

Reviewer 1:

We would like to thank the reviewer for the helpful and detailed comments. In the following we carefully address the comments point by point. First we repeat the reviewer's comment (R#...) and then give our response (A).

R#1: This paper describes an attempt to estimate CO₂ emissions from coal-burning power plants in Germany using either airborne remote sensing measurements or airborne in-situ sampling within the plume. The data analysis and emission estimates use either a mass-balance approach or a Gaussian plume assumption. The uncertainty analysis leads to relative errors on the order of 10-15% and the estimate comparison to reported emissions are consistent with this relative errors. This paper contains a lot of interesting material for an evaluation of the potential and difficulty of such approaches for the estimate of CO₂ emissions from point sources. It should eventually be published. On the other hand, I have been very disappointed by the manuscript presentation that looks more like a experiment report than a scientific paper. Also, the paper lacks AMTD Interactive comment Printer-friendly version Discussion paper conciseness and many figures are not necessary. Although the paper combines the in-situ and the remote-sensing approaches, there is no real discussion of the pros and cons of both.

A: We added a discussion on the pros and cons of the respective methods in the revised manuscript:

This case study illustrates the advantages and disadvantages of the used methods. The remote sensing approach offers the possibility to perform many flight legs in a short period of time. This is necessary to reduce the uncertainty as can also be seen from Fig. 14. The multiple transects allow for the application of the Gaussian plume model to a multi source setup which simultaneously retrieves the emission rates from several sources.

While for MAMAP the plume model usually utilises a priori information on the source location, an imaging instrument with sufficient spatial resolution and sensitivity (similar to, for example, AVIRIS-NG (Thompson et al., 2015; Frankenberg et al., 2016), though having a lower sensitivity compared to MAMAP) is able to determine the source location from the data directly and can acquire more data on shorter time scales potentially reducing uncertainties on derived emission estimates. Furthermore imaging instruments offer the possibility of mapping large areas in a survey for unknown sources.

However, there is generally the need for wind information which originates from models and/or in-situ measurements. The analysis in this study shows an overestimated model wind speed of about 6% (or about 0.4m/s) which is smaller than the uncertainty on wind speed. So in this case relying on the model alone may be sufficient. In a former study of similar setup (Krings et al., 2013) the error was about 10% or (0.7m/s). A wider and systematic analysis on the accuracy of the model wind is needed to assess to what extent additional wind (profile) measurements are dispensable. This will also become more relevant with regard to observations of localized sources by current and upcoming satellite missions with increased accuracy and spatial resolution. In these cases additional wind measurements will generally not be available.

The remote sensing instrument MAMAP measures solar backscattered electromagnetic radiation in the short-wave infrared. To simplify the radiative transfer calculations, cloud free atmospheres are selected to avoid the radiative transfer issues associated with solar

electromagnetic radiation passing through clouds. The selection of the measurement day for this study was largely driven by this requirement. This generally involves a more convective and therefore thicker boundary layer making the gathering and analysis of the in-situ measurements more complex.

In contrast to the remote sensing measurements of the entire vertical column, in-situ measurements need to sample the plume with flight legs at different altitude levels. As a result of the time needed to complete a representative vertical cross section of measurements, only a limited number of repeated measurements are typically feasible. Interpolations within the cross sections and extrapolations to the surface and sometimes to the top of the plume have to be applied. This also applies for this study, where the boundary layer reached into restricted airspace. However, the in-situ method has the advantage of delivering vertically and horizontally resolved information in conjunction with co-located wind information, which can be readily used to infer a flux estimate. The high intrinsic sensitivity enables the detection of elevated trace gas levels also at great distances to the source. Errors on the inversion results from in-situ and remote sensing data are rather similar.

R#2: Although the in-situ wind speeds are used for the interpretation of the remote sensing data, the in-situ concentration observations (Figure 5) should have been used, I think, for a discussion on the validity of the Gaussian plume model.

A: The Gaussian shape of the vertical distribution can only be observed close to the source. Due to the “reflection” of the plume off the surface and off the top of the boundary layer, the CO₂ gets mixed rather rapidly. About 2 km downwind of the source the CO₂ is well mixed according to the Gaussian model, i.e. on average and not necessarily for any snapshot in time. The in-situ measurements, for example, downwind of power plant Niederaußem (Fig. 5) are more than 2 km downwind of the source so that no distinct Gaussian shape in the vertical concentration profile is to be expected.

