effort, but it would be good to know how necessary these extra precautions are with regard to calculating the true emissions. Furthermore, when using mass balance techniques at any point to verify emissions from a policy enactment standpoint, we need to be as efficient as possible with resources. So, using the single transects within the boundary layer from each flight we calculated the emissions under the simple assumption of a perfectly mixed boundary layer. The PBLH was kept constant for all transects. Figure 1 shows the results of the single transect mass balance calculations for the two flights on June 6, 2018. The average of the single transects (blue) is always well within the uncertainty range of the kriging mass balance results (red). Nevertheless, single transect emission estimates deviate up to 40% in both directions from the kriging estimate for CH_4. This deviation is much larger than the kriging estimate uncertainty. Deviations are largest for transects close to the PBLH when the concentration gradient between the boundary layer and free troposphere is also large, e.g. the highest CH_4 transects. Thus, when calculating emissions from single transects the flight altitude should be well below the PBLH to avoid sampling free tropospheric air masses. On the other hand, these results discourage single transect mass balance estimates anyway."

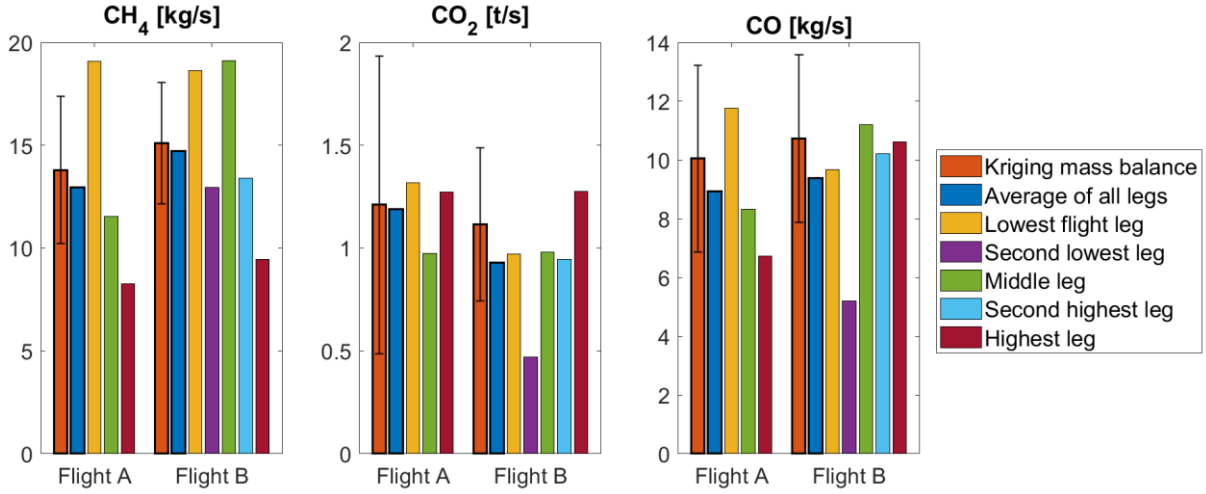


Figure 1: Mass balance results for single transects compared to the average of all single transects and the kriging mass balance result from Section 3.4.

We also added the following sentence to the Conclusion:
"The calculation of emission estimates from single flight transects is not advisable, because the single transect estimates showed deviations from their mean and the kriging method of more than 40% in both directions."

Grammar
Line 43: change "affect" to "affecting". – *Done.*
Line 473: "don't" to "do not". – *Done.*

Review 2 by Anna Karion:

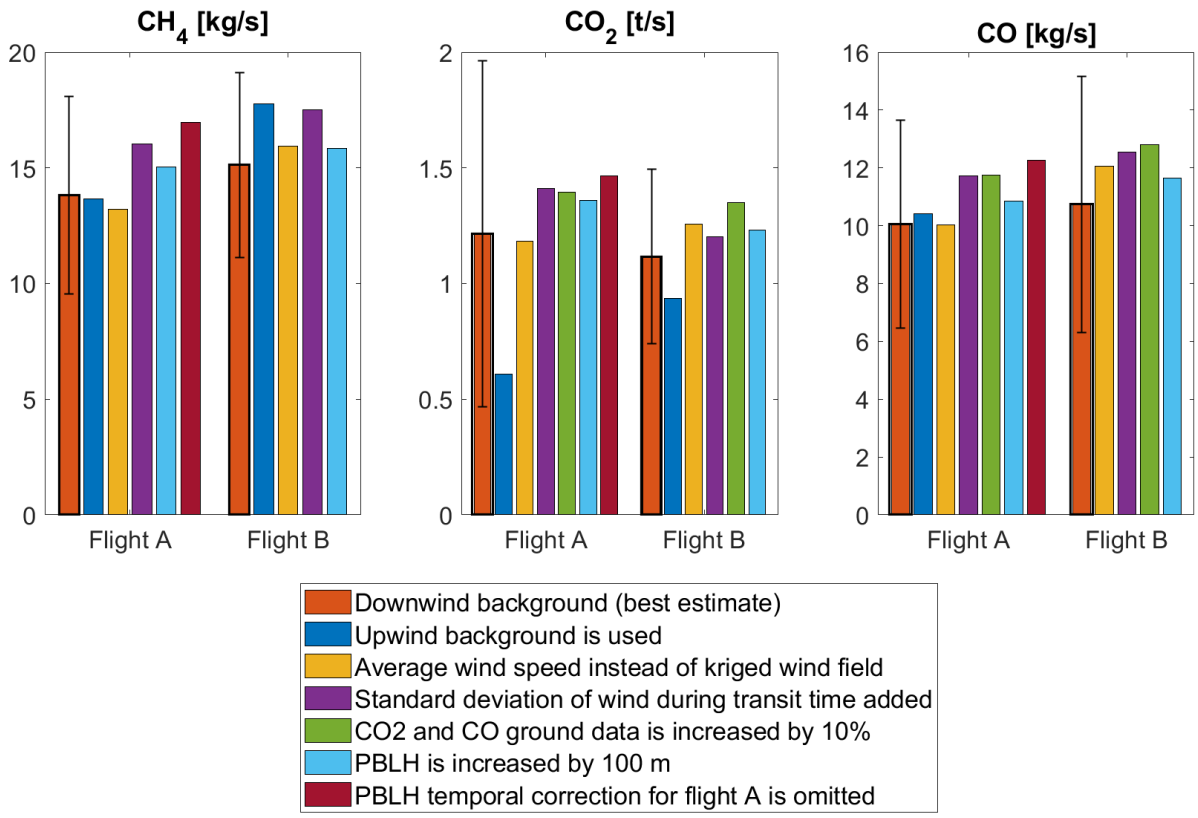
This is a well-written, well-researched manuscript and should be published in ACP after 2 major corrections noted below. Overall, the authors have done a lot of work on this flight data (although only 2 flights). But especially given that there are only two flights, there are two additional revisions that should be considered.

1) Characterization of the wind conditions prior to the flights. This is done qualitatively ("relatively steady"), but not quantitatively, and can be used in the sensitivity analysis and uncertainty calculation. The mass balance

requires the wind to be constant (Eulerian equation). If it is not constant over the transit time, there is uncertainty.

Thank you for this comment! It is true that the wind prior to the flight is very important for the mass balance. Here we are in the advantageous position of having wind profile measurements of the region for the entire day. We added the wind speed during the four hours prior to the downwind sampling to Table 3. As suggested, we added another sensitivity analysis to the uncertainty calculation, using the standard deviation of the wind lidar wind speed measurements in the four hours before the downwind sampling as temporal wind speed uncertainty estimate. We added a paragraph in Section 3.4 describing this sensitivity test and updated Figure 7 to include this uncertainty measure: "One assumption for a mass balance calculation is that the wind is constant during the time it takes for the gases to be transported from the emission source to the observation location. In reality the wind field is subject to considerable variability. In our case we were able to assess this temporal variability from the wind lidar observations. To account for wind variability, we calculated the standard deviation of wind speed within the boundary layer and added it to the kriged wind field used in the mass balance calculation. This introduced an uncertainty of 17% and 15% to the morning and afternoon flight results, respectively."

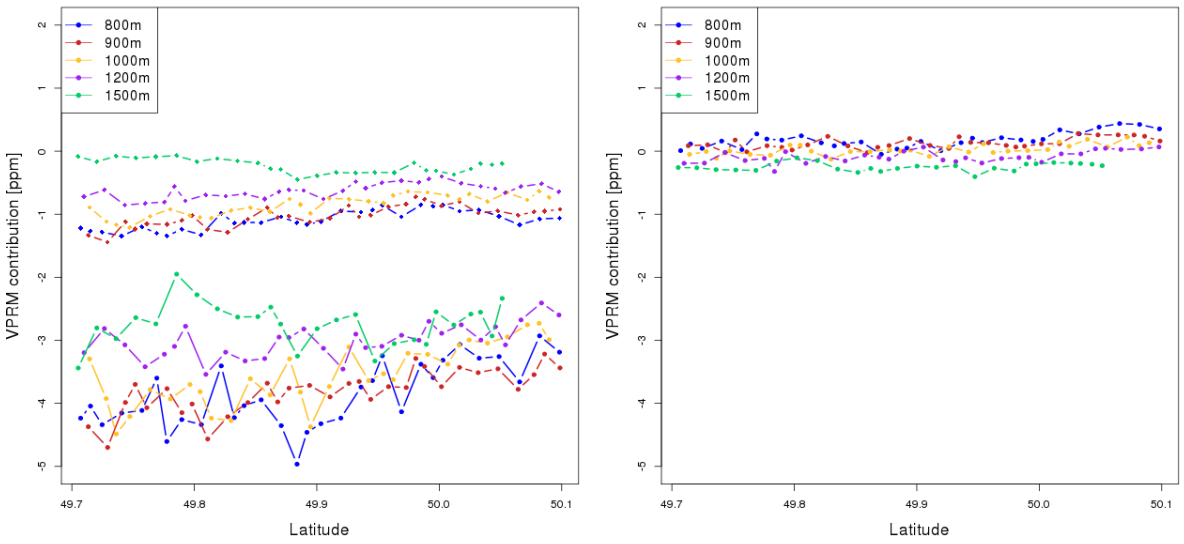
The uncertainty due to the wind variability, here in purple bars, is proportional to the wind speed uncertainty, because the wind speed is a linear factor in the mass balance equation. The addition of this sensitivity test increased the uncertainty of the mass balance estimates. The new uncertainty values have been updated throughout the manuscript.



2) Use of biosphere model in sensitivity analysis. This is less important to the final result, but in the sensitivity analysis using the upwind transect as the background condition, the method of use of the biosphere model is confusing. The authors only subtract the biospheric influence one hour prior to measurement, because the upwind measurement was made one hour prior. This needs more explanation to me at least. However, since this method is only used to look at the sensitivity of background choice, it is a more minor issue than the above. More discussion on both these points is below, with other detailed/minor comments.

The biosphere model VPRM calculates the contribution of biogenic sources to the CO_2 mole fractions in the atmosphere, by parameterizing the gross ecosystem exchange and ecosystem respiration fluxes. In our case we calculate contributions to downwind observations because when using an atmospheric background calculated

from the upwind flight track, this biogenic uptake changes the background relative to the enhancements caused by anthropogenic CO_2 emissions. Ideally, a mass balance flight should sample the same air mass upwind and downwind of the emission sources. Then the contribution of the biogenic sink could be inferred from the trajectories between the upwind and downwind flight track as a sum over the entire transit time.
We acknowledge that by using the model in the manner described we have added a layer of complexity to a phenomenon already difficult to begin with. Using the model, in theory, is supposed to allow us to do more than just grossly estimate biospheric influence assuming heterogeneity in space and time - we also aimed to take into account its spatio-temporal variability. We believed that for our study they might have been of some importance, as the area has many fragmented forests intertwined with substantial urban- and rural-type areas.
*As was mentioned above, the optimal use of the model in the method described would require for the upwind track to be flown in exactly Lagrangian manner. Specifically, this could be done by either 1) flying at the *same upwind location*, however four hours before the downwind measurements were taken, or 2) flying at the *same time* (i.e. as was in fact flown), but closer to the downwind track at the time (along the red line from Fig S2). Each of those options would allow then for a direct link between the simulated mole fractions and those measured aboard the aircraft, i.e. both values would be available at the same time and location, which would, in theory, minimize the errors caused by the inhomogeneities in the spatial distribution of the biospheric sources. Neither of those strategies was adopted, and even in case they were, we would still have to face other challenges: for scenario 1) measurements done earlier, during PBL development, tend to be poorly represented in the models, for scenario 2) the area constrained by the mass-balance method would not have encompassed the full study area – c.f. Fig S2.*
We have thus decided to adopt a hybrid approach, in which we assume that we can still link the measurements to our model quasi-directly, despite the fact that the model results are simulated for a location several tens of kilometers away from the actual upwind measurement location. It should be noted that it is quietly assumed here, that the biospheric fluxes are spatially homogeneous, i.e. fluxes in the area between the downstream transect and the location of the trajectories one hour before are similar to fluxes in the area between the location of the trajectories one hour before and the upstream transect. This explanation has been added to Section 2.2.2.



In fact, the above figures show that there is very little difference: The left figure shows the biospheric signal accumulated along the 1 hour back-trajectories for each receptor point along the downwind transect at different heights (small diamonds, values between 0 ppm and -1.5 ppm) together with those accumulated along the full trajectories between upwind and downwind transect (values between -2 and -5 ppm). Note that almost all of the differences between these are related to the longer duration of the full trajectories of 3.5 hours on average. The right figure shows the differences between 1 h and full-period accumulations after normalization (i.e. the VPRM CO_2 contributions are divided by the duration of the trajectory). This clearly shows that there are very small differences in fluxes over the area along the 1 h trajectories and along the full trajectories connecting upwind and downwind transects.

L23: "estimates... which are well within the range"... – *Done.*

L60: e.g. reads awkwardly, I would write "which are used in climate projections, for example." – *Done.*

L87 should say "using an airborne eddy covariance" (or maybe "the airborne eddy covariance..."). – *Done.*

Figure 1: Could there be an inset to show the larger map where this is located? The google earth image does not actually show any location boundaries etc, or lat/lon indicators. – *Done.*

L102: make and model or reference for "well-established" CRDS? Were all 3 gases measured by both instruments? [I see now this is further described in Section 2, so perhaps just refer here to that section].
We removed "…using a well-established cavity ring-down spectrometer (CRDS) and a modified quantum and interband cascade laser spectrometer (QCLS, Kostinek et al., 2019)." Since this is elaborated in Section 2.

L108 - this is a very nice aspect of this study that is usually not done! – *Thanks!*

L112: Perhaps Chapters should be "Sections"? (this is up to the editors). Later the text does refer to "Sections", so consistency would be good. – *Done.*

L124: Earlier several remote sensing instruments were mentioned, but this work only focuses on the in-situ measurements. Is there a reason for this?
During CoMet a wide range of instruments were employed. Each of the methods delivered individual results that will all be published within the CoMet Special Issue. Including remote sensing into this study would have significantly enlarged the manuscript since there are several issues that need to be addressed. Nevertheless, a comparison of in situ and remote sensing observations is also planned.

L138 I don't think there should be a comma after both. – *We deleted this.*

L165-166: The wind speed (and direction) also must remain constant over the transport time. This is a key assumption that the flux measured out of the downwind plane (or "wall") is equal to the emissions flux from the surface. Variability from this assumption is likely to happen, so that should be accounted for in the uncertainty analysis. Using wind speeds measured by the aircraft in the afternoon may not be the correct approach if they are not representative of the wind field over the whole domain over the transit time. If downwind wind is used, it should be shown that the wind speed was constant (or what the variability was) through the time it took for the air to transit the domain. Fig 6(a) in Karion et al., ACP, 2019 (https://doi.org/10.5194/acp-19-2561-2019) shows (granted, in an extreme case with long residence time) the difference between true emissions and measured emissions in a downwind plane using a forward model. Given the relatively small domain here, it's likely fine to assume steady winds over transport time, but this should be stated (and the variability of the wind included in uncertainty/sensitivity analysis).
We included this uncertainty. See our comment above.

L211, using the upwind as a "check" has been done before, citing some literature here would strengthen the justification of this choice.
We added another sentence on the upwind transect and a reference in the manuscript.

L229 "form" should be "from" – *Done.*

Section 2.2.2. It is not clear to me why only the last hour of the footprint is used to estimate influence of biogenic fluxes on the downwind transect. For the final mass balance, the edges of the plume are used as the background, so I don't understand why the time between the upwind and downwind sampling matters. For the case of using the upwind transect, I am still not clear on the use of the 1 hour time frame. More on this in the next section.
We added a clarification that the biogenic uptake is only necessary for the use of an upwind background mole fraction. Please also see our answer to your general comment on this matter.

L264: Very nice detailed error analysis: This is where I believe that you should include the variability of the wind over the transport time, not just measurement error. Perhaps added in quadrature to the 0.3 m/s, and constant over the whole grid. Or, now reading 2.3.2, perhaps it should be considered in the sensitivity portion, as it would be a systematic error.
We included it as a sensitivity test.

