September 30, 2017

Dear Editor,

We have now uploaded our comments to the points raised by the two referees and in the additional comment by L. Golston. We carefully addressed each individual comment.

These comments were in fact very helpful to revise our manuscript and as a consequence, we in particular

- substantially improved the structure of the manuscript following the guidelines given by the two referees,
- substantially improved and added additional analyses for the in-situ emission rate estimation method as requested by the referees and in the additional comment by L. Golston,
- we refined the remote sensing section with an additional analysis of the emissions for the power plant furthest downwind ("Frimmersdorf") and included an additional error component for the remote sensing mass balance approach as suggested by referee #1,
- addressed all of the referee comments which were quite detailed and substantial, and which indeed lead to a significantly improved manuscript as we believe.

The present document includes all our comments to the issues raised by the referees and in the additional comments. At the end of this document a marked up manuscript version can be found highlighting all modifications made for the revised version.

Best regards,

Thomas Krings on behalf of the authors

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 CO_2 gets mixed rather rapidly. About 2 km downwind of the source the CO_2 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 zres (vertical) x hres (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 CO_2 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 CO_2. 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: RMS[(model-measurement)/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 CO_2 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 CO_2 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 CO_2 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 CO_2 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 CO_2 gets mixed rather rapidly. About 2 km downwind of the source the CO_2 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 CO_2 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 flew 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: Marginaly 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.

Reviewer 2

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:The manuscript "Airborne remote sensing and in-situ measurements of atmospheric CO2 to quantify point source emissions" by Krings et al presents results from an airborne campaign, inferring point source fluxes of CO2 using both mass balance approaches as well as a Gaussian plume modeling of remotely sensed total column averages.
Even though the data presented here is indeed interesting, I tend to agree with reviewer 1 that it often reads much like a report and would need restructuring and more concise (and precise!) language.

A: We improved on the overall structure to meet the guidelines given by the two reviewers.

R#2: At times, the authors get caught up in details that are not entirely relevant to the study at hand, e.g. Figs 9-12 are too detailed or not necessary (9-10) or misplaced in the respective section

A: We removed Figures 6, 7, 9, 10, 17 and moved Figure 12 to the description of the target area, respectively. 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#3: I would suggest putting a general description of the domain as well as the data right at the beginning (e.g. showing MAMAP footprints as well as in-situ ground projections on the map in Figure 1. It would greatly help setting the stage for the discussion.

A: Figure 1 has been updated accordingly.

R#4: Some more specific comments: Abstract last sentence: this is a sudden topic break and needs some rephrasing

A: We rephrased the abstract:
Reliable techniques to infer greenhouse gas emission rates from localised sources require accurate measurement and inversion approaches. In this study airborne remote sensing observations of CO_2 by the MAMAP instrument and airborne in-situ measurements are used to infer emission estimates of carbon dioxide released from a cluster of coal fired power plants. The study area is complex due to sources being located in close proximity and overlapping associated carbon dioxide plumes. For the analysis of in-situ data, a mass balance approach is described and applied. Whereas for the remote sensing observations an inverse Gaussian plume model is used in addition to a mass balance technique. A comparison between methods shows that results for all methods agree within 10% or better for cases where in-situ measurements were made for the complete vertical plume extent. The computed emissions for individual power plants are in agreement with results derived from emission factors and energy production data for the time of the overflight.

R#5: Page 3, line 30: wheras (there are many small things like this or "straight forward", which is one word. I won't go into more details, the copy-editor should catch those at a later stage but some sentences are too literal translations from german.

A: We carefully went through the manuscript to improve wording.

R#6: Section 4: I think this is poorly described and justified and I urge the authors to consider revisiting the differences between their approach and the paper Levi Golston mentioned.

A: This was done very extensively.

R#7: E.g.
I) you mention "Kriging" is not necessarily the best suitable approach. If you provide critic, you have to back it up either with an analysis or a citation. There is also no real explanation what kind of interpolation schemes you are using (apart from the boundary voxels).