For better assessment of validity of the horizontal plume model, on the other hand, which is used to infer the emissions, the model result is now shown in addition to the data for the individual remote sensing flight legs (see also R#48).

R#3: It is really not clear why the emission from the Frimmersdorf power plant could not be estimated with the remote sensing technique (P17-18). Indeed, there are many flight tracks downwind of this power plant that, in principle, could be used for that purpose. I assume that the authors have attempted an inversion, with no success, so that they chose to discard this estimate. Their experience on that particular aspect should be clearly stated to help in the design of future similar campaigns. Perhaps only flight tracks within a few kilometers from the emission can be used?

A: In this study, we previously discarded the flight tracks further downwind due to large data gaps as stated in the main text (p17 L17 – p18 L3). In addition, the measured concentrations downwind of Frimmersdorf are an integrated composite of all sources upwind. For example, the power plant Niederaußem is more than 10 km upwind of these tracks so that part of the enhancements might already be dispersed to the flanks of the transects which we require for normalization – as mentioned in the text.

However, as the reviewer encourages us to investigate the measurements from further downwind, we relaxed the signal threshold for the first three tracks downwind of power plant Frimmersdorf to a minimum of 3000 counts and the inclination filter to 15°. In this way we ensure that a sufficient set of measurements, even if of lower quality, are available for interpretation. The mass balance flux estimates for these tracks are shown in an updated result

plot. Although the result shows some scatter, the average is reasonable. Since we did not apply our usual quality filter, these results have to be interpreted with more caution.

We did not apply the Gaussian plume method for these data as that would require to mix data which were subject to different filter criteria.

A brief discussion has been added to the revised manuscript.

R#4: The paper is strangely organised. The method for the remote sensing approach is mostly described in the “Result” section. Besides, it is rather strange to have in situ measurements, such as Figure 10, presented in the Remote Sensing section rather than the In situ section.

A: Agreed, the description of the remote sensing method was moved to Section 4. The original reason to have Figure 10 in the remote sensing section was that it is specifically referred to here. The Figure was removed for the revised version. Figure 11, however, was left in the remote sensing section. Although it shows in-situ data, the plot specifically addresses the remote sensing analysis.

R#5: There is a need to show early in the manuscript (section 3) the flight track (both in-situ and remote sensing), similar to Figure 12, as well as the location of the “virtual wall” that was chosen for the mass balance estimate

A: The new figures 1 to 3 are providing this now in a clear manner. The concept of the 'wall' was abandoned. We describe it now as cross-sections with maximum distances from which a projection along the wind was allowed. The concept is the same, but, in other words. This point is commented later again.

R#6: In the “wall” approach (Figure 4), I could not understand why several cells are considered in the along wind direction. Why not assume that the cell dimensions are z_{res} (vertical) x h_{res} (cross wind) x d (along wind).

A: As mentioned in the answer to R#5, the wording has changed. Along with this are the cross-sections in the new Figures 3, and 4 to 8. All these examples are with real data, and not anymore a schematic figure like Fig. 4, that was obviously confusing.

R#7: I could not understand the discussion on page 7 lines 27-30. The dimension of the wall is not provided.

A: See answer to R#6.

R#8: Detailed comments: P2L15: Are you suggesting that thermal infrared observations provide valid concentration estimates in the presence of clouds?

A: Yes, as long as the clouds are not in between target and instrument optics. However interpretation of thermal infrared measurements depends on the thermal contrast, as there is no signal if, for example, the CO₂ has the same temperature as the surface (Young, 2002).

In contrast, short-wave infrared observations as used in this study require backscattered sunlight. Clouds may block the solar electromagnetic radiation or increase the radiative transfer complexity in the determination of the path of the electromagnetic radiation due to (multiple) scattering.