L283: Table 1 repeated. – *Corrected.*

L328: no comma after 'cases'. – *Corrected.*

L388: Agree! This addresses my issue above about steady winds over the transit time - so the section 2.2 above should mention this as well. In table 3, I would recommend also showing the mean wind and standard deviation for the Lidar over the 3-4 hours transit time, not only the flight time (to address my earlier comment). Several times this section says "relatively steady" - using the lidar data, this can be quantified and stated here.
Please see our answer above.

L429-435: This is one issue with Kriging in space samples that were conducted at different times! It seems that hopefully the correction did the job, but in general, one might think this is a reason to do separate mass balance calculations for each transect (with each its own PBL depth), and then average. But either way, changing conditions are difficult to deal with, and the authors have made a decent attempt at accounting for this, so no need to do this.
We now added an extra-study where we used the individual tracks using the same PBLH for all of them. Results are listed in the new Section 3.5.

Figure 5: These are nice useful figures. What is the averaging time prior to Kriging? Or is it the native 0.5-Hz data that is Kriged?
The averaging time prior to Kriging is 20 seconds. This is stated in the method section 2.2

L490. To clarify the method for using VPRM above, is it not accounted for at all in the version where the edges are used as background, and only when the upwind background is used? – *Correct, the VPRM model is only used for the upwind background.*
This makes sense if the assumption is that the uptake is the same at the edges and in the center, which in this case is a reasonable assumption (maybe not in an urban area). Does the enhancement from VPRM not change across the transect? That would then justify not including it when using the edges as background.

The VPRM implementation used in our study did, in fact, take into the account the variable land-use and vegetation types (including also urban areas) in the area of influence of our observation. However, the resulting simulated uptake signal only shows a small trend along the transects as can be seen in this figure. Thus, we calculated the downwind background for both edges and used an average of these values to account for the gradient.
Please also note our answer to the general comment above.

Regarding the upwind background use of VPRM: I am still confused by the use of one hour only, and given how variable the uptake is in time due to the diurnal cycle of fluxes, it's not clear how to deal with this. If biogenic fluxes were constant in time and space, then the full trajectories (footprints) should be used until the location of the upwind leg (…).
This is not exact. Even if the fluxes were constant in space and time, then applying VPRM over the trajectory connecting upwind and downwind points would still simulate a too large uptake. To visualize this, we can imagine our study area as a part of an infinite, perfectly flat and perfectly homogeneous forest. Over such an area the mole fractions of CO_2 would be independent of space (as much as stochastic mixing allows, naturally) – but would still retain its time-dependence, as the diurnal biospheric cycle still works. Doing the same measurements of mole fractions over our flight tracks, we would still have a single hour of temporal difference in

observations, and four-hour difference in the air-mass flow between measurement locations, during which the biosphere was able to uptake more CO_2. In such a scenario, if we would use the model and predict the change of mole fraction as suggested by the reviewer, we would have simulated a drop of mole fraction that was caused by four hours of uptake, which is incompatible with the measurements that we did. In this case we should again compare against influence simulated from a single hour. The simulated mole fractions would perfectly fit the observations thanks to spatial homogeneity.
However, in more realistic scenarios of inhomogeneous biospheric fluxes, the spatial decoupling between observations and simulated single-hour trajectories will cause errors that might be more challenging to quantify. This is in fact the situation that we are dealing with here – see the discussion and figure under the general comment above.

(*continued*) Otherwise, I think this is too complex to handle in the method it has been handled here. If the upwind leg were flown one hour *after* the downwind legs then how would this work? I am wondering if the upwind transect just cannot be used other than to show the lack of significant upwind fossil sources. Please explain/justify the use of the one hour time frame, and how this would work if the upwind leg were flown simultaneously or after the downwind leg - would you not subtract VPRM at all in that case? Seems wrong.
In case the upwind track would have been flown one hour after the downwind sampling, the CO_2 background mole fractions in the boundary layer would have further decreased due to continued biogenic uptake. In this case the upwind background observations would be even lower than the downwind background and would need to be corrected with one hour of biogenic uptake. Using the model to estimate this correction becomes meaningless, as we cannot then establish a link between the downwind and upwind mole fractions using our model framework.

L515: Definitely agree with this statement on biosphere-atmosphere fluxes being complex! – *Thanks.*

L520 (sensitivity using average wind speed during downwind legs): This is close to what I suggested earlier. However, this is the average over the flight time, not transport time. One could use the transit time instead (from lidar perhaps, or from the model), as was done by Karion 2013 & 2015, to see the effect. What matters more is what the wind was at the location and time of emission - if it was high, then less CH4 was picked up in that air mass, if low, vice-versa.
We added a sensitivity test using wind variability during the transit time.

SI: Table 1 can the caption also explain the offset of local time from UTC so the reader can quickly translate the UTC hours to local? – *Done.*

How frequent is "frequently calibrated" (approximately - daily? hourly?). Text S1 gives the mole fractions of the cylinders in ppm for both gases, but the text says ppm/ppb. Should be ppm/ppm.
The drift with time uncertainty is estimated using the flight time, so presumably the calibrations did not occur during flight, so frequently is > 2.5 hours?
We calibrated after every second flight. This has been added to the text.

Estimating CH_4, CO_2, and CO emissions from coal mining and industrial activities in the Upper Silesian Coal Basin using an aircraft-based mass balance approach

Alina Fiehn[1], Julian Kostinek[1], Maximilian Eckl[1], Theresa Klausner[1], Michał Gałkowski[2,3], Jinxuan Chen[2], Christoph Gerbig[2], Thomas Röckmann[4], Hossein Maazallahi[4], Martina Schmidt[5], Piotr Korbeń[5], Jarosław Nęcki[3], Pawel Jagoda[3], Norman Wildmann[1], Christian Mallaun[6], Rostyslav Bun[7,8], Anna-Leah Nickl[1], Patrick Jöckel[1], Andreas Fix[1], Anke Roiger[1]

[1] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2] Max-Planck-Institut für Biogeochemie (MPI-BGC), Jena, Germany
[3] Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Kraków, Poland
[4] Institute for Marine and Atmospheric research Utrecht, Utrecht University, Utrecht, The Netherlands
[5] Institute of Environmental Physics, University of Heidelberg, Heidelberg, Germany
[6] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Flugexperimente, Oberpfaffenhofen, Germany
[7] Department of Applied Mathematics, Lviv Polytechnic National University, Ukraine
[8] Faculty of Applied Sciences, WSB University, Dąbrowa Górnicza, Poland

Correspondence to: Alina Fiehn (alina.fiehn@dlr.de)

Abstract. A severe reduction of greenhouse gas emissions is necessary to reach the objectives of the Paris Agreement. The implementation and continuous evaluation of mitigation measures requires regular independent information on emissions of the two main anthropogenic greenhouse gases, carbon dioxide (CO_2) and methane (CH_4). Our aim is to employ an observation-based method to determine regional-scale greenhouse gas emission estimates with high accuracy. We use aircraft- and ground-based in situ observations of CH_4, CO_2, carbon monoxide (CO), and wind speed from two research flights over the Upper Silesian Coal Basin (USCB), Poland, in summer 2018. The flights were performed as a part of the Carbon Dioxide and Methane (CoMet) mission above this European CH_4 emission hot spot region. A kriging algorithm interpolates the observed concentrations between the downwind transects of the trace gas plume and then the mass flux through this plane is calculated. Finally, statistic and systematic uncertainties are calculated from measurement uncertainties and through several sensitivity tests, respectively.

For the two selected flights, the in situ derived annual CH_4 emission estimates are 13.8 ± ~~34~~.~~6~~ 3 kg/s and 15.1 ± ~~34~~.0 kg/s, which ~~is~~ are well within the range of emission inventories. The regional emission estimates of CO_2, which were determined to be 1.21 ± 0.~~72~~ 75 t/s and 1.12 ± 0.~~37~~ 38 t/s, are in the lower range of emission inventories. CO mass balance emissions of 10.1 ± 3.~~2~~ 6 kg/s and 10.7 ± ~~24~~.~~9~~ 4 kg/s for the USCB are slightly higher than the emission inventory values. The CH_4 emission estimate has a relative error of ~~21~~26-~~26~~31%, the CO_2 estimate of ~~33~~37-~~60~~62%, and the CO estimate of ~~27~~36-~~32~~41%. These errors mainly result from the uncertainty of atmospheric background mole fractions and the changing planetary boundary layer height during the morning flight. In the case of CO_2, biospheric fluxes also add to the uncertainty

and hamper the assessment of emission inventories. These emission estimates characterize the USCB and help to verify emission inventories and develop climate mitigation strategies.

1 Introduction

One of the main objectives of the Paris Agreement is to keep the global temperature rise well below 2°C compared to pre-industrial levels (UNFCCC, 2015). This ambitious goal can only be reached by a severe reduction of greenhouse gas emissions. The development of efficient mitigation strategies and the implementation and management of long-term policies requires consistent, reliable, and timely information on emissions of the two main anthropogenic greenhouse gases, carbon dioxide (CO_2) and methane (CH_4). Carbon monoxide (CO) can be used as an additional tracer for comparison with emission inventories and as proxy for CO_2 from fossil fuel combustion. It is produced from the incomplete combustion of fossil fuels and biomass and reacts with the hydroxyl radical (OH), thus ~~affects~~ affecting the main sink of CH_4.

The globally averaged atmospheric abundances of CO_2 and CH_4 have increased by 47% to 407.8 ± 0.1 ppm and by 159% to 1869 ± 2 ppb, respectively, in the period 1750 to 2018 (WMO, 2019). The relative contribution of individual sources and sinks to atmospheric CH_4 is still highly uncertain and the factors that affect these sources and sinks are not fully understood (Saunois et al., 2019). After a period of stable mole fractions since 2000, the atmospheric abundance of CH_4 has started to increase again in 2007, and after 2014 the increase intensified yet again (Nisbet et al., 2014; Nisbet et al., 2016). The reason for this increased growth is currently investigated in several studies, which partly contradict each other by discussing biogenic sources, fossil fuel emissions and/or a decrease in the OH sink (Hausmann et al., 2016; Schaefer et al., 2016; Saunois et al., 2017; Turner et al., 2017; Worden et al., 2017; Nisbet et al., 2019).

Atmospheric emission inventories for trace species are usually based on *bottom-up* data-based approaches. Here, emissions for individual facilities, sectors, or sources are compiled into a comprehensive database. If direct emission data is not available, they are often calculated using activity data, like the mass of coal extracted, together with emission factors. For Annex I countries, sector specific emissions of greenhouse gases have to be reported annually under the United Nations Framework Convention on Climate Change (UNFCC). Other countries are encouraged to report national totals of emissions. Bottom-up inventories can thus include single-source emissions, national totals, or can be disaggregated on different spatial scales. These gridded emission inventories commonly use national emission totals and distribute them across each country using proxy data like population density or single facility locations. This method is used to compile emission inventories, which are used ~~e.g.~~ in climate projections, for example. The neglect of regional differences and the uncertainties in the proxy data and emission factors introduce high uncertainties into the emission inventories at grid cell level (Janssens-Maenhout et al., 2019). Without accurate emission estimates it is challenging to create reliable future climate projections and develop efficient mitigation strategies.

Therefore, there is a strong need for an independent and objective verification of emissions from individual sources or source regions based on atmospheric observations, usually referred to as *top-down* approaches. Top-down studies based on satellite data provide information on global and regional scales. For methane, emission quantification of individual sources has

recently been demonstrated on very large point sources (Pandey et al., 2019; Varon et al., 2019), but quantification of smaller sources is still difficult. Here, airborne measurements reveal more detailed insights on smaller scales, because in situ measurements allow the study of emission sources with high spatial resolution and accuracy. High precision measurements of atmospheric concentration can be used for the top-down estimation of emissions from specific regions or sectors using atmospheric inversion models (Gurney et al., 2002; Thompson et al., 2014; Bergamaschi et al., 2018), and for the validation of numerical models used to calculate atmospheric abundances based on bottom-up emission inventories (Krinner et al., 2005; O'Shea et al., 2014). Airborne measurements provide highly valuable data for an independent assessment of anthropogenic CH_4, CO_2 and CO emissions, because the majority of these emissions originate from a small fraction of the globe, namely fossil fuel exploitation facilities, cities, and power plants. Airborne measurements have shown to be useful in emission assessment of anthropogenic emissions from several sectors, including landfills (Cambaliza, 2015; Krautwurst et al., 2017) and oil and gas production regions (Karion et al., 2015; Yuan et al., 2015; Alvarez et al., 2018; Barkley et al., 2019). Plant et al. (2019) and Ren et al. (2018) showed that North American cities emit more CH_4 than suspected, because of underestimation of natural gas leakage or lack of inclusion of end use emissions.

Aircraft top-down approaches can be used in several ways to obtain greenhouse gas flux estimates. One way is the mass balance approach, where the emissions are estimated from observed in situ mole fractions and wind speeds in the target region. Different flight patterns are used for mass balance studies: A single downwind flight transect in the approximate vertical center of the boundary layer (Karion et al., 2013) or several transects of the plume at the same height but different distances from the source (Turnbull et al., 2011) are sufficient in case of a well-mixed planetary boundary layer (PBL). A better understanding of vertical trace gas distribution is achieved by several transects at different heights but the same distance (Cambaliza, 2015; Karion et al., 2015; Pitt et al., 2019). Single point sources or small areas can be assessed by circular flight paths at different heights (Conley et al., 2017; Tadić et al., 2017; Ryoo et al., 2019). ~~Using~~ The airborne eddy covariance technique~~,~~ can directly infer vertical fluxes ~~can be inferred directly~~ (Hiller et al., 2014; Yuan et al., 2015). Further techniques for airborne emission estimation include active and passive remote sensing instruments (Amediek et al., 2017; Krautwurst et al., 2017). All methods can be combined with inverse modelling to derive emission distributions (Kort et al., 2008; Polson et al., 2011; Brioude et al., 2013; Xiang et al., 2013; Cui et al., 2015).

This study is part of the Carbon Dioxide and Methane (CoMet) mission. The goal of CoMet is to develop and evaluate methods for the independent monitoring of greenhouse gas emissions and to provide data for satellite validation. CoMet combined a suite of airborne active (lidar) and passive (spectrometers) remote sensors with in situ instruments to provide local- to regional-scale data about atmospheric concentrations of CO_2 and CH_4 and to derive emissions on different spatial scales. One of the foci of CoMet was the Upper Silesian Coal Basin (USCB), located in southern Poland, which represents one of the largest European CH_4 emission sources with a total of around 500 kt CH_4/a (~3% of European CH_4 emissions), emitted from about 40 hard coal mines (EEA, 2019). CH_4 is released from the coal deposits and bedrock before and during mining and ventilated to the atmosphere through individual ventilation shafts due to safety reasons (Figure 1). The USCB is also a heavily industrialized urban agglomeration of >2 million inhabitants. During the CoMet mission in early summer

2018, we performed airborne in situ measurements of CH_4, CO_2 and CO aboard the DLR aircraft Cessna Grand Caravan 208B ~~using a well-established cavity ring-down spectrometer (CRDS) and a modified quantum and interband cascade laser spectrometer (QCLS, Kostinek et al., 2019)~~.

During ten research flights conducted in May and June 2018, we studied emissions from coal mine ventilation shafts, power plants and other industrial facilities in the USCB region by using an airborne mass balance approach. Depending on the wind situation, different areas of the USCB region were targeted. To account for the lower part of the emission plume not accessible by aircraft, a number of vans equipped with mobile in situ measurement systems conducted ground-based measurements in a coordinated manner. Here we present trace gas observations from the two mass balance flights targeting the emissions of the entire USCB, one in the morning and one in the afternoon of the same day, June 6, 2018. ~~Since the morning is not an ideal time for the in situ mass balance method because of the growing convective planetary boundary layer, we consider the estimate from the afternoon flight to be more reliable. However, we describe the morning flight as well and consider its results as additional information.~~ In ~~Chapter~~ Section 2 we present the observational data used in this study to derive emission estimates, a theoretical description of the mass balance method including the statistical interpolation method kriging together with the uncertainty analysis, and an overview of emission inventories available for the USCB. ~~Chapter~~ Section 3 contains the results of the mass balance flights. It includes a presentation of the meteorological situation, as well as the mass balance estimate and its uncertainties. ~~Chapter~~ Section 4 compares our mass balance emission estimate with current emission inventories. A conclusion is given in ~~Chapter~~ Section 5.

2 Data and methods

2.1 Observational data

During the CoMet 1.0 campaign several aircraft and ground based instruments were used to extensively sample greenhouse gas emissions of the USCB in early summer 2018. Here we present measurements taken aboard the DLR Cessna Grand Caravan 208B (Caravan). The Caravan was based in Katowice, Poland, from May 29 to June 13, 2018. Ten research flights were conducted in the USCB targeting different parts of the USCB. The flight paths were planned using a CH_4-plume forecast provided by the online-coupled, three times nested global and regional MECO(n) model (Nickl et al., 2019). For our estimation of entire USCB emissions, we use airborne in situ observations from two flights on June 6, 2018, one in the morning (09:22 - 11:45 UTC, 11:22 – 13:45 CEST) and one in the afternoon (13:01 - 15:28 UTC, 15:01 – 17:28 CEST), in the following referred to as flights A and B, respectively. Figure 1 shows the flight track of flight B on a map with the CH_4 emission sources. Both flights were designed in a box pattern with an upwind leg in the northeast approximately in the middle of the PBL and the downwind wall in the southwest with flight transects at several heights. CH_4, CO_2, and CO enhancements were clearly observed in the downwind wall. The flights were conducted in coordination with ground-based teams, which drove the instrumented vans below the upwind and downwind legs. Their tracks and sampled CH_4 mole

fractions for the afternoon flight are shown in Figure 1. For the emission estimation, we selected ground-based data according to closeness in time. Sampling times for flight and ground-based data are listed in Table S1.