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. Our method is now clearly described, and displayed in the new figures 4 to 8. 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 average and interpolate 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: The presented method is not better than Kriging, but, Kriging should not be regarded as the only option. This is also true after studying Gordon et al. (2015) in detail. 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#8: II) You mention that you include turbulent fluxes. This is very interesting and I was excited but then I didn't see any further analysis. Did you compute the differences with or without turbulent components? What is the relative error in your case? Can you plot c-mean(c) vs v-mean(v) for some voxels to show the correlations as expected for turbulent fluxes?

A: Yes, this is done now. The different results can be found in a detailed Table as a supplement since it would be too much for Table 3 in the publication. The differences with and without turbulent fluxes were very small, and – surprisingly – negative, i.e. the turbulent contributions are rather from dilution (entrainment) than adding to the flux.

R#9: III) Wind speed seems to be a dominant error, do the others actually matter? You will need to provide realistic estimates regarding kriging and turbulent fluxes, otherwise the reader won't be able to judge the importance (even though this specific case study might not lend itself to extrapolation to a general case). Page 5, line 9: As above, please show what error you incur by doing mean(v)*mean(c) vs. mean(v*c)

A: This is now explicitly done by providing our standard fluxes (averages plus inter- and extrapolations of local mass x wind) plus 'flux 2', which was calculated after the averaging of the mass- and wind-field by our method, and by Kriging.

R#10: Figure 4: Please add color-bar and make this a realistic example based on real data.

A: The confusion about the old Figures 4 and 5 should now be eliminated. The new figures are much clearer and more consistent because it is clearly visible now which were the original measurements (Figs. 2 and 3), and how they were treated on the grids. This allowed to omit Fig. 4 which only showed the concept.

R#11: How many data points to average do you typically have per voxel?

A: A grid cell which was crossed once or twice is averaging the 5 Hz data (concentrations and wind), resulting in typically 10 to 20 data points when hres and zres were 100 m. However, the size of the cells was varied between 100 and 200 m for the sensitivity analysis, resulting in 10 to 40 points.

R#12: Page 9, line 8: Conditio sine qua non: Even though I have a "Grosses Latinum", I had to look it up again. Please rephrase in plain english, esp. as it is here used in a rather trivial way, not warranting the grandiose latin phrase ;-).

A: Was revised.

R#13: One might argue though that precision "could" be important if it is really bad while accuracy won't matter. This could be a factor when flying very cheap instruments on small unmanned aerial vehicles near the plume. So I would keep the discussions as general as possible.

A: We agree in principle. However, the absolute accuracy of the concentrations remains less important when subtracting the background based on the same measurements. This is also true for expensive and relatively stable instruments.

R#14: Page 9, line 17: Chimneys: It would be good to discuss how well you could measure the fluxes if the emissions were to happen at the surface. What would this imply for the in-situ based approach and potential flight-paths.

A: We discussed this in the referenced ESA report for CH_4, which was emitted from coal shafts close to the surface. Of course in such cases it is more important to have the lowest flight track as low as possible (50 m above ground with a special permission). However, in this campaign, where we flew to many sources for which we had no low-flying permission, this was restricted to 150 m. This was irrelevant for the CO_2 from high chimneys though. Another factor is the stability of the boundary layer. Since we have chosen daytimes for flying, when the convection reached at least the 150 mAGL, extrapolations with constant flux for the CH_4 (not in this paper), and diminishing or constant concentrations for the CO_2 (both options for checking the sensitivity) were applied. The constant flux for the CH_4 was applied because an expected enhancement of concentrations near the surface was compensated by the diminishing wind speed. For more reliable results for CH_4 we would prefer mobile measurements near the surface, which some of the authors did during a separate campaign in summer 2016. This aspect is now mentioned in the revised text.

R#15: Figure 5: This figure confuses me. I "assume" the dots are actual measurement locations.