R#9: P4L12 : Please provide a valid argumentation why the method used in the manuscript is better than krigging

A: As a main part of the revision, we did the whole calculations with our linear inter- and extrapolation method with only four rules, and with Kriging. There are two aspects to distinguish: (i) about the inter- and extrapolation. By using Kriging as another method we have shown that the difference is small. (ii) More important seems to be averaging and interpolation of fluxes, instead of averaging mass- and wind-fields before calculating the fluxes. Especially when the latter was done by Kriging, the deviation from the ensemble of other solutions is increasing, most likely due to the fact, that small artefacts in the individual fields are increasing the errors. Bottom line: Our method is not better than Kriging, but, we should not regard Kriging as the only option. This is also true after studying Gordon et al. (2015) in detail. We think that the new text is much clearer in showing and discussing these details. Since a complete revision was performed, and the separation of more individual sources was possible, the results as displayed in table 3 and Figure 18 were updated. The details are presented in a separate supplement

R#10: P7L16-22. Not clear why there is a need to have the virtual wall oriented precisely crosswind. It seems more important to have the wall aligned with the flight tracks

A: We disagree with. Aligning with the flight track is possible as long as there is only one track, or several perfectly stacked above each other. Then it does not make a difference. However, in a real case, with a flight pattern that was not ideal by several reasons, it is very important to have cross sections exactly perpendicular to the wind during the time of observations, because otherwise, maxima on different flight tracks would add in different grid cells. This can be avoided when the projection to the cross-sections (we do not call it 'wall' anymore) is along the wind, to a perpendicular plane. The new figures 2 and 3 should explain this.

R#11: P8L9: Could not understand

A: Should now be clear with the new explanations about the inter- and extrapolations, and the percentage of directly measured fluxes in relation to the extrapolations below and above the flight tracks.

R#12: P9L2-3 : Could not understand

A: When the (systematic) error in the wind speed measurement is 0.5 m/s, this would modify the total flux in a 5 m/s wind by 10%. It is less if the error is non-systematic. However, this is a worst-case estimate. With the same 0.5 m/s in error, a flux in a 10 m/s would be wrong by 5% only, but, by 25% in a flow of 2 m/s. On the other hand, 0.5 ppm error would only contribute an error of 1% for the flux in a typical moderate plume with 50 ppm enhanced CO₂. We argue here, that under the conditions of these measurements (plume enhancements usually higher than 50 ppm) and wind speeds around or above 5 m/s, the errors of the measurements (instrumental errors, both systematic and stochastic) are contributing a maximum of 10 % and that the wind is more critical than the concentrations. This finding is well in agreement with Gordon et al. (2015).

R#13: P9L6: What about the sampling of the plume and its variability ? What is the variability of the concentration with the wall cells?

A: This is now clearly shown in Figures 4 to 8, and with the initial data in Figures 2 and 3.

R#14: P9L15: It is rather difficult to understand that there is an uncertainty about the top of the mixing layer, but not on the flux close to the surface. Either one assumes that there is little vertical mixing, in which case only the flight track at a level close to that of the chimney, or there is mixing that transfers CO₂ both high in the mixing layer and towards the surface (see P9L30).

A: Primarily we show that the amount of fluxes coming from the extrapolation is only 10, and 14% in the two budgets that were directly measured (more in those that were derived as sums or differences). The different methods are discussed: Fluxes or concentrations staying constant or diminishing to zero above background, etc. Finally we attributed half of the extrapolated amounts to the overall error (Table 3), which means that the extrapolation has an uncertainty of 50%. More details are discussed in the revised text and should be clearer from the new figures.

R#15: P9L20: Instationarity of the source is mentioned. What is the variability of the source according to the power plant management ?

A: We added the information to the revised version. For this case study, the source variation based on energy production was less than 0.5% for Niederaußem, Neurath new and old blocks. For Frimmersdorf the variability was about 4% but with considerably lower total fluxes.

R#16: P10L5-10. Although it is not stated clearly, I understand that the discussion is for various days. The paper should rather provide the result for the particular day that is analysed in the manuscript, and make a single sentence for the other days.

A: Except for the reference case of power plant Weisweiler in the new Table 3, all results and discussions are about this specific day. This is clearer now in the revised version.