Additionally, three Doppler wind lidar instruments Leosphere Windcube 200S were stationed at Rybnik, Wisła Mala and Krzykawka to measure vertical profiles of wind speed, wind direction and turbulence parameters (Figure 1). Details on the CoMet lidar wind measurement setup and the planetary boundary layer height (PBLH) determination are given in Wildmann et al. (2020) and Luther et al. (2019).

A sophisticated suite of instruments aboard the Caravan gathered both, meteorological parameters and trace gas concentrations. A 5-hole probe, connected to a pressure transducer, is mounted on a nose boom under the left wing of the aircraft and measured the three dimensional wind vectors. The temperature, pressure, and humidity sensors and the calibration of the wind measurement system are described in detail by Mallaun et al. (2015). A flight-ready CRDS analyzer (G1301-m, Picarro) was installed in the cabin of the aircraft. It measured CH_4, CO_2 and water vapor at a frequency of 0.5 Hz with cavity ring-down spectroscopy. Trace gas concentrations for water vapor were corrected according to Rella et al. (2013). The calibration and uncertainty assessment were conducted in analogy to Klausner et al. (2020), who used the same instrument, aircraft, and calibration technique. Details specific to the CoMet set-up can be found in the Supplement (Table S2 and Text S1). CO is measured with a modified QCLS (Aerodyne) that also records CO_2, CH_4, ethane (C_2H_6), and nitrous oxide (N_2O) (Kostinek et al., 2019). Furthermore, a dry air sampler with 12 glass flasks (1 l) was installed aboard the Caravan, which were filled during the flight and later analyzed in the laboratory at Max-Planck Institute for Biogeochemistry for trace gas concentrations and isotopic signatures (CH_4, CO_2, CO, N_2O, H_2, SF_6, $\delta^{13}C$-CO_2, $\delta^{18}O$-CO_2, $\delta^{13}C$-CH_4, $\delta^{2}H$-CH_4). However, in this study we focus only on the continuous in situ observations, while the results of ethane measurements and isotopic signatures will be published in a follow-up study.

Ground-based CH_4 data were recorded by three teams using vans equipped with different CRDS analyzers (Picarro G2201-i, AGH University and University of Heidelberg; G2301, Utrecht University). The group from the AGH University measured below the upwind leg and groups from University of Heidelberg and Utrecht University sampled below the downwind tracks. For traceability between airborne and ground-based systems, an instrument intercomparison was conducted with the same four gas cylinders.

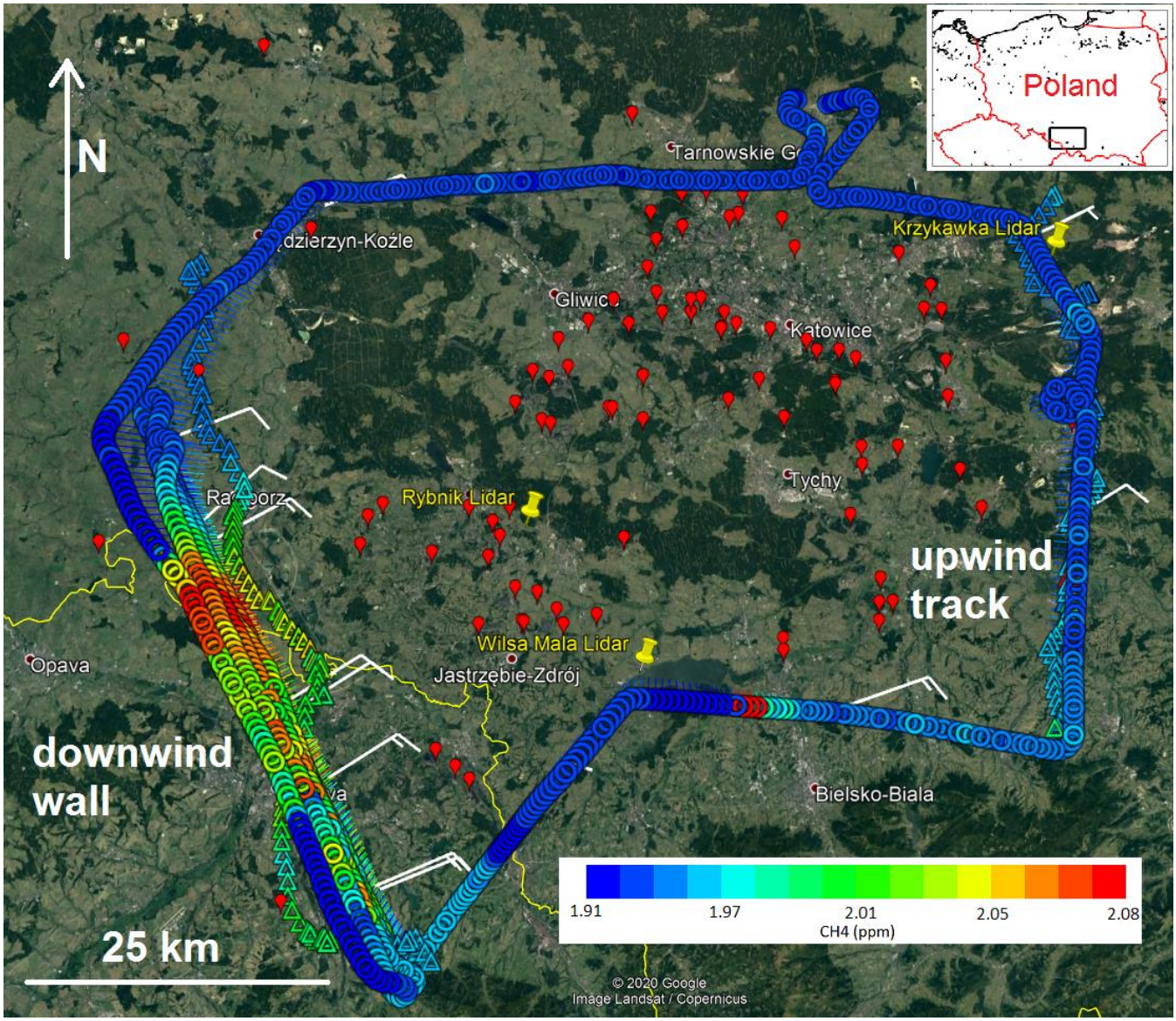


Figure 1: Flight track for flight B, color-coded with in situ measured CH_4 mole fractions. The wind was blowing from the north-east over the USCB (as indicated by the white wind barbs) carrying emissions to the south-west. Airborne observations averaged over 20 s are displayed as circles and mobile ground observations averaged over 80 s below the upwind track and the downwind wall are marked as triangles. Red markers show the locations of active coal mine shafts from the CoMet v2 inventory.

2.2 Mass balance method

We use a mass balance method to calculate emission estimates for the USCB from two flights conducted on June 6. This approach is subject to several assumptions. First, the wind speed, wind direction, emissions, and the PBLH should remain

constant over the sampling time. Second, the trace gas plume has to be discernible from the atmospheric background. Third, there shouldn't be any entrainment/detrainment into the free troposphere and the lifetime of the species must be much longer than transport and sampling times. Finally, the trace gas plume should be well-mixed between the lowest flight track and the ground. These criteria are most likely to be met in the early afternoon, when the PBL has reached its maximum height and does not rise any further. The PBLH generally increases during the morning; hence afternoon flights are preferred over morning flights for mass balance studies. For our morning flight, we determine the temporal change of the PBLH during sampling to be 20% of its final height. We apply a correction to the observed trace gas enhancements to account for this change (see Sect. 3.2).

In our approach we calculate the mass flux of each trace gas (CO_2, CH_4, and CO) through a vertical surface along the downwind flight tracks, here called "wall" (see Figure 1). The wall stretches from the ground to the top of the PBL. Since the downwind measurements, ground-based and airborne, were not taken exactly on this wall, as a first step, all data used in the calculation, are projected onto the closest point of the wall and then interpolated to fill the entire wall using the well-known kriging approach. The flux through the wall is defined by

$$F = \int_{z=\text{ground}}^{z=PBLH} \int_{x=S}^{x=N} \Delta c_{x,z}\, v_{x,z}\, dx\, dz, \quad (1)$$

where $\Delta c_{x,z}$ is the concentration enhancement of the trace gas above the background at each grid point, while $v_{x,z}$ describes the wind speed component at each grid point perpendicular to the wall. The integration area is defined by the ground, the PBLH, and the edges of the wall to the south *S* and north *N* (see bottom right panel of Figure 2). The PBLH is determined from the vertical gradient of potential temperature, measured during profile flight sections, and the times when the top of the PBL was crossed in the wall. During the afternoon flight the PBL top was crossed three times in the wall and from this information the slanted boundary layer height could be well constrained.

The concentration enhancements Δc are calculated from observed, interpolated mole fractions m and the background mole fraction m_0 of the trace gases using linear temperature and pressure profiles deduced from the airborne measurements:

$$\Delta c = (m - m_0)\, M\, \frac{p}{R\,T}. \quad (2)$$

Here, *M* is the gas molecular weight, *p* the pressure, *R* the universal gas constant, and *T* the temperature in Kelvin.

To retrieve trace gas mole fractions m and wind speed v on the wall between the actual flight tracks, we use the kriging interpolation method with a stochastic Gaussian model. Kriging creates a grid of estimated values from data points with sparse spatial coverage and also gives standard errors for these values. We use a modified version of the EasyKrig software (© Dezhang Chu and Woods Hole Ocean Institution). For more details see Mays et al. (2009) and Pitt et al. (2019), who previously used this software in an aircraft mass balance study.

For CH_4, not only the mole fraction measured along the flight transects but also the data of the ground-based measurements is included in the kriging. Although CO_2 was also measured on the ground by the same instruments, the data cannot be used because it is heavily influenced by the surrounding car traffic. For ground-based CO_2, neither large scale enhancements nor background concentrations could be discerned. We chose the CH_4 observations along the ground track closest in time to the

airborne measurements. The data is projected onto the downwind wall, averaged over 20 seconds and then interpolated horizontally to regular distances before kriging. Airborne data is averaged over 10 second intervals in order to reach similar spatial resolution to the ground-based data. Only data below the PBLH is included in the kriging process. We then closely followed the approach described in Pitt et al. (2019). The kriging output fields of CH_4, CO_2, CO mole fractions and perpendicular wind speed are given at a grid resolution of 0.1° in latitudinal direction and 20 m in the vertical.

2.2.1 Downwind and upwind background determination methods

For the mass balance approach, the background mole fraction m_0 of the trace gases needs to be determined. Here we compare two methods: (i) background estimated from the downwind wall's edges and (ii) background estimated from the upwind leg. The downwind background method assumes that the boundary layer height remains constant for the time of sampling within the wall, while the upwind method requires the boundary layer to stay at the same height for the whole flight time and ideally a quasi-Lagrangian sampling of the same air mass in the upwind and downwind transects. Thus, the less strict criteria of the downwind background method are more likely to be met in real conditions and we will use this method in our best estimate and the upwind background as a sensitivity test. The downwind method also requires that there are no sources upwind of the area of interest which would create a complex concentration pattern flowing into the domain. To show this we used our upwind flight transect similar to previous studies (Karion et al., 2013; Heimburger et al., 2017).

In order to determine the downwind background mole fraction from the wall's edges, we evaluate the variability of the CH_4 observations within the PBL. The background is separated from the plume using the standard deviation within a 2 min interval for airborne and 10 min interval for ground-based data. Starting at the edges of the wall, the interval is moved towards the center. We define the boundary between CH_4 atmospheric background and plume where the standard deviation surpasses 3.4 ppb CH_4. The average CH_4 background standard deviation is 2.9 ppb. The CO_2 background section is adopted from the CH_4 background, because the variability in the background is too high for this approach to be applicable. The CO background threshold for the 2 min interval is 4.5 ppb with an average background standard deviation of 3.5 ppb. We average all background mole fraction observations within the PBL to the south and north of the plume separately. The mean of these two values is considered as the average background for the downwind method. Thus, we assume a linear spatial gradient in the trace gas background.

The second way of determining the atmospheric background mole fraction uses the observations within the boundary layer from the upwind flight transect, which was flown about 15 minutes before the downwind wall and is here used in a sensitivity study. Methodologically, we define a perpendicular inflow transect according to the prevalent wind direction, and project the upwind measurements onto this line (Supplement Figure S1). After interpolation to regular distances, the average inflow mole fraction represents the upwind trace gas background. This approach has the advantage that sources upwind of the area of interest can be identified through potential enhancements in the upwind transect and are excluded from the emission estimate. On the other hand, the upwind background assumes that the same air masses are sampled in the up- and

downwind, which is not true for our two flights, since the air masses needed approximately 3-4 hours to travel form the upwind to the downwind measurement location, while the aircraft only needed 15 minutes. The maximum time separation between up- and downwind sampling is 1.5 h. Thus, our sampling is not strictly Lagrangian (i.e. air mass following) and changes in boundary layer background concentrations over time may affect the emission estimates using the upwind background method. Another disadvantage of using upwind background concentrations with respect to CO_2 is the necessity to account for large scale ground fluxes like the biogenic uptake of CO_2, which is discussed in the next section.

2.2.2 Simulation of biogenic uptake of CO_2

We derive the influence of biogenic uptake of CO_2 from a combination of backward trajectories, calculated using the Stochastic Time-Inverted Lagrangian Transport (STILT, Lin et al., 2003) model, and biospheric fluxes from Vegetation Photosynthesis and Respiration Model (VPRM; Mahadevan et al., 2008). STILT was set up with receptors distributed along the flight track of the downwind wall and from each receptor, we then release 100 particles in the model. To drive the trajectory simulations, we used output of ECMWF HRES short-term forecasting system (approx. 9 km x 9 km spatial resolution, 137 vertical levels), preprocessed to assure mass-conservation of the wind fields. The median locations of the particle ensemble then constitute the median trajectories (Figure S2). ~~As we are only interested in the influence from our domain of interest, we truncate the median trajectories at their position closest to the upwind measurement.~~ The optimal use of the model in the method described would require for the upwind track to be flown in exactly Lagrangian manner, sampling the same air mass upwind and downwind of the sources. In our case, we have a single hour of temporal difference in the observations, and a four-hour difference in the air-mass flow between measurement locations, during which the biosphere was able to uptake CO_2. For the difference in background mole fractions, the hour of biogenic uptake between upwind and downwind observations is relevant.~~The time lag between upwind and downwind sampling is approximately one hour, thus, t~~ The biospheric VPRM contribution to the downwind measurements is ~~finally~~ calculated using the footprint derived from the last hour of each trajectory, multiplied with the VPRM fluxes corresponding in time and location. ~~-~~We decided on this hybrid approach, in which we assume that we can still link the measurements to our model quasi-directly, despite the fact that the model results are simulated for a location several tens of kilometers away from the actual upwind measurement location. It should be noted that it is assumed here, that the biospheric fluxes are spatially homogeneous. We add this contribution to the downwind CO_2 observation, only when using an upwind background, and then use these values for the interpolation with kriging.

2.3 Error estimate

For an error estimate of the derived mass flux we consider the statistical error of the input data and the systematic error of the method.

2.3.1 Statistical error

The statistical error of our approach is determined using error propagation in the flux equation (Equations 1-2). The uncertainty calculation of the concentration enhancement $u_{\Delta c}$, the flux density uncertainty u_{Fd} and the final flux uncertainty u_F are described by equations 3-5:

$$\Delta c = c - c_0 \;\rightarrow\; u_{\Delta c} = \sqrt{{u_c}^2 + {u_{c_0}}^2}\;; \tag{3}$$

$$Fd = \Delta c * v \;\rightarrow\; u_{Fd} = \sqrt{\left(\frac{u_{\Delta c}}{\Delta c}\right)^2 + \left(\frac{u_v}{v}\right)^2} * Fd\;; \tag{4}$$

$$F = \sum_i Fd_i * A \;\rightarrow\; u_F = \sqrt{\sum_i \left(u_{Fd_i}\right)^2} * A\;. \tag{5}$$

The first two uncertainties are calculated for each grid point of the wall surface; the final flux uncertainty u_F is the combination of the single uncertainties. The trace gas uncertainty u_c and wind speed uncertainty u_v are a combination of measurement and kriging uncertainties expressed as kriging standard error (KSE):

$$u_{c/v} = u_{\text{measurement}} + \text{KSE} = u_{\text{measurement}} + \sqrt{u_{\text{kriging}} \cdot var(\Delta c)} \tag{6}$$

The measurement uncertainty $u_{\text{measurement}}$ has been determined to 1.1 nmol mol^{-1} (hereafter referred to as ppb) for CH_4, 0.15 µmol mol^{-1} (hereafter referred to as ppm) for CO_2 (Table S2, Text S1), and 7 ppb for CO (Kostinek et al., 2019). The wind speed measurement uncertainty u_v has been assessed to be 0.3 m/s for each of the horizontal components (Mallaun et al., 2015). The uncertainty of the interpolation and extrapolation kriging method is output by EasyKrig as a gridded field of normalized variance values u_{kriging}. To retrieve the gridded KSE (see Figure S4), which is the equivalent to the standard deviation, we multiply the kriging error output u_{kriging} by the variance of the kriging input dataset Δc and then take the square root (Equation 6). The background mole fraction uncertainty u_{c_0} is here defined as the standard deviation of all data points contributing to the background calculation (see Table 4). The uncertainty of the grid cell area *A* is assumed to be zero.

2.3.2 Systematic error

We conducted several sensitivity tests in order to test the robustness of our mass balance method and to determine its systematic error. These sensitivity tests are described and discussed in Sect. 3.4. We assume all systematic errors to be independent and calculate the total absolute systematic error as the square root of the sum of squared individual differences from the best estimate, which treats the data as described in Sect. 2.2 with a downwind trace gas background.

2.4 Bottom-up emission inventories

Several inventories of greenhouse gas and air pollutant emissions exist for the USCB. They vary in spatial and temporal resolution, as well as in the time for which they are available. Table 1 gives an overview of the six inventories we use in this study for comparison with top-down derived CH_4, CO_2, and CO emissions in the USCB region.

Table 1: Overview of emission inventories used in this study. The year states the last year, for which data are available.

Inventory	Year	Resolution	Coverage	Gases
E-PRTR v16 (EEA, 2019)	2017	point	Europe	CH_4, CO_2, CO
CoMet v2 (internal inventory)	2016	point	Silesia, CZ Moravia	CH_4, CO_2
Scarpelli CH4 (Scarpelli et al., 2020)	2016	0.1° x 0.1°	Global	CH_4 (Oil, Gas, Coal)
CAMS-REG v3.1 (Granier et al., 2019)	2016	0.1° x 0.05°	Europe	CH_4, CO_2, CO
EDGAR v5/v4.3.2 (Crippa et al., 2018; Janssens-Maenhout et al., 2019)	see right	0.1° x 0.1°	Global	CH_4 (2015), CO_2 (2018), CO (2012)
GESAPU (Bun et al., 2019)	2010	15`` x 15`` (~400 m)	Poland, Ukraine	CH_4, CO_2, CO

The first point source inventory listed in Table 1~~Table 1~~ is the European Emission Release and Transfer Register (E-PRTR). It results from the Regulation (EC) No 166/2006 which implements the United Nations Economic Commission for Europe (UNECE) PRTR Protocol under which industrial facilities have to report their emissions to air if they exceed a threshold of 100 t/a for CH_4, 100 kt/a for CO_2, and 500 t/a for CO. Annual data can be downloaded from the European Environmental Agency's website (EEA, 2019). More information on the E-PRTR is given via its website: https://prtr.eea.europa.eu/ (last accessed: 24 February 2020).

The CoMet v2 inventory is a point source inventory based on the E-PRTR 2016 emissions created by the CoMet team especially for this campaign. It comprises anthropogenic sources of CH_4 and CO_2 in the USCB and its vicinity. The largest difference between the E-PRTR and the CoMet inventory is that E-PRTR considers each coal mine as one single point source, often located at the mining operator headquarters, whereas in the CoMet inventory individual ventilation shafts were visually geo-localized using Google Earth. Then, the emission value of each mine was evenly distributed between all ventilation shafts belonging to that mine. Active Czech coal mines in the Ostrava region did not report any CH_4 emissions to

E-PRTR but were assumed to emit the same amount of CH_4 per ton of extracted coal as Polish mines. We deduced a factor of 11.8 ± 5.2 kg CH_4 per ton of extracted coal for the USCB mines listed in Table S3 and applied this value to the Czech mines of Karvina, Karkov, CSM, and Paskov. The locations of the fourteen listed landfills and waste disposal sites were checked against satellite imagery. Their CH_4 emission is assumed to be 3.3 kt/a, which is less than 1% of the total USCB emissions.

Scarpelli et al. (2020) published the newest gridded emission inventory available for comparison within in this study. It only contains CH_4 emissions from oil, natural gas and coal exploitation. But since these are the main sources (87% according to CAMS) of CH_4 emissions in the USCB, values are comparable to the total of other inventories. Scarpelli et al. (2020) use the national totals of emissions reported to the UNFCCC and distribute them according to the positions of relevant infrastructure. Uncertainties of the emissions are based on the emission factor uncertainties from the Intergovernmental Panel on Climate Change (IPCC) and are given as gridded information. Averaged over the USCB, the given relative error standard deviation for CH_4 emissions is 60.9%.

The Copernicus Atmospheric Monitoring System (CAMS) regional emission inventory (CAMS-REG-GHG/AP; Granier et al., 2019) is based on the TNO-MACC inventories (Kuenen et al., 2014). This inventory offers a resolution twice as high as the Scarpelli and EDGAR inventories. The inventory was also constructed by using the reported emission national totals by sector and spatially distributing them consistently across all countries by using proxy parameters.

The most widely used gridded emission inventory is probably the Emission Database for Global Atmospheric Research (EDGAR, https://data.europa.eu/doi/10.2904/JRC_DATASET_EDGAR) global emission inventory. The most recent version 5.0 (https://edgar.jrc.ec.europa.eu/overview.php?v=50_GHG) includes emissions of the three major greenhouse gases CO_2, CH_4 and N_2O. It is based on the previous EDGAR version 4.3.2 (Janssens-Maenhout et al., 2019). We use the CO emissions from the air pollutant inventory (Crippa et al., 2018) from version 4.3.2. The most recent year of emission data is 2015 for CH_4, 2018 for CO_2, and 2012 for CO. In EDGAR, annual country-specific emissions are derived from international activity data and emission factors, which are then distributed in time and space using monthly shares and spatial proxy datasets. The data includes uncertainty factors per species for three types of countries: OECD countries of 1990, countries with economies in transition in 1990, and the remaining countries in development. European emissions from EDGAR in 2012 have standard deviations of 16 % for CH_4, 2.5 % for CO_2 (Janssens-Maenhout et al., 2019), and 65 % for CO (Crippa et al., 2018).

The GESAPU inventory (Bun et al., 2019) has been created for Ukraine and Poland only for the reference year 2010. Originally, it is a point, line, and area source inventory based on shapefiles. The advantage of this type of information is that it has a very high resolution, but can also be gridded with any spatial resolution and orientation. The GESAPU inventory comprises all sectors of anthropogenic emissions. Here we use a gridded version of the emissions with a resolution of 15 arc seconds (approximately 296 m x 463 m for the region).

Figure 2 shows the spatial distribution of CH_4 emissions as given by the six inventories. Point sources from E-PRTR and CoMet inventory are displayed as black markers while the background colors give the gridded inventory values. Although the inventories generally agree on the locations of CH_4 emissions, there are several cases, where sources seem to be missing.

Regarding point sources, E-PRTR (top left) has fewer individual sources than the CoMet inventory (top right) due to the separation in single ventilation shafts. Additional mines in the CoMet inventory include the four Czech mines and the four ventilation shafts of the Brzeszcze mine around 19.15°E and 49.95°N. The gridded Scarpelli (top left) emission distribution for CH_4 does not represent the point sources well. There are no emissions north of 50.2°N although several mines are located in this northern area. Generally the CAMS (top right) emission maxima seem to represent the point source locations better than the Scarpelli or EDGAR (bottom left) emission distribution, with the exception of the Czech mines, which are included in Scarpelli and EDGAR, but not in CAMS. In the GESAPU inventory, the high CH_4 emissions associated with mining activities were visualized by overlaying marker for sources above 1 kt/a on the gridded emission map. These are fewer high emitting sources than in the E-PRTR inventory. This could be caused by consolidation and separation of mines between 2010 and 2017, the respective years for the data. Two flights (on June 6, 2018), which are shown as blue tracks in Figure 2, were designed to capture the emissions of the region during north-easterly wind conditions.

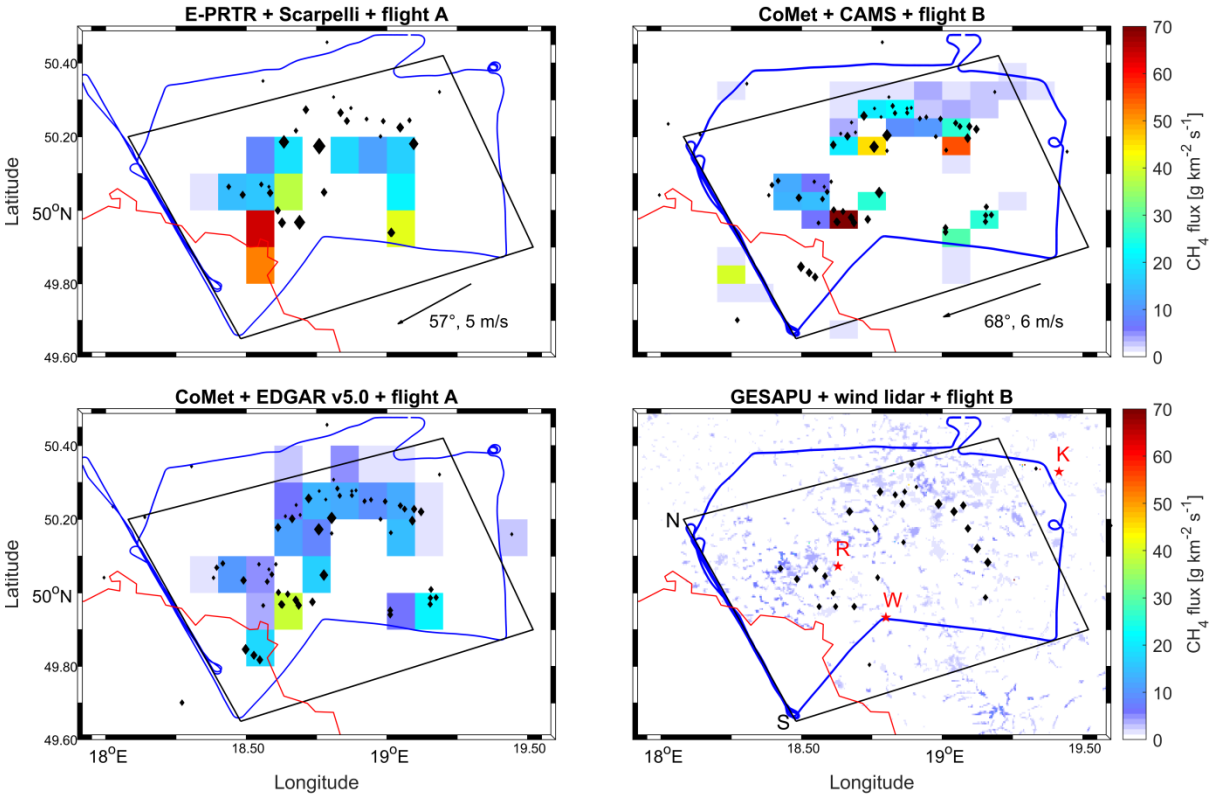


Figure 2: CH_4 emission distribution of inventories in the USCB. Background colors give emissions from gridded inventories Scarpelli, CAMS, EDGAR, and GESAPU, while the markers are sized according to the emissions of the point source inventories E-PRTR and CoMet. Additionally, we added GESAPU sources above 1 kt/a CH_4 as markers for better visibility. The black boxes denote the emission area for comparison with the mass balance estimate via aircraft. The blue lines show the flight tracks of the flights A and B on June 6, 2018, used in the mass balance and the arrows in the top two panels show the mean wind direction during the two flights. The red line denotes the Polish-Czech border. Red stars in the bottom right panel show the locations of the wind lidar instruments (R: Rybnik, W: Wilsa Mala, K: Krzykawka). Also marked in this panel are the southern and northern edges of the downwind wall S and N.

The CO_2 and CO emission distribution in the inventories is displayed in Figure 3. CO_2 point sources (from E-PRTR and CoMet) agree well with EDGAR and CAMS, except for the strong CO_2 and CO emissions associated with the Lagisza power plant and Acelor Mittal steel factory at 50.34°N and 19.28°E, which are correctly placed in the northeast corner of the flight track in E-PRTR, CoMet and GESAPU. Instead, EDGAR and CAMS include an emission hot spot to the southeast and east, respectively, of this location, that is not associated with a point source. The Rybnik power plant, located in the central western USCB, is the strongest point source emitter of CO_2 in all inventories. CO has one emission hot spot in the USCB, namely the Acelor Mittal steel factory next to the Lagisza power plant with 137 kt/a in E-PRTR 2017. This source is not represented in EDGAR and shifted to the east in CAMS. GESAPU includes this source, but with much lower emissions of 63 kt/a.

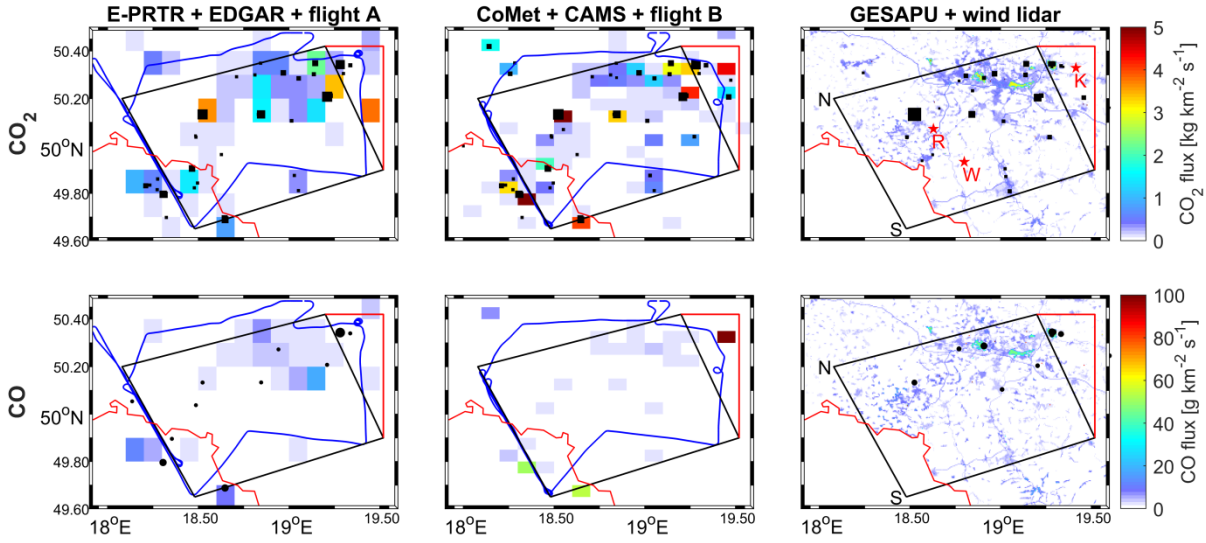


Figure 3: Like Figure 2 but for CO_2 and CO. GESAPU sources above 0.1 Mt/a and 1 kt/a for CO_2 and CO, respectively, are added as markers. The straight red lines show the addition to the mass balance area necessary because of misplaced sources.

To compare the emission inventories with our mass balance flights, the emissions of each inventory are summed up within an area representative of the flight track and wind direction (more details see Sect. 3.3), which is marked by the black boxes in Figure 2 and 3. Since some of the CO_2 and CO sources are obviously misplaced in the gridded inventories, but really lie within our mass balance area, we enlarged the mass balance area toward the east in order to include these sources into the USCB sum. These enlargements are marked by red lines in Figure 3. Although missing sources influence the comparison between inventories and the emission estimate via aircraft, the misplacements might not, since misplaced emissions are now within the enlarged mass balance area.

For each inventory, the total annual emission from the enlarged area including the reported uncertainty is given in Table 2. These values include emissions from all sectors available in the inventories (see also discussion in Sect. 4). Scarpelli assumes the highest emissions for CH_4, followed by CAMS and CoMet. GESAPU features the lowest CH_4 emissions, which

might partly arise from the sources in the Czech Republic, which are not covered in the inventory. Highest CO_2 emissions are assumed by the EDGAR inventory. CO emissions are highest in CAMS, closely followed by GESAPU.

Table 2: Annual emission totals in the USCB area for different emission inventories and trace gases.

Inventory	CH_4 [kt/a]	CO_2 [Mt/a]	CO [kt/a]
E-PRTR	448	37.0	144
CoMet	581	39.1	--
Scarpelli	685 ± 456	--	--
CAMS	621	51.5	329
EDGAR	556 ± 89	59.0 ± 1.5	236 ± 154
GESAPU	405	56.8	291

3 Results

3.1 Meteorological situation

The meteorological conditions have to fulfil certain criteria for a feasible mass balance calculation. On June 6, 2018, the weather conditions for an airborne mass balance experiment in the USCB were advantageous due to relatively constant wind speed and wind direction over the sampling time. The PBLH changed considerably during flight A in the morning, but was rather constant during flight B in the afternoon.