R#17: P10L30 “to the top of the well mixed boundary layer”. In situ measurements shown in the manuscript (Figures 4 and 5) clearly show that the boundary layer is not “well mixed”

A: We agree, that "well mixed" was not the best expression to describe the situation within the plume relatively close to the sources. We do not use it anymore in this context. However, it applies for the boundary layer in terms of water vapor or aerosols on the regional scale, enabling us to estimate the top of the actual convective boundary layer. Convective dispersion was evidently acting within this layer below 1300 mAMSL. Please also note that Fig. 4 was conceptual (now replaced). The heterogeneity of the plumes is clearly stated, and is the prime reason for the method applied for the inter- and extrapolation.

R#18: P12L1. Although section 5.2 is supposed to show “results”, it actually mostly describes the method.

A: Agreed, this section was moved to the methods chapter.

R#19: P12L24. Justification not clear. 0.9% relative to what ?

A: The RMS is relative to the model: $\text{RMS}[(\text{model}-\text{measurement})/\text{model}]$ where the choice of 0.9% follows from Figure 6. This is where a strong decrease in fit quality begins. However, also from Figure 6 it can be gathered that not many data is affected by this. As the reviewer is of the opinion Figure 6 is “definitely not useful” (see below) we removed the Figure and added following lines to the revised version:

Filtering, based on the spectroscopic fit quality, has been applied rejecting measurements with a root mean square (RMS) value of the differences between measurement and model after the fit larger than 0.9% relative to the model affecting about 0.1% of the total measurements. The threshold was empirically determined from the distribution of RMS values ordered by size (compare also Krings et al., 2011, 2013).

R#20: P12L26: There is no justification for the removal of data “close to saturation”. As long as there is no saturation, these data should have a high SNR. Please justify

A: We added more information to the revised version:

Filtering of the data accounts for not only SNR but also whether linear full well is achieved. For the full well ADC range chosen by the manufacturer a non-linear behavior could be observed for very high detector fillings. Therefore data with very high filling factors are excluded from further processing. However, out of all measurements, the chosen maximum threshold value affects only 4 single measurements (all in one burst) during the whole measurement period.

R#21: P14L1. It is said that the elevated XCO₂ are well aligned with the wind field from the power plants, but it seems to me that the high value are further North-East that what would be expected

A: Considering the complexity of the atmosphere (turbulence or puffiness of air masses with high CO₂ concentrations) we consider the average alignment of the overall plume structure of all power plant emission with the determined wind direction to be quite good as can also be seen from Figure 13.

R#22: P14L6: It is said that the boundary layer depth is important to compute the wind field. However, the in situ measurements clearly show that the boundary layer is not well mixed.

A: The boundary layer depth is used to determine up to which height the released CO₂ may disperse following the vertical Gaussian plume model depending on, for example, distance to the source and atmospheric stability. It does not imply or assume that released CO₂ is instantly well mixed. See also our answer to R#17. However, as can be clearly seen from the in-situ vertical cross-section in the Fig. 5, the CO₂ increase is indeed reaching up to the highest available in-situ legs. Please keep in mind that the Fig. 4 on the other hand is conceptual and not based on actual data. Figure 4 was replace for the revised version to avoid confusion.

R#23: P15L1 : Figure 11 shows that there was a significant decrease of the wind speed during the time of the in situ measurements. This should affect the intensity of the plume and I am surprised this was not discussed in the in-situ section.

A: It is not completely clear what the reviewer means by “intensity of the plume”. The in-situ method considers concentration and wind speed measured simultaneously, so decreased concentrations with higher wind speed will still yield the same flux.

R#24: P17L4. Are there really any significant difference for the modelled wind speed over the 10 km area?

A: The standard deviation over the measurement area for the model layer shown in Figure 8 is about 5.8%. This is mentioned in the revised version.

R#25: P17L8. Section 5.2.4 is supposed to be a “Result” section. Yet, a large fraction of it describes the method

A: Agreed. Was moved to Section 4.

R#26: P17L16: It is said that the measurements are

A: Unfortunately the reviewer’s comment is not complete here. We have checked the corresponding part of the manuscript and did not identify any obvious issues.

R#27: P18L4. Description of the wind speed estimate. It is said that a Gaussian profile for the concentration is assumed. Yet, the in situ measurement do not show such Gaussian profile. It would be nice to compare the assumed vertical distribution of CO₂ with the in situ measurements.