The wind lidar measurements at Rybnik airport were located close to the center of our in situ wall (Figure 2) and can be used to assess the wind history over the entire measurement day. Vertical profiles of wind speed and wind direction show that during the previous night a low-level jet blew over the area with wind speeds of more than 10 m/s, in the morning the wind slowed down to around 5 m/s and then accelerated to 6-7 m/s around 13:00 UTC (Figure 4, Table 3). The boundary layer wind direction was between 50° and 70° over the entire day. The nightly low-level jet prevented accumulation of emissions, and the slowing down around 6:00 UTC provided relatively constant wind speeds for four hours before we started our downwind sampling at 10:00 UTC. This steady wind history prior to the flight is crucial for the mass balance approach, because of the assumptions stated in Section 2.2. During this time emissions from the farthest shafts (75 km from downwind wall) were able to travel from emission to observation location at constant wind speed and direction. A comparison of aircraft observations in the downwind wall and wind lidar averages during the observation times is given in Table 3. Observed wind speeds with the lidar are within the range of aircraft observed wind speeds. Generally wind speeds in the southern USCB were about 1 m/s higher than in the northern part of the USCB.

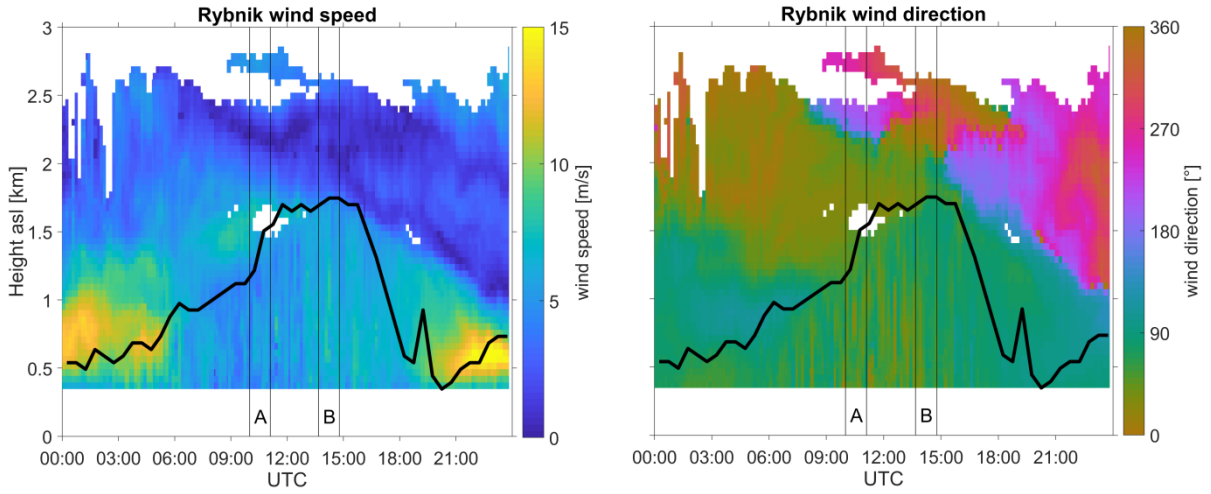


Figure 4: Wind speed and direction at Rybnik measured with a Doppler wind lidar on June 6, 2018. The bold line denotes the PBLH determined from the eddy dissipation rate and the thin vertical lines illustrate the downwind wall sampling times of flights A and B.

Table 3: Overview of wind data and PBLH from aircraft averaged within the downwind wall and wind lidar observations at Rybnik. Aircraft data give uncertainty ranges due to measurement uncertainty and wind lidar data state a standard deviation of the measurements within the PBL. The wind speed obtained from the lidar is additionally as average over the 4 hours previous to the downwind sampling.

	Mean wind speed perpendicular [m/s]		Wind dir. [°]		PBLH [km asl]	
	Aircraft	Wind lidar	Aircraft	Wind lidar	Aircraft	Wind lidar
Flight A (morning)	4.8 ± 0.3 to 5.7 ± 0.3	5.0 ± 0.9 and 5.1 ± 0.9 (4h)	48 ± 2	57 ± 15	0.9 ± 0.05 and 1.25 ± 0.05	1.2 ± 0.05 to 1.5 ± 0.05
Flight B (afternoon)	5.8 ± 0.3 to 7.0 ± 0.3	6.4 ± 0.8 and 5.7 ± 1.0 (4h)	62 ± 2	68 ± 12	1.3 ± 0.05 and 1.8 ± 0.05	1.7 ± 0.05

The diurnal development of the PBLH, with a maximum of 1.7 km above sea level (asl), is discernible from the wind lidar observations. The PBLH measured by the wind lidar increased from 1.1 to 1.5 km during the sampling of flight A, but remained relatively constant at 1.7 km during flight B. We also determined the PBLH from two vertical aircraft profiles of potential temperature, observed before and after the sampling of the downwind wall (Figure S3). Before flying the wall pattern, we obtained a vertical profile in the southern part of the USCB area (around 18.2°E, 49.8°N). After finishing the wall pattern a northern profile was sampled on the way back to Katowice airport (18.2°E, 50.3°N). During both flights, the

PBLH was about 400 m lower in the southern part than in the northern part of the USCB. Thus, the PBLH data in Table 3 describes a latitudinal gradient for the aircraft, and temporal changes from the wind lidar.

Furthermore, for the mass balance, we assumed no entrainment from the free troposphere during sampling time. This assumption is supported by a strong capping inversion at the PBLH observed in the aircraft profiles (Figure S3). Still, since the PBLH was increasing during the sampling for flight A, there was considerable entrainment of free tropospheric air into the mixed layer. The correction we applied for this temporal change of the PBLH is described in the following section. The uncertainty related to this correction is assessed in the sensitivity test (Sect. 3.4) concerning the temporal PBLH variability.

3.2 Kriging results

For our mass balance, we use airborne in situ observations from two flights on June 6, 2018. CH_4, CO_2, and CO enhancements were clearly observed in the downwind wall. The ground-based teams drove below the upwind and downwind legs using the closest highways and national roads. Halfway through the southern track we ascended and descended to derive the height of the PBL based on meteorological measurements. Above the PBL, observed CH_4 and CO concentrations were lower than within the PBL, while CO_2 concentrations were higher.

In a first step of emission estimation for the entire USCB (as described in Sect. 2.2) the observed data in the downwind wall is inter- and extrapolated using the kriging algorithm (Figure 5). Details of the kriging parameters can be found in the supplement (Text S2). Mole fractions in the wall are cut off below the ground, above the PBL, and to the south and north of the flight legs (points S and N).

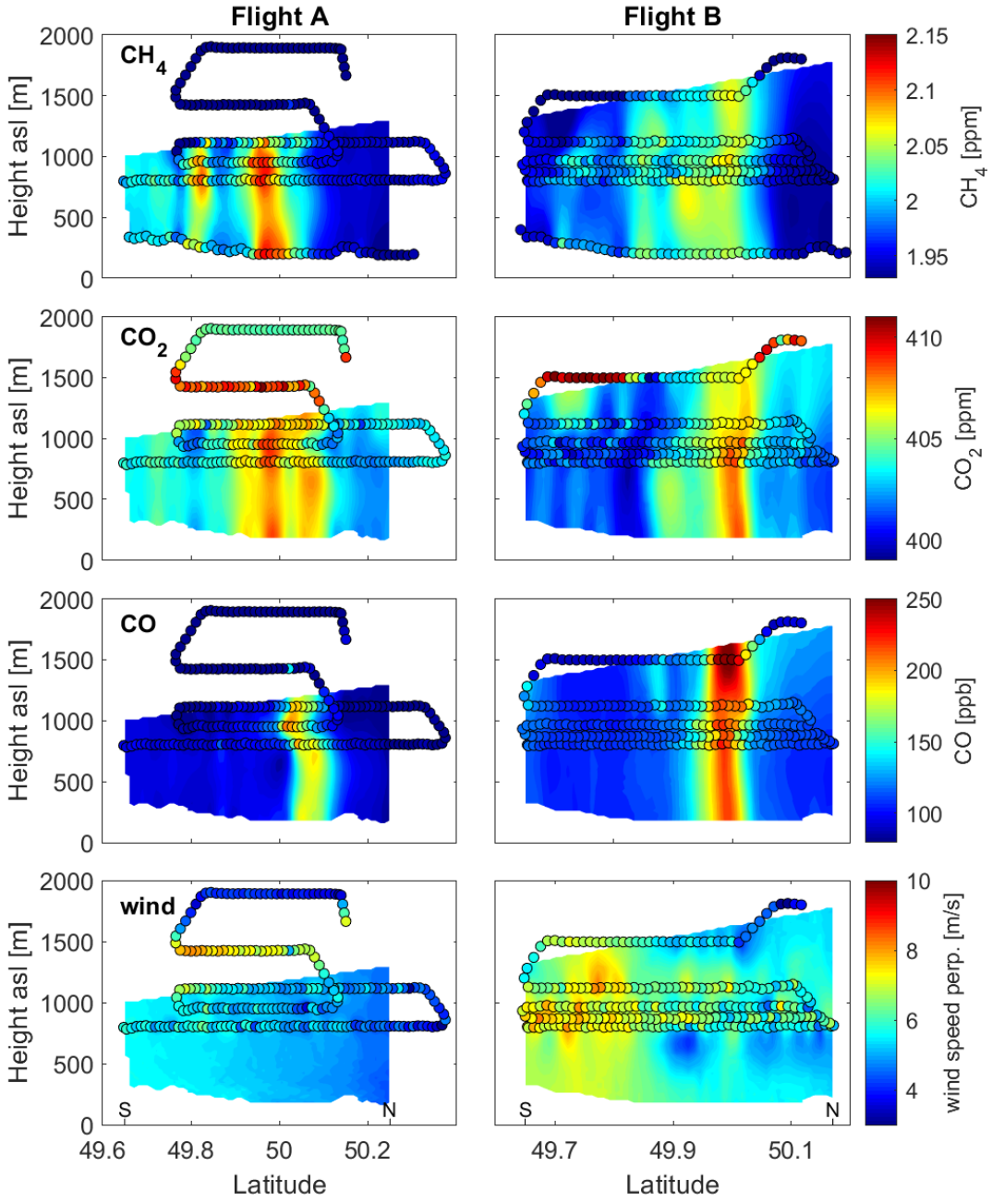
For the morning flight A, the trace gas plumes reach from the ground to the top of the PBL. The transects on the ground and at 800 m show the highest CH_4 maxima (Figure S4). At 1000 m and 1100 m the maximum enhancements are lower. The same is true for the CO_2 and CO enhancements. This is probably caused by the growing PBLH during the flight. During the downwind measurement of the morning flight A, the height of the PBL increased from 1.2 km asl (0.9 km above ground level (agl)) to 1.5 km asl (1.2 km agl), which is an increase of 20%. The lowest transect (800 m) was sampled first in the shallowest PBL. The two upper transects were sampled about half an hour later, when the PBLH had increased by about 20%. Thus the emissions from the USCB were mixed within a much smaller volume during the lowest transect, than during the following two. The ground-based sampling of the morning flight took place between 9:00 and 10:40 UTC. Two cars started in the center of the downward projected flight track and moved away from each other to the south and north. Thus, the central part was sampled first, during low PBLH conditions. To account for the low PBLH during the first flight transect and the ground-based sampling, we apply a correction factor of -20% to the ground observations and the lowest flight transect. Figure S4 shows the original, uncorrected observational data, while Figure 5 shows the corrected values. Corrected enhancements are in the order of 0.16 ppm CH_4, 7 ppm CO_2 and 130 ppb CO.

During the afternoon flight B, the CH_4 plume is evenly distributed between the ground observations and the lowest flight track at 800 m (Figure 6). Thus, we assume good vertical mixing within the PBL and use the same CO_2 and CO mole

fractions at the ground as in the lowest flight transect. Trace gas enhancements are in the order of 0.12 ppm CH_4, 6 ppm CO_2 and 120 ppb CO, thus, lower than during the morning flight. The main CH_4 plume is located at 50.0°N with a secondary plume around 49.8°N. There are two CO_2 plumes at 50.0°N and 50.1°N. The CO plume is located at 50.0°N.

The horizontal wind speed shows a latitudinal gradient with higher wind speeds in the south than in the north for both flights. This gradient is preserved when using a kriged wind field for flux calculation instead of an average wind speed for the whole downwind wall (as discussed in Sect. 3.4).

Error estimates from the interpolation and extrapolation are retrieved from the kriging software as gridded fields (see Figure S5). The KSE generally increases with distance to the measurement locations and is highest at the ground for CO_2, CO and wind speed because no ground-based measurements were available for these parameters.



460 **Figure 5: Mole fractions and perpendicular wind speed in the downwind in situ wall from observations (circles) and inter- and extrapolation with a kriging algorithm (shading). The CH_4 wall incorporates ground-based measurements. For CO_2 and CO the ground mole fraction is assumed to be the same as in the lowest flight track. The wind extrapolation does not use any information below the lowest flight track.**

3.3 Background mole fractions

We applied both the downwind and the upwind method (see Sect. 2.2.1) to determine atmospheric background mole fractions of trace gases. Average background mole fractions and standard deviations for both methods are summarized in Table 4. Figure 6 shows the observed PBL mole fractions of CH_4, CO_2, and CO at different heights for flight B. The highest transect (light blue), originally planned in the free troposphere above the PBL, turned out to partially be within the PBL, but the southern and northern end were sampled in the free troposphere. The background mole fractions according to the downwind method are displayed as dotted lines. For flight A, the background could not be reached to the south of the downwind wall and only background values from the north were used for CH_4 and CO_2 (Figure S4).

Table 4: Average background mole fractions and their standard deviations calculated with the downwind and upwind methods.

	Downwind background			**Upwind background**		
	CH_4 [ppm]	CO_2 [ppm]	CO [ppb]	CH_4 [ppm]	CO_2 [ppm]	CO [ppb]
Flight A	1.941 ± 0.005	402.7 ± 0.9	82.5 ± 8.9	1.944 ± 0.006	404.6 ± 1.0	81.6 ± 8.5
Flight B	1.944 ± 0.007	401.8 ± 0.7	110.5 ± 5.2	1.936 ± 0.004	402.8 ± 1.8	-

The upwind mole fractions (black lines) were shifted to the corresponding latitudes of the downwind wall based on the wind direction. The CH_4 upwind mole fractions follow the same north-south gradient as the downwind background (Figure 6, top). Around 49.94°N the CH_4 mole fraction is slightly enhanced in the upwind. There is a similar enhancement around 50.13°N in flight A (Figure S4). Due to the projection, these would be between 50.2°N and 50.3°N on the inflow track. The only source upwind of the inflow track in the inventories is the Trzebinia mine and power plant at 19.44°E and 50.16°N. We use the ground-based observations below the upwind track (grey line) to confirm our aircraft observations. They show similar absolute values and a similar north-south trend to the airborne track. Additionally, there are three spikes between 49.73° and 49.78°N. These locations correspond to an inflow latitude of around 50.0°N and probably originate from sources close by, since they do not appear in the airborne observations. The largest peak most likely originates from the coal processing and waste water treatment facilities right upwind to the measurement route at 50.027°N and 19.438°E.

The CO_2 upwind background is higher than downwind mole fractions at both ends of the measurement transects but lower in the center, where the downwind plume was observed. The average upwind background of CO_2 is 2 ppm and 1 ppm higher than the downwind background for flight A and B, respectively. This discrepancy is caused by the biogenic uptake of CO_2 between the upwind and downwind transects. The impact of the biogenic sink is discussed below.

Upwind CO observations during flight B do not cover the complete transect due to a start delay of the QCLS. Thus, we did not use the CO upwind background for this flight. The CO upwind observations for flight A show small variations resulting

in a background standard deviation of about 9 ppb. Here, the upwind CO measurements are smaller than downwind background values.

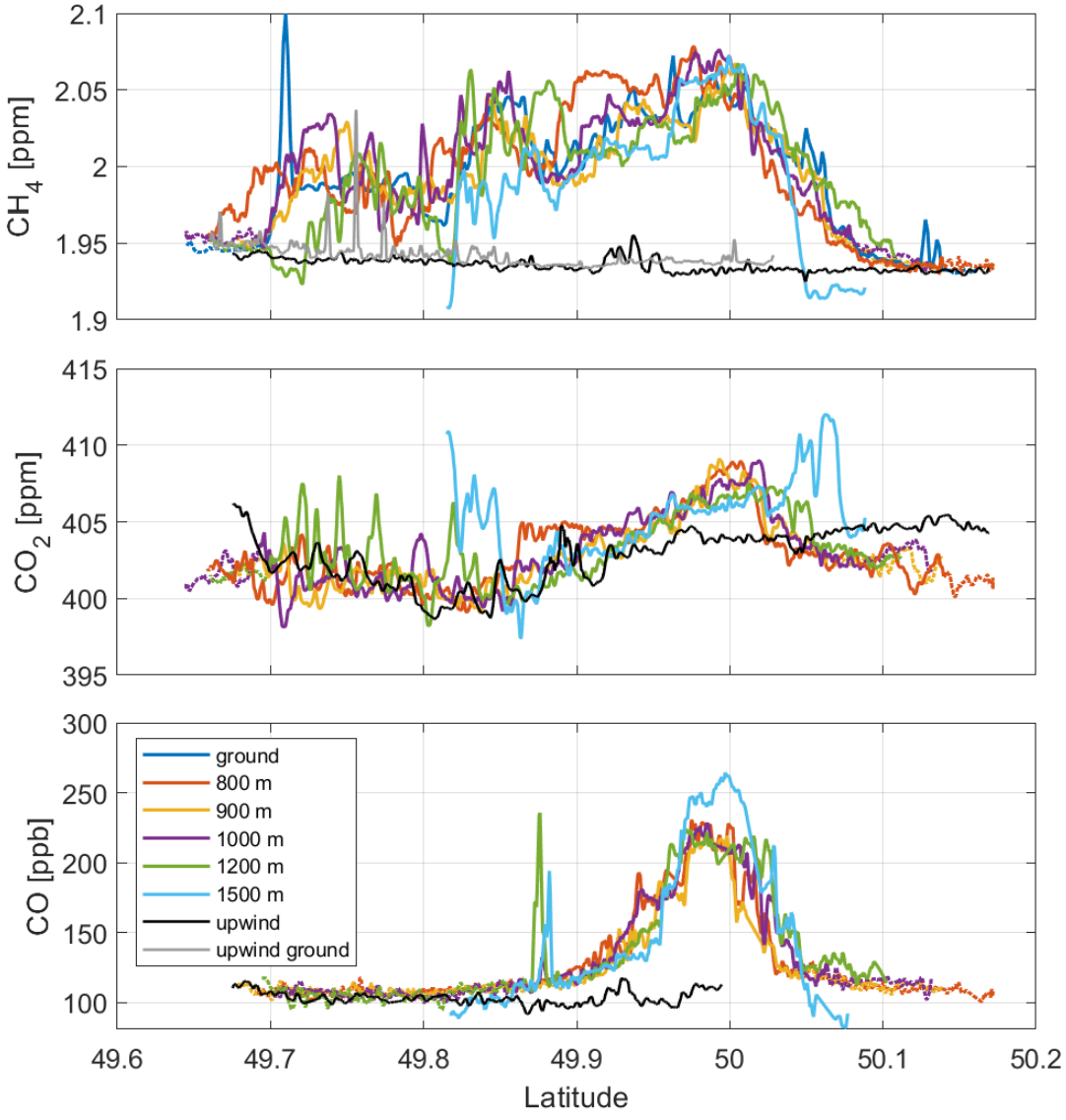


Figure 6: Mole fractions of CH_4, CO_2, and CO at different heights above mean sea level within the PBL downwind of the sources for flight B. Background mole fractions according to the downwind method are displayed as dashed part of the lines at the edges. Additionally, the background according to the upwind method is shown in black and grey. Upwind data has been shifted to the respective downwind latitude. The CO upwind background stops at 50° N due to an instrument start up delay on this part of the track.

The upwind background method calls for an estimate of the biogenic uptake of CO_2. We estimate this uptake from the STILT trajectories and the VPRM model (see Sect. 2.3.2). Figure S2 exemplary shows the truncated trajectories for the

800 m altitude transect of flight B. Trajectories for other transects and flight A are very similar. The biogenic uptake for each trajectory is determined from the last hour of transport. By subtracting the VPRM uptake from the corresponding downwind measurement (as the uptake is negative), one can obtain a downwind CO_2 concentration without biospheric influence. This uptake is on average 1.00 ppm for flight A and 0.95 ppm for flight B.

3.4 USCB emission estimate

From the two mass balance flights on June 6, 2018, we determined the total USCB emissions of CH_4, CO_2, and CO. Figure 7 summarizes the best-estimate emissions and the sensitivity calculations (see Sect. 2.3). The uncertainty of the best-estimate includes the statistical error, calculated from the uncertainties of the input parameters and the systematic error calculated from the sensitivity tests. The CH_4 emission estimates for the entire USCB on June 6 are 13.8 ± ~~3~~4.~~6~~ 3 kg/s and 15.1 ± ~~3~~4.0 kg/s for flights A and B, respectively. This is a difference of 9% between the two flights. The CO_2 emission estimates are 1.21 ± 0.~~72~~ 75 t/s and 1.12 ± 0.~~37~~ 38 t/s for the two flights, also with a difference of 9%, but with the morning flight results being higher. Finally, CO emissions from the USCB were calculated to be 10.1 ± 3.~~2~~ 6 kg/s and 10.7 ± ~~3~~4.~~9~~ 4 kg/s for flight A and B, respectively. The discrepancy between them is 6%.

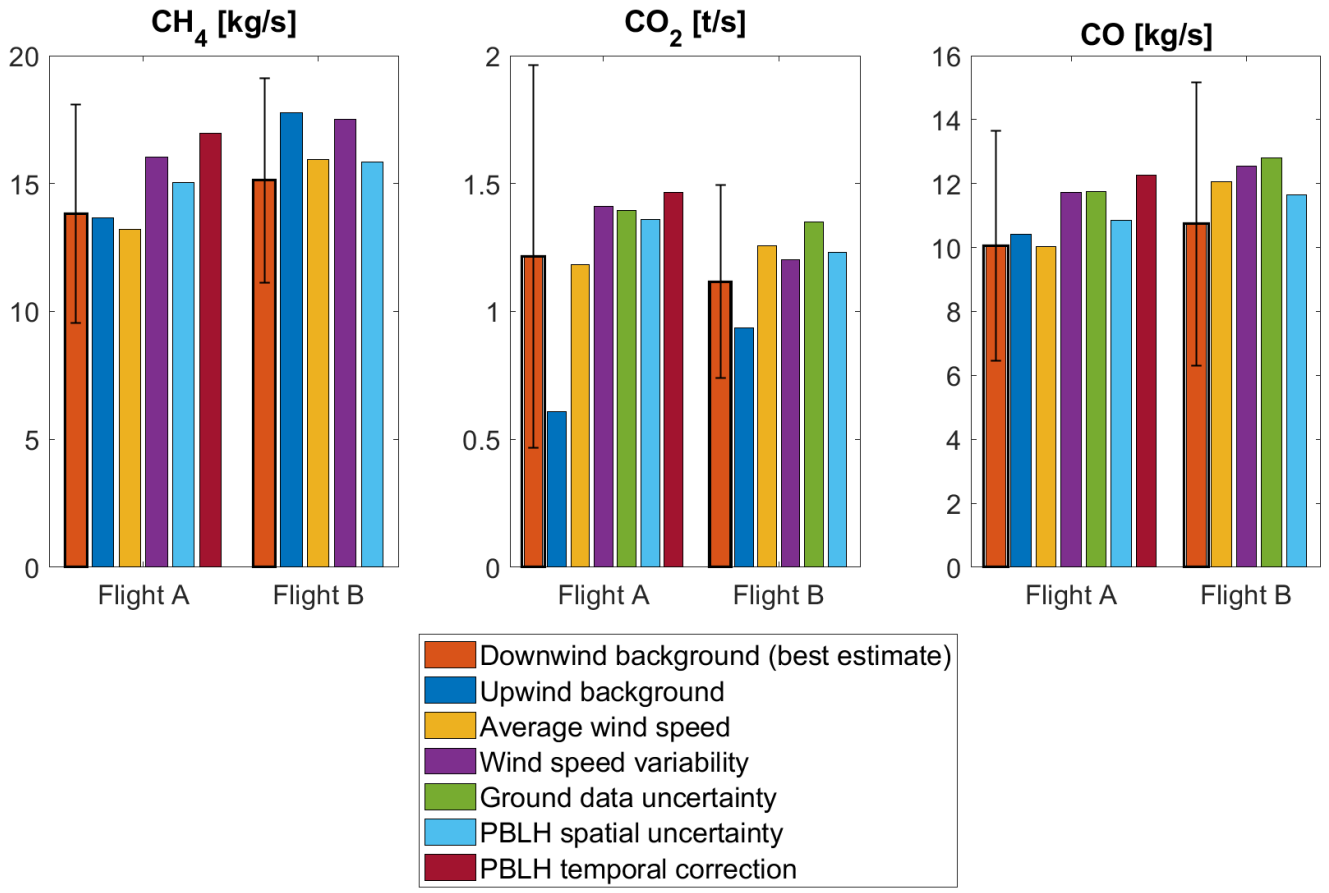


Figure 7: USCB emission estimates on June 6, 2018, using an airborne mass balance approach including several sensitivity tests.

We determined the systematic errors with several sensitivity tests applied to the treatment of different variables during the mass balance calculation (Figure 7). Systematic errors are calculated as emission difference between the best estimate mass balance using downwind background as described in Sect. 2.3 and the sensitivity studies:

1) Upwind background method

This background method leads to almost the same CH_4 emission estimate for flight A. The flight B estimate is 18% larger than the best estimate, showing that the assumption of a linear background gradient is not true for this case. The CO_2 emission estimate using an upwind background is 50% and 16% smaller than the best estimate for flights A and B, respectively. Especially for flight A, the upwind CO_2 mole fractions in the PBL might be enhanced due to a shallower PBLH. Also, the experiment was not conducted in a Lagrangian way, meaning that the sampling time difference between upwind and downwind does not match the travel time of the air. With potentially inhomogeneous biosphere-atmosphere fluxes this could cause a problem. For CO the upwind background method yields an emission estimate difference of 3% for

flight A. For flight B we did not calculate a CO emission estimate because of an incomplete upwind measurement (Figure 6). In general, CO upwind and downwind background data is quite similar.

2) Average wind speed

The impact of wind measurement treatment on the estimated mass fluxes was tested by using the averaged observed wind speed instead of the kriged wind field. This technique could for example be employed if no wind measurements were available and average model winds need to be used. The emission estimates for the morning flight are up to 4% lower and for the afternoon flight up to 13% higher than for the best estimate. Here the systematic change in the emission estimates is caused by the location of the plume in the wind field. During flight A, the plumes were located where the wind speed was slightly higher than average (see Figure 5). Using the average wind speed, thus, results in a reduction of the emission estimates. During flight B, the plume locations were in a slow wind region with higher wind speeds to the south, especially for the CO_2 and CO plume. Using averaged wind speed, thus, enhanced the emission estimate. We highlight the importance of measuring the wind speed simultaneously with the mole fractions and using this spatial knowledge in the flux calculation.

3) Wind speed variability

One assumption for a mass balance calculation is that the wind speed and direction are constant during the time it takes for the gases to be transported from the emission source to the observation location. In reality the wind field can be subject to considerable variability. In our case we were able to assess this temporal variability from the wind lidar observations. To account for wind variability we calculated the standard deviation of wind speed during the four hour transit time within the boundary layer and added it to the kriged wind field used in the mass balance calculation. This introduced an uncertainty of 17% and 15% to the morning and afternoon flight results, respectively.

4) Ground data uncertainty

Since we did not use CO_2 and CO from the mobile ground measurements, we calculated the sensitivity of our approach to the precise knowledge of ground-based data for CO_2 and CO. Assuming a 10% uncertainty of the ground value enhancements and increasing the kriging input ground values by this factor, results in a systematic error of 15-20%. This shows that a good approximation, or even better a measurement, of mole fractions below the lowest flight track is important for exact emission estimates.

5) PBLH uncertainty

Another sensitivity of our method is related to the knowledge of the PBLH and its variability. Its exact determination in the downwind wall is only possible when we cross its top during ascents or descents. This occurred once during the morning flight and three times in the afternoon. The PBLH is further constrained by vertical profiles before and after sampling the downwind wall and through the wind lidar observations. This data hints at temporal and spatial variations in the PBLH (see Sect. 3.1). Based on this data we assign an uncertainty estimate of 100 m to PBLH. We account for the spatial PBLH uncertainty in the emission estimate by using a boundary layer 100 m higher than our best estimate. This is realized through

cutting off the flux density field at this increased boundary. For this sensitivity test, discrepancies are between 5% and 12% for all three gases.

6) Temporal PBLH variability

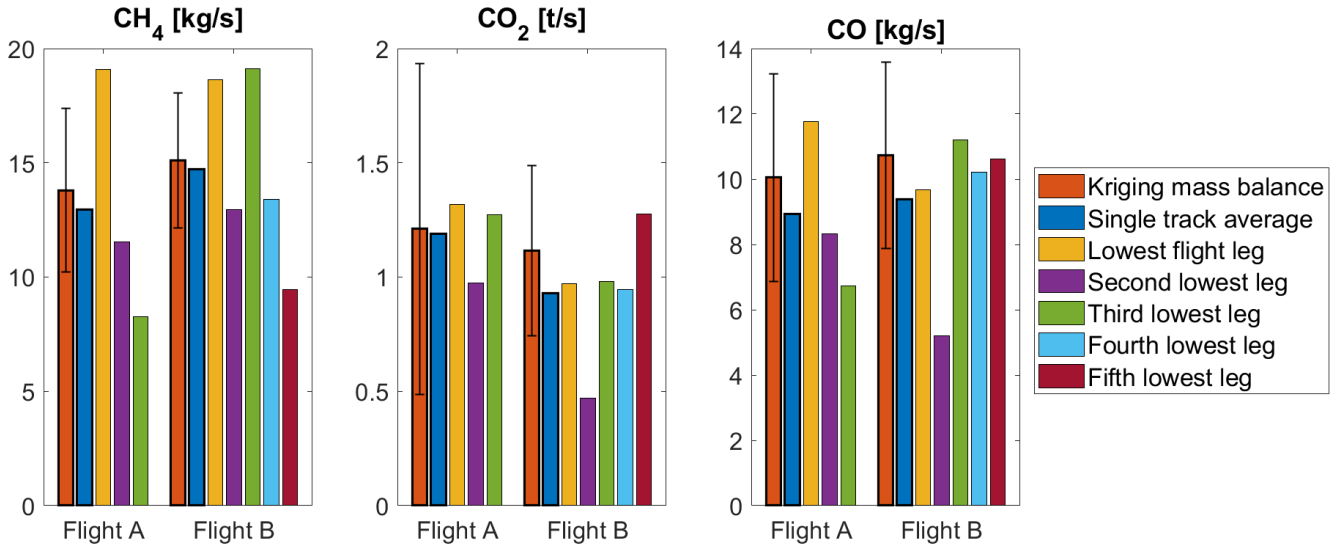
The last sensitivity test accounts for the temporal variation of the PBLH during the morning flight A. The PBLH showed a temporal variability of 300 m, quantifiable from wind lidar measurements. We assess the uncertainty caused by the temporally increasing PBLH for the morning flight by omitting the trace gas enhancement correction described in Sect. 3.2. The systematic error for flight A is between 21% and 23%.

On average, the uncertainty of the background mole fraction (up to 50%)~~, and~~ the uncertainty of mole fractions at the ground (15-20%), and the wind variability (15-17%) have the highest impact on the systematic uncertainty. For flight A, the changing PBLH introduces an additional 21-23% uncertainty to the emission estimates. Assuming that the single systematic uncertainties are independent of each other, the total systematic error of the emission estimate is calculated as the square root of the sum of squared individual uncertainties and is added to the statistical uncertainty. The statistical error is 1% for CH_4 and around 3% for CO_2 and CO and, thus, small compared to the systematic errors of this approach. It is added to the systematic error to obtain the total error of the emission estimates. The CH_4 emission estimate has a total relative error of ~~26~~31% and ~~21~~26%, the CO_2 estimate of ~~60~~62% and ~~33~~37% and the CO estimate of ~~32~~36% and ~~37~~41% for flights A and B, respectively. The errors are ~~always~~ mostly larger for flight A than for flight B, since the afternoon flight is more suitable for a mass balance experiment due to the temporally constant PBLH.

3.5 Single transect emission estimates

This detailed emission estimate, as described above, can help to understand uncertainties of a mass balance study in cases where less information is available. We ensured the validity of the mass balance technique by performing multiple vertical transects and even driving underneath the flight path to capture the signal at the surface layer. Many mass balance studies do not put in this level of effort, hence we can estimate how necessary these extra precautions are with regard to calculating the true emissions. Furthermore, when using mass balance techniques at any point to verify emissions from a policy enactment standpoint, resources should be used as efficient as possible. By using the information from single transects within the boundary layer of each flight we calculated the emissions under the simple assumption of a perfectly mixed boundary layer. The PBLH was kept constant for all transects. Figure 8 shows the results of the single transect mass balance calculations for the two flights on June 6, 2018. The average of the single transects (blue) is always well within the uncertainty range of the kriging mass balance results (red). Nevertheless, by assuming that one individual transect is representative for the entire PBL, transect emission estimates deviate up to 40% in both directions from the kriging estimate for CH_4. This deviation is much larger than the kriging estimate uncertainty. Deviations are largest for transects close to the PBLH when the concentration gradient between the boundary layer and free troposphere is also large, e.g. the highest CH_4 transects. Thus,

when calculating emissions from single transects the flight altitude should be well below the PBLH to avoid sampling free tropospheric air masses. On the other hand, these results discourage using single transect mass balance estimates anyway.