A: We repeat here, what we answered to comment R#2:
The Gaussian shape of the vertical distribution can only be observed close to the source. Due to the “reflection” of the plume off the surface and off the top of the boundary layer, the CO₂ gets mixed rather rapidly. About 2 km downwind of the source the CO₂ is well mixed according to the Gaussian model, i.e. on average and not necessarily for any snapshot in time. The in-situ measurements, for example, downwind of power plant Niederaußem (Fig. 5) are more than 2 km downwind of the source so that no distinct Gaussian shape in the vertical concentration profile is to be expected.

R#28: Also, the fact that there is little vertical gradient in the wind speed makes this discussion somewhat unnecessary.

A: That is to some degree true. On the other hand, a complete description of the method should involve the estimation of the average wind speed. For the revised version we condensed the discussion.

R#29: How is done the weighting to derive a mean wind speed ?

A: We extended the main text:
The emitted CO₂ was then distributed using a vertical Gaussian dispersion with the stability parameter resulting from the 2D horizontal Gaussian plume inversion model. This

information could be used to obtain an altitude weighted mean wind speed for the remote sensing cross sections through the plume based on relative concentrations per altitude layer.

R#30: P18L18. “Very unstable atmospheric conditions”. Is that consistent with the observed meteorological conditions on that day?

A: The convective dispersion leads to unstable atmospheric conditions. Note also that the derived stability is an effective parameter that also subsumes other effects such as increased flow gas temperature or even changes in wind direction that may lead to additional plume broadening and dispersion. We are more explicit about that in the revised version.

R#31: P23L22: The authors state the error analysis leads to an uncertainty on the order of 10% for the mass balance approach. This is in contradiction, I believe, with the results shown in Figure 15 that show larger variations for the various leg estimates. For instance, three legs can be used to estimate the emissions from Niederaussem. There is a factor of 2 between the largest and the smallest. This appears contradictory with the error analysis, in particular since several of the error sources are biases and cannot explain a difference between the estimates from two nearby legs. I am surprised this is never discusses in the text.

A: The reviewer is right. Our sensitivity study for the remote sensing mass balance approach did indeed not take into account any statistical errors. The magnitude of the flight track to flight track variability shows furthermore how critical it is to have a sufficient number of flight tracks to obtain an accurate estimate. As there are only few flight tracks per power plant, the error is naturally quite large. This is now discussed in the revised version, the error analysis was updated and the updated Figure 15 includes these uncertainties.

R#32: P21L1: The whole section 5.2.5 is poorly written.

A: The section was shortened and improved.

R#33: P24L4: “can differ more than 20% for individual power plants”. So what are you saying here ? Are you suggesting that the reported emissions shown in the paper (Figure 18) can be off by that much ?

A: Not at all. We explicitly wrote:

“The error on power generation itself is generally about 1% (compare also Krings et al., 2011) and the annual error of derived emissions is required to be within 2.5% (European Commission, 2007). **The error for the time of the overflight is most likely not much larger**, although comparisons between U.S. inventories based on monitoring of stack gases with inventories based on emission factors can differ more than 20% for individual power plants (Ackerman and Sundquist, 2008).”

In summary:

- (1) We have no indication to believe that the error is larger than what is required.
- (2) There is a publication that found differences of 20% between different methods.

R#34: In the following, I make comments on the figures. I strongly believe that several of them are not useful whereas other could bring additional information
Figure 1 : Figure 2 : Limited usefulness

A: Figure 1 was updated to contain also the in-situ tracks and the new Figures 2 to 8 are replacing those that were questioned.

R#35: Figure 4: Provide colour scale

A: The old figures 4 and 5 are replaced by the new figures 2 to 8, which should be much clearer now. All color scales are provided.

R#36: Figure 5 : Is this figure supposed to show the same data as in Figure 4 ?

A: See above. No, it was not the same data, and we agree that the old Figure 5 was confusing because the concentrations between the cells are smoothed by the graphics program. We are sure that the new figures are much clearer and more consistent because it is clearly visible now which were the original measurements (Figures 2 & 3), and how they were treated on the grids. This allowed us to omit the old Fig. 4 which was only showing the concept.

R#37: I cannot recognize any feature. I strongly recommend to show the value of the measurements within the circles that are used for the interpolations.