595 **Figure 8: Mass balance results for single transects through the plumes compared to the average of all single transects and the kriging mass balance result from Section 3.4.**

4 Comparison with bottom-up inventories

Hereafter we compare our airborne top-down emission estimate for the USCB with the bottom-up emission inventories
600 described in Sect. 2.1. Both emission values, the bottom-up inventory and the top-down mass balance estimate, are based on different methods and assumptions which hamper a one-by-one comparison. Especially differences in the temporal resolution of the two methods create a problem in case emissions are subject to strong temporal fluctuations such as a seasonal or diurnal cycle. Aircraft-borne top-down methods can only provide snapshot emission estimates, which for a comparison need to be scaled to the temporal resolution of the emission inventories. At the same time, bottom-up inventories
605 also include uncertainties, for example in the emission factors which are often derived from process studies and are then used to derive annual sums. For this comparison, we scale our mass balance emission estimate, based on a snapshot of one day in the early summer, to an annual emission estimate. We assume this scaling to be representative, because of the nature of the USCB emissions. In general, coal mining activities continue all year round and the power plants using the excavated coal are continually operated base load facilities. Still, it is known that CH_4 emissions from individual ventilation shafts vary on
610 weekly to monthly scale, when mines open new longwall excavation areas and ventilation increases. However, since we study emissions on a regional scale (including ~35 mines), we argue that emissions from individual shafts vary independently and therefore variations cancel out to a large extent. According to the CAMS inventory (Figure S6), industrial

emission, including coal mine exhaust, make up 87% of USCB CH_4 emissions, with the waste sector (11%) and fugitives (2%) being the other contributors. Thus, we assume our CH_4 emission estimate to be largely representative for the entire year. CO_2 emissions attributed to public power generation (65%) and residential heating (6%) do have an annual cycle. The other contributions include industry (21%) and transportation (7%). The CO emissions result to 30% from residential combustion with annual cycle with the remainder from public power (3%), industry (54%), and road transport (13%) without annual cycle. Thus, there is an annual cycle for CO_2 and CO and our summer measurements likely underestimate the annual value. Additionally, gridded inventories need to be treated with caution when used in region-specific studies (Janssens-Maenhout et al., 2019). These inventories distribute national emission totals onto a grid using proxy data. Most of the uncertainty of the grid cell level data originates from the uncertainty in the proxy data (Hogue et al., 2016). Furthermore, the comparison of inventories from 2010 with observational estimates from 2018 is not consistent and we treat comparisons to the GESAPU inventory with caution.

Our airborne mass balance CH_4 emission estimate on June 6, 2018, of 436 ± ~~115~~ 135 kt/a and 477 ± ~~101~~ 126 kt/a for flights A and B, respectively, is in the lower range of inventory emissions (Figure ~~8~~9). E-PRTR emission estimates are similar to our estimate, despite the omitted sources with emissions lower than the threshold of 0.1 kt/a. The CoMet emission inventory is higher than both mass balance estimates, but within the error range of flight B. Compared to E-PRTR from 2017, the CoMet inventory includes several mines in Poland that reported higher CH_4 emissions in 2016 than in 2017, three additional Czech mines, and four landfills within the mass balance area. Scarpelli, CAMS and EDGAR CH_4 estimates are also higher than our mass balance results. The GESAPU inventory states the lowest emissions, which may result from the missing emissions from Czech mines (estimated to be around 70 kt/a).

Our CO_2 aircraft mass balance emission estimates of 38.3 ± ~~22~~23.~~8~~ 6 Mt/a and 35.2 ± 11.~~7~~ 9 Mt/a agree with all inventories within the reported errors of the measurements. These errors are large, especially for the morning flight. Under very good conditions it is possible to report results that can inform about the quality of emission inventories, but issues like the biospheric fluxes of CO_2 and annual cycles of emissions impede comparisons to annual emission inventory values.

The CO emission estimates of 317 ± ~~100~~ 114 kt/a and 339 ± ~~90~~ 139 kt/a from the aircraft mass balance on June 6, 2018, are at the upper end of the emission inventories. Especially the E-PRTR emission estimate for 2017 is much lower than the mass balance result. This point source inventory does not include emissions from the transport and residential sector, which together comprise 42% of USCB CO emissions according to CAMS (Figure S6), which explains the discrepancy. CAMS, EDGAR, and GESAPU inventories are in the range of the emission estimates, but due to the annual cycle in residential combustion we suspect that these inventories underestimate CO emissions from the USCB.

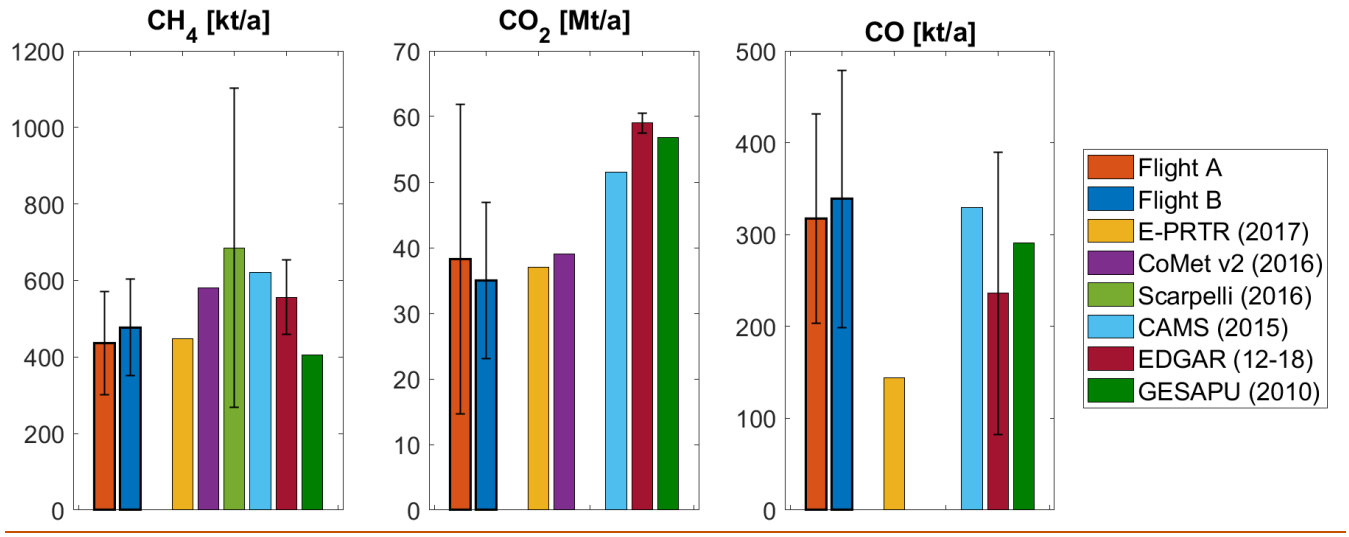


Figure 9~~8~~: Comparison of USCB emission estimates of the CoMet mass balance flights A and B with bottom-up emission inventories. Error bars show one standard deviation of the estimates, where available.

5 Summary and Conclusions

In times of rising atmospheric concentrations of greenhouse gases and countries trying to reduce their associated emissions, it is important to develop an independent and objective emission monitoring system. During the CoMet campaign the European CH_4 emission hot spot Upper Silesian Coal Basin (USCB) was sampled by in situ techniques as well as passive and active remote sensing on ground and from aircraft. From two flights A and B around the USCB, conducted on June 6, 2018, combined with vehicle-based ground measurements, we determined a regional emission estimate of CH_4, CO_2, and CO for the entire USCB using in situ data and a mass balance approach. The plumes of all three trace gases could be observed and separated from the atmospheric background in all downwind transects. For the morning flight A, a trace gas enhancement correction was employed to account for the temporal change of PBLH during the sampling. We employed a kriging algorithm for the interpolation of observed CH_4, CO_2, CO and wind speed between the flight transects and towards the ground. CH_4 ground-based observations confirmed the existence of a well-mixed PBL with similar trace gas enhancements at the ground and in the aircraft transects. From the kriged fields we calculated the USCB emission estimate as the mass flux through the downwind wall for each flight. Using error propagation and several sensitivity tests we carefully determined the total error of our mass-balance approach. The CH_4 emission estimate has a total relative error of ~~21~~26-~~26~~31%, the CO_2 estimate of ~~33~~37-~~60~~62% and the CO estimate of ~~32~~36-~~37~~41%. These uncertainties are mainly caused by the background determination, wind speed variability, and the missing knowledge of mole fractions below the lowest flight track for CO_2 and CO. The higher uncertainty values apply to the morning flight estimate, because the temporal variation of the PBLH introduced a large error. Thus, we highlight the importance of a constant PBLH over time, knowledge of trace gas

mole fractions at the ground and the exact knowledge of background mole fractions. The large uncertainties in the CO_2 estimate are dominated by the uncertainties in biospheric CO_2 fluxes. These estimates could be improved by performing flights in wintertime, when the biospheric fluxes are negligible. Flights during different seasons would also better constrain the annual cycle in CO_2 emissions from the residential sector. The calculation of emission estimates from single flight transects is not advisable, because the single transect estimates showed deviations from their mean and the kriging method of more than 40% in both directions.

The CoMet in situ CH_4 emissions estimates from June 6, 2018, of 13.8 ± ~~34~~.~~6~~ 3 kg/s and 15.1 ± ~~34~~.0 kg/s for flight A and B, respectively, are in the lower range of the six presented emission inventories. This agreement of our independent USCB emission estimate with the bottom-up coal mining emission reports indicates that this sector of emissions is well understood and monitored on regional scales. The emissions of CO_2 were determined to be 1.21 ± 0.~~72~~ 75 t/s and 1.12 ± 0.~~37~~ 38 t/s. The estimate from the second flight constrains the emissions to the lower end of inventory values. The gridded inventories, which report higher emissions than our estimate, do not include an annual cycle in the residential combustion emissions of CO_2. This might be reflected in our low summer emission estimate. In general, an airborne mass balance estimate for CO_2 on these spatial scales is difficult due to inhomogeneous biospheric uptake. CO mass balance emissions of 10.1 ± 3.~~2~~ 6 kg/s and 10.7 ± ~~34~~.~~9~~ 4 kg/s for the USCB on June 6, 2018, are much higher than the E-PRTR point source inventory, which does not include residential combustion and road transport emissions, and are still in the upper range of the gridded emission inventory values. The comparison between the snapshot top-down emission estimate and annual bottom-up inventories is influenced by the temporal variability of emissions in the USCB. Therefore, additional measurements during different seasons are needed to finally confirm bottom-up emission inventories.

Our airborne in situ mass balance method describes a measurement and evaluation strategy, which can be applied for various emission sources on a local to regional scale. In this case, we provide an independent bottom-up emission assessment for the USCB, which also serves as a point of reference for other state of the art techniques, like airborne lidar and passive spectroscopy. A comparison of in situ and remote sensing emission estimation techniques will follow in future studies.

Independent top-down validation of emissions in industrialized countries can confirm the statistical approaches used in bottom-up inventories. Once facility locations and activity, technology and abatement information becomes available for other countries or regions, the confirmed emissions from industrialized areas will help to improve global emission inventories used in climate projections. These will in turn help policy makers to develop efficient climate mitigation strategies. Consistent, reliable, and timely information on greenhouse gas emissions will allow the implementation, evaluation, and management of long-term policies that might allow keeping the global temperature rise below 2°C above pre-industrial levels.

Author contributions

Alina Fiehn, Julian Kostinek, and Maximilian Eckl performed the trace gas measurements, calibrations and data preparation. Alina Fiehn analyzed the data and drafted the manuscript. Theresa Klausner provided shaft-wise E-PRTR geolocation and emission data retrieved from the E-PRTR dataset and the Polish State Mining Authority for 2014. This dataset was updated and expanded by Michal Galkowski to the used version 2. Michal Galkowski, Jinxuan Chen, and Christoph Gerbig provided STILT and VPRM simulations and helpful discussions on biogenic uptake of CO_2. Thomas Röckmann coordinated the deployment of ground-based measurements and helped with data evaluation and interpretation. Hossein Maazallahi, Martina Schmidt, Piotr Korben, and Jaroslaw Necki conducted ground-based in situ measurements in the field during the campaign and collected and shared the data. Pawel Jagoda conducted ground-based observations and supported the aircraft observations through coordinating the airport communications at Rybnik airport. Norman Wildmann took an active part during the campaign deploying the wind lidar and retrieving and providing wind lidar data. Christian Mallaun supervised the wind measurements onboard the Cessna Caravan and prepared the data. Rostyslav Bun provided a gridded version of the GESAPU emission inventory for the USCB. Anna-Leah Nickl and Patrick Jöckel devised, set up, and supervised the forecasting system that allowed flight planning for the CoMet campaign in the USCB. Andreas Fix coordinated all CoMet campaign contributions. Anke Roiger developed the research idea and coordinated the CoMet Cessna campaign operations. All authors contributed to the interpretation of the results and the improvement of the manuscript.

Competing interests: The authors declare that they have no conflict of interest.

Acknowledgement: The authors especially thank DLR-FX for the campaign cooperation, especially the pilots Thomas van Marwick and Philipp Weber and the group of Ralph Helmes, Andreas Giez, Martin Zöger, and Martin Sedlmeir. We would like to thank Joseph Pitt for providing an updated version of the kriging package and giving advice on its usage. We acknowledge funding for the CoMet campaign by BMBF (German Federal Ministry of Education and Research) through AIRSPACE (FKZ grants no. 01LK1701A and 01LK1701C). We thank DLR VO-R for funding the young investigator research group "Greenhouse Gases". The ground-based measurements on vehicles were funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie ITN project Methane goes Mobile – Measurements and Modelling ($MEMO^2$; https://h2020-memo2.eu/) grant agreement no. 722479. The authors acknowledge ECCAD for archiving and distributing the CAMS emission inventories.

Data Accessibility: The data can be inquired directly from the authors and will be made public with the final manuscript.

References

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., Davis, K. J., Herndon, S. C., Jacob, D. J., Karion, A., Kort, E. A., Lamb, B. K., Lauvaux, T., Maasakkers, J. D., Marchese, A. J., Omara, M., Pacala, S. W., Peischl, J., Robinson, A. L., Shepson, P. B., Sweeney, C., Townsend-Small, A., Wofsy, S. C., and Hamburg, S. P.: Assessment of methane emissions from the U.S. oil and gas supply chain, Science, 361, 186-188, 10.1126/science.aar7204, 2018.

Amediek, A., Ehret, G., Fix, A., Wirth, M., Büdenbender, C., Quatrevalet, M., Kiemle, C., and Gerbig, C.: CHARM-F – a new airborne integrated-path differential-absorption lidar for carbon dioxide and methane observations: measurement performance and quantification of strong point source emissions, Appl. Opt., 56, 5182-5197, 10.1364/AO.56.005182, 2017.

Barkley, Z. R., Lauvaux, T., Davis, K. J., Deng, A., Fried, A., Weibring, P., Richter, D., Walega, J. G., DiGangi, J., Ehrman, S. H., Ren, X., and Dickerson, R. R.: Estimating Methane Emissions From Underground Coal and Natural Gas Production in Southwestern Pennsylvania, Geophysical Research Letters, 46, 4531-4540, 10.1029/2019gl082131, 2019.

Bergamaschi, P., Karstens, U., Manning, A. J., Saunois, M., Tsuruta, A., Berchet, A., Vermeulen, A. T., Arnold, T., Janssens-Maenhout, G., Hammer, S., Levin, I., Schmidt, M., Ramonet, M., Lopez, M., Lavric, J., Aalto, T., Chen, H., Feist, D. G., Gerbig, C., Haszpra, L., Hermansen, O., Manca, G., Moncrieff, J., Meinhardt, F., Necki, J., Galkowski, M., O'Doherty, S., Paramonova, N., Scheeren, H. A., Steinbacher, M., and Dlugokencky, E.: Inverse modelling of European CH4 emissions during 2006–2012 using different inverse models and reassessed atmospheric observations, Atmos. Chem. Phys., 18, 901-920, 10.5194/acp-18-901-2018, 2018.

Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S. W., Evan, S., McKeen, S. A., Hsie, E. Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NOx and CO2 and their impacts, Atmos. Chem. Phys., 13, 3661-3677, 10.5194/acp-13-3661-2013, 2013.

Bun, R., Nahorski, Z., Horabik-Pyzel, J., Danylo, O., See, L., Charkovska, N., Topylko, P., Halushchak, M., Lesiv, M., Valakh, M., and Kinakh, V.: Development of a high-resolution spatial inventory of greenhouse gas emissions for Poland from stationary and mobile sources, Mitigation and Adaptation Strategies for Global Change, 10.1007/s11027-018-9791-2, 2019.

Cambaliza, M. O. L., Shepson, P.B., Bogner, J., Caulton, D.R., Stirm, B., Sweeney, C., Montzka, S.A., Gurney, K.R., Spokas, K., Salmon, O.E., Lavoie, T.N., Hendricks, A., Mays, K., Turnbull, J., Miller, B.R., Lauvaux, T., Davis, K., Karion, A., Moser, B., Miller, C., Obermeyer, C., Whetstone, J., Prasad, K., Miles, N. and Richardson, S.: Quantification and source apportionment of the methane emission flux from the city of Indianapolis, Elem Sci Anth, 3, 000037, 10.12952/journal.elementa.000037, 2015.

Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D. H., Sweeney, C., Herndon, S., Schwietzke, S., Pétron, G., Pifer, J., Kort, E. A., and Schnell, R.: Application of Gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, Atmos. Meas. Tech., 10, 3345-3358, 10.5194/amt-10-3345-2017, 2017.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., and Janssens-Maenhout, G.: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2, Earth Syst. Sci. Data, 10, 1987-2013, 10.5194/essd-10-1987-2018, 2018.

Cui, Y. Y., Brioude, J., McKeen, S. A., Angevine, W. M., Kim, S.-W., Frost, G. J., Ahmadov, R., Peischl, J., Bousserez, N., Liu, Z., Ryerson, T. B., Wofsy, S. C., Santoni, G. W., Kort, E. A., Fischer, M. L., and Trainer, M.: Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin, Journal of Geophysical Research: Atmospheres, 120, 6698-6711, 10.1002/2014JD023002, 2015.

EEA: European Pollutant Release and Transfer Register (E-PRTR): https://www.eea.europa.eu/data-and-maps/data/member-states-reporting-art-7-under-the-european-pollutant-release-and-transfer-register-e-prtr-regulation-18, access: 20 September 2019, 2019.

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version), Copernicus Atmosphere Monitoring Service (CAMS), 2019.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415, 626-630, 10.1038/415626a, 2002.

Hausmann, P., Sussmann, R., and Smale, D.: Contribution of oil and natural gas production to renewed increase in atmospheric methane (2007–2014): top–down estimate from ethane and methane column observations, Atmos. Chem. Phys., 16, 3227-3244, 10.5194/acp-16-3227-2016, 2016.

Heimburger, A., Harvey, R. M., Shepson, P. B., Stirm, B. H., Gore, C., Turnbull, J., Cambaliza, M. O. L., Salmon, O. E., Kerlo, A.-E. M., Lavoie, T. N., Davis, K. J., Lauvaux, T., Karion, A., Sweeney, C., Brewer, W. A., Hardesty, R. M., and Gurney, K. R.:

Assessing the optimized precision of the aircraft mass balance method for measurement of urban greenhouse gas emission rates through averaging, Elem Sci Anth, 5, https://doi.org/10.1525/elementa.134, 2017.

Hiller, R. V., Neininger, B., Brunner, D., Gerbig, C., Bretscher, D., Künzle, T., Buchmann, N., and Eugster, W.: Aircraft-based CH4 flux estimates for validation of emissions from an agriculturally dominated area in Switzerland, Journal of Geophysical Research: Atmospheres, 119, 4874-4887, 10.1002/2013jd020918, 2014.

Hogue, S., Marland, E., Andres, R. J., Marland, G., and Woodard, D.: Uncertainty in gridded CO2 emissions estimates, Earth's Future, 4, 225-239, 10.1002/2015ef000343, 2016.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, Earth Syst. Sci. Data, 11, 959-1002, 10.5194/essd-11-959-2019, 2019.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., 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, Geophysical Research Letters, 40, 4393-4397, doi:10.1002/grl.50811, 2013.

Karion, A., Sweeney, C., Kort, E. A., Shepson, P. B., Brewer, A., Cambaliza, M., Conley, S. A., Davis, K., Deng, A., Hardesty, M., Herndon, S. C., Lauvaux, T., Lavoie, T., Lyon, D., Newberger, T., Pétron, G., Rella, C., Smith, M., Wolter, S., Yacovitch, T. I., and Tans, P.: Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region, Environmental Science & Technology, 49, 8124-8131, 10.1021/acs.est.5b00217, 2015.

Klausner, T., Mertens, M., Huntrieser, H., Baumann, R., Galkowski, M., Kuhlmann, G., Fiehn, A., Jöckel, P., Pühl, M., and Roiger, A.: Urban greenhouse gas emissions from the Berlin area: A case study on aircraft-based CO2 and CH4 in situ observations in summer 2018, Elementa, 2020.

Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehrkorn, T., Daube, B. C., Kaplan, J. O., Houweling, S., and Wofsy, S. C.: Emissions of CH4 and N2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations, Geophysical Research Letters, 35, 10.1029/2008gl034031, 2008.

Kostinek, J., Roiger, A., Davis, K. J., Sweeney, C., DiGangi, J. P., Choi, Y., Baier, B., Hase, F., Groß, J., Eckl, M., Klausner, T., and Butz, A.: Adaptation and performance assessment of a quantum and interband cascade laser spectrometer for simultaneous airborne in situ observation of CH4, C2H6, CO2, CO and N2O, Atmos. Meas. Tech., 12, 1767-1783, 10.5194/amt-12-1767-2019, 2019.

Krautwurst, S., Gerilowski, K., Jonsson, H. H., Thompson, D. R., Kolyer, R. W., Iraci, L. T., Thorpe, A. K., Horstjann, M., Eastwood, M., Leifer, I., Vigil, S. A., Krings, T., Borchardt, J., Buchwitz, M., Fladeland, M. M., Burrows, J. P., and Bovensmann, H.: Methane emissions from a Californian landfill, determined from airborne remote sensing and in situ measurements, Atmos. Meas. Tech., 10, 3429-3452, 10.5194/amt-10-3429-2017, 2017.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, 10.1029/2003gb002199, 2005.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003-2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963-10976, 10.5194/acp-14-10963-2014, 2014.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres, 108, 10.1029/2002jd003161, 2003.

Luther, A., Kleinschek, R., Scheidweiler, L., Defratyka, S., Stanisavljevic, M., Forstmaier, A., Dandocsi, A., Wolff, S., Dubravica, D., Wildmann, N., Kostinek, J., Jöckel, P., Nickl, A. L., Klausner, T., Hase, F., Frey, M., Chen, J., Dietrich, F., Nęcki, J., Swolkień, J., Fix, A., Roiger, A., and Butz, A.: Quantifying CH4 emissions from hard coal mines using mobile sun-viewing Fourier transform spectrometry, Atmos. Meas. Tech., 12, 5217-5230, 10.5194/amt-12-5217-2019, 2019.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochemical Cycles, 22, 10.1029/2006gb002735, 2008.

Mallaun, C., Giez, A., and Baumann, R.: Calibration of 3-D wind measurements on a single-engine research aircraft, Atmos. Meas. Tech., 8, 3177-3196, 10.5194/amt-8-3177-2015, 2015.

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, Environmental Science & Technology, 43, 7816-7823, 10.1021/es901326b, 2009.

Nickl, A. L., Mertens, M., Roiger, A., Fix, A., Amediek, A., Fiehn, A., Gerbig, C., Galkowski, M., Kerkweg, A., Klausner, T., Eckl, M., and Jöckel, P.: Forecasting of regional methane from coal mine emissions in the Upper Silesian Coal Basin using the on-line nested global regional chemistry climate model MECO(n)(MESSy v2.53), Geosci. Model Dev. Discuss., 2019, 1-29, 10.5194/gmd-2019-303, 2019.

Nisbet, E. G., Dlugokencky, E. J., and Bousquet, P.: Methane on the Rise—Again, Science, 343, 493-495, 10.1126/science.1247828, 2014.
Nisbet, E. G., Dlugokencky, E. J., Manning, M. R., Lowry, D., Fisher, R. E., France, J. L., Michel, S. E., Miller, J. B., White, J. W. C., Vaughn, B., Bousquet, P., Pyle, J. A., Warwick, N. J., Cain, M., Brownlow, R., Zazzeri, G., Lanoisellé, M., Manning, A. C., Gloor, E., Worthy, D. E. J., Brunke, E.-G., Labuschagne, C., Wolff, E. W., and Ganesan, A. L.: Rising atmospheric methane: 2007–2014 growth and isotopic shift, Global Biogeochemical Cycles, 30, 1356-1370, 10.1002/2016gb005406, 2016.
Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very Strong Atmospheric Methane Growth in the 4 Years 2014–2017: Implications for the Paris Agreement, Global Biogeochemical Cycles, 33, 318-342, 10.1029/2018gb006009, 2019.
O'Shea, S. J., Allen, G., Gallagher, M. W., Bower, K., Illingworth, S. M., Muller, J. B. A., Jones, B. T., Percival, C. J., Bauguitte, S. J. B., Cain, M., Warwick, N., Quiquet, A., Skiba, U., Drewer, J., Dinsmore, K., Nisbet, E. G., Lowry, D., Fisher, R. E., France, J. L., Aurela, M., Lohila, A., Hayman, G., George, C., Clark, D. B., Manning, A. J., Friend, A. D., and Pyle, J.: Methane and carbon dioxide fluxes and their regional scalability for the European Arctic wetlands during the MAMM project in summer 2012, Atmos. Chem. Phys., 14, 13159-13174, 10.5194/acp-14-13159-2014, 2014.
Pandey, S., Gautam, R., Houweling, S., van der Gon, H. D., Sadavarte, P., Borsdorff, T., Hasekamp, O., Landgraf, J., Tol, P., van Kempen, T., Hoogeveen, R., van Hees, R., Hamburg, S. P., Maasakkers, J. D., and Aben, I.: Satellite observations reveal extreme methane leakage from a natural gas well blowout, Proceedings of the National Academy of Sciences, 116, 26376-26381, 10.1073/pnas.1908712116, 2019.
Pitt, J. R., Allen, G., Bauguitte, S. J. B., Gallagher, M. W., Lee, J. D., Drysdale, W., Nelson, B., Manning, A. J., and Palmer, P. I.: Assessing London CO2, CH4 and CO emissions using aircraft measurements and dispersion modelling, Atmos. Chem. Phys., 19, 8931-8945, 10.5194/acp-19-8931-2019, 2019.
Plant, G., Kort, E. A., Floerchinger, C., Gvakharia, A., Vimont, I., and Sweeney, C.: Large Fugitive Methane Emissions From Urban Centers Along the U.S. East Coast, Geophysical Research Letters, 46, 8500-8507, 10.1029/2019gl082635, 2019.
Polson, D., Fowler, D., Nemitz, E., Skiba, U., McDonald, A., Famulari, D., Di Marco, C., Simmons, I., Weston, K., Purvis, R., Coe, H., Manning, A. J., Webster, H., Harrison, M., O'Sullivan, D., Reeves, C., and Oram, D.: Estimation of spatial apportionment of greenhouse gas emissions for the UK using boundary layer measurements and inverse modelling technique, Atmospheric Environment, 45, 1042-1049, 10.1016/j.atmosenv.2010.10.011, 2011.
Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, Atmos. Meas. Tech., 6, 837-860, 10.5194/amt-6-837-2013, 2013.
Ren, X., Salmon, O. E., Hansford, J. R., Ahn, D., Hall, D., Benish, S. E., Stratton, P. R., He, H., Sahu, S., Grimes, C., Heimburger, A. M. F., Martin, C. R., Cohen, M. D., Stunder, B., Salawitch, R. J., Ehrman, S. H., Shepson, P. B., and Dickerson, R. R.: Methane Emissions From the Baltimore-Washington Area Based on Airborne Observations: Comparison to Emissions Inventories, Journal of Geophysical Research: Atmospheres, 123, 8869-8882, doi:10.1029/2018JD028851, 2018.
Ryoo, J. M., Iraci, L. T., Tanaka, T., Marrero, J. E., Yates, E. L., Fung, I., Michalak, A. M., Tadić, J., Gore, W., Bui, T. P., Dean-Day, J. M., and Chang, C. S.: Quantification of CO2 and CH4 emissions over Sacramento, California, based on divergence theorem using aircraft measurements, Atmos. Meas. Tech., 12, 2949-2966, 10.5194/amt-12-2949-2019, 2019.
Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., Dlugokencky, E. J., Etiope, G., Bastviken, D., Houweling, S., Janssens-Maenhout, G., Tubiello, F. N., Castaldi, S., Jackson, R. B., Alexe, M., Arora, V. K., Beerling, D. J., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Crevoisier, C., Crill, P., Covey, K., Frankenberg, C., Gedney, N., Höglund-Isaksson, L., Ishizawa, M., Ito, A., Joos, F., Kim, H. S., Kleinen, T., Krummel, P., Lamarque, J. F., Langenfelds, R., Locatelli, R., Machida, T., Maksyutov, S., Melton, J. R., Morino, I., Naik, V., O'Doherty, S., Parmentier, F. J. W., Patra, P. K., Peng, C., Peng, S., Peters, G. P., Pison, I., Prinn, R., Ramonet, M., Riley, W. J., Saito, M., Santini, M., Schroeder, R., Simpson, I. J., Spahni, R., Takizawa, A., Thornton, B. F., Tian, H., Tohjima, Y., Viovy, N., Voulgarakis, A., Weiss, R., Wilton, D. J., Wiltshire, A., Worthy, D., Wunch, D., Xu, X., Yoshida, Y., Zhang, B., Zhang, Z., and Zhu, Q.: Variability and quasi-decadal changes in the methane budget over the period 2000–2012, Atmos. Chem. Phys., 17, 11135-11161, 10.5194/acp-17-11135-2017, 2017.
Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murgia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y.,

Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000-2017, Earth Syst. Sci. Data Discuss., 2019, 1-136, 10.5194/essd-2019-128, 2019.

Scarpelli, T. R., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Sheng, J. X., Rose, K., Romeo, L., Worden, J. R., and Janssens-Maenhout, G.: A global gridded (0.1° × 0.1°) inventory of methane emissions from oil, gas, and coal exploitation based on national reports to the United Nations Framework Convention on Climate Change, Earth Syst. Sci. Data, 12, 563-575, 10.5194/essd-12-563-2020, 2020.

Schaefer, H., Fletcher, S. E. M., Veidt, C., Lassey, K. R., Brailsford, G. W., Bromley, T. M., Dlugokencky, E. J., Michel, S. E., Miller, J. B., Levin, I., Lowe, D. C., Martin, R. J., Vaughn, B. H., and White, J. W. C.: A 21st-century shift from fossil-fuel to biogenic methane emissions indicated by 13CH4, Science, 352, 80-84, 10.1126/science.aad2705, 2016.

Tadić, J. M., Michalak, A. M., Iraci, L., Ilić, V., Biraud, S. C., Feldman, D. R., Bui, T., Johnson, M. S., Loewenstein, M., Jeong, S., Fischer, M. L., Yates, E. L., and Ryoo, J.-M.: Elliptic Cylinder Airborne Sampling and Geostatistical Mass Balance Approach for Quantifying Local Greenhouse Gas Emissions, Environmental Science & Technology, 51, 10012-10021, 10.1021/acs.est.7b03100, 2017.

Thompson, R. L., Ishijima, K., Saikawa, E., Corazza, M., Karstens, U., Patra, P. K., Bergamaschi, P., Chevallier, F., Dlugokencky, E., Prinn, R. G., Weiss, R. F., O'Doherty, S., Fraser, P. J., Steele, L. P., Krummel, P. B., Vermeulen, A., Tohjima, Y., Jordan, A., Haszpra, L., Steinbacher, M., Van der Laan, S., Aalto, T., Meinhardt, F., Popa, M. E., Moncrieff, J., and Bousquet, P.: TransCom N2O model inter-comparison – Part 2: Atmospheric inversion estimates of N2O emissions, Atmos. Chem. Phys., 14, 6177-6194, 10.5194/acp-14-6177-2014, 2014.

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, 10.5194/acp-11-705-2011, 2011.

Turner, A. J., Frankenberg, C., Wennberg, P. O., and Jacob, D. J.: Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl, Proceedings of the National Academy of Sciences, 114, 5367-5372, 10.1073/pnas.1616020114, 2017.

UNFCCC: Adoption of the Paris Agreement, Conference of the Parties: Twenty-first session, Paris, France, 30 November–13 December 2015, 2015.

Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S., Aben, I., Scarpelli, T., and Jacob, D. J.: Satellite Discovery of Anomalously Large Methane Point Sources From Oil/Gas Production, Geophysical Research Letters, 46, 13507-13516, 10.1029/2019gl083798, 2019.

Wildmann, N., Päschke, E., Roiger, A., and Mallaun, C.: Towards improved turbulence estimation with Doppler wind lidar VAD scans, Atmos. Meas. Tech. Discuss., 2020, 1-30, 10.5194/amt-2020-8, 2020.

WMO: WMO Greenhouse Gas Bulletin, The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2018, 2019.

Worden, J. R., Bloom, A. A., Pandey, S., Jiang, Z., Worden, H. M., Walker, T. W., Houweling, S., and Röckmann, T.: Reduced biomass burning emissions reconcile conflicting estimates of the post-2006 atmospheric methane budget, Nature Communications, 8, 2227, 10.1038/s41467-017-02246-0, 2017.

Xiang, B., Miller, S. M., Kort, E. A., Santoni, G. W., Daube, B. C., Commane, R., Angevine, W. M., Ryerson, T. B., Trainer, M. K., Andrews, A. E., Nehrkorn, T., Tian, H., and Wofsy, S. C.: Nitrous oxide (N2O) emissions from California based on 2010 CalNex airborne measurements, Journal of Geophysical Research: Atmospheres, 118, 2809-2820, 10.1002/jgrd.50189, 2013.

Yuan, B., Kaser, L., Karl, T., Graus, M., Peischl, J., Campos, T. L., Shertz, S., Apel, E. C., Hornbrook, R. S., Hills, A., Gilman, J. B., Lerner, B. M., Warneke, C., Flocke, F. M., Ryerson, T. B., Guenther, A. B., and de Gouw, J. A.: Airborne flux measurements of methane and volatile organic compounds over the Haynesville and Marcellus shale gas production regions, Journal of Geophysical Research: Atmospheres, 120, 6271-6289, 10.1002/2015jd023242, 2015.