A: See answers to R#35 and R#36.

R#38: What is the link between this figure and the “wall” approach shown in Figure 4 ?

A: See answers to R#35 and R#36.

R#39: Figure 6: Definitely not useful. Not clear what is really shown (ie RMS of what, relative to what ?)

A: This Figure justifies the 0.9% filter on the RMS of (model-measurement)/model. It is quite instructive documenting the good data quality. The Figure was removed anyhow. See also comment to R#19.

R#40: Figure 7: Definitely not useful

A: Was removed.

R#41: Figure 8: Is it really XCH₄ as indicated in the legend, or XCO₂ ? Why no color scale ?

A: Typo was corrected and the color scale added.

R#42: Figure 9: Marginally usefull. The text could simply say that the in situ measurements (potential temperature and aerosol) provide no useful information to determine the top of the boundary layer up to 1100 m

A: The Figure was removed and the text updated accordingly.

R#43: Figure 10 : Should definitely be presented in the “in situ” section, together with Figure 5, and not in the remote sensing section.

A: Agreed. However, the Figure was removed for the revised version.

R#44: Figure 11: Difficult to read. Values for the X-axis could be simpler (e.g. 5/10/15)

A: Scale has been adjusted.

R#45: Figure 12: Should be shown early on in manuscript.

A: We agree. The Figure has been replaced by similar Figures shown at the beginning.

R#46: Why is the color scale not adjusted to the data (no observation before 12)

A: This is to have the same color scale for all associated plots making comparisons between plots easier.

R#47: Figure 13: I suggest to reduce the range of the color bar to 0.99-1.02 and have color lines (Gaussian plume model) for 1.005, 1.01, 1.015 and 1.02

A: Adding more lines will make the plot unreadable. In combination with the complementary Figure 14 which now contains also the model result (see below), the information from Fig. 13 should be sufficiently detailed.

R#48: Figure 14: I strongly suggest to add, on each of these graphs, a line showing the result of the modeling according to the Gaussian plume approach. Also, add a horizontal line to show the 0.

A: The Figure has been updated accordingly.

R#49: Figure 15 : State explicitly in the legend that each symbol corresponds to a flux estimate derived from a given aircraft leg.

A: Done.

R#50: Figure 16 : Not useful

R#51: Figure 17 : Not useful

A: For our analysis we make choices for both wind direction and grid size and considered it reasonable to justify our decision and investigate the impact and sensitivity. Fig. 16 is furthermore important because it shows that wind direction can in principal be fitted directly to the data, which we now explicitly point out in the revised version. We kept Figure 16 but removed Figure 17 as suggested.

References:

Ackerman, K. V. and Sundquist, E. T.: Comparison of Two U.S. Power-Plant Carbon Dioxide Emissions Data Sets, *Environ. Sci. Technol.*, 42, 5688–5693, doi:10.1021/es800221q, 2008.

Gordon, M., Li, S.-M., Staebler, R., Darlington, A., Hayden, K., O'Brien, J., and Wolde, M.: Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations, *Atmospheric Measurement Techniques*, 8, 3745–3765, doi:10.5194/amt-8-3745-2015, <https://www.atmos-meas-tech.net/8/3745/2015/>, 2015.

Krings, T., Gerilowski, K., Buchwitz, M., Reuter, M., Tretner, A., Erzinger, J., Heinze, D., Pflüger, U., Burrows, J. P., and Bovensmann, H.: MAMAP – A new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: retrieval algorithm and first inversions for point source emission rates, *Atmos. Meas. Tech.*, 4, 1735–1758, doi:10.5194/amt-4-1735-2011, 2011.

Krings, T., Gerilowski, K., Buchwitz, M., Hartmann, J., Sachs, T., Erzinger, J., Burrows, J. P., and Bovensmann, H.: Quantification of methane emission rates from coal mine ventilation shafts using airborne remote sensing data, *Atmos. Meas. Tech.*, 6, 151–166, doi:10.5194/amt-6-151-2013, 2013.

Young, S. J.: Detection and quantification of gases in industrial-stack plumes using thermal infrared hyperspectral imaging. Aerospace Report No. ATR-2002(8407)-1. El Segundo, Calif: The Aerospace Corporation, 2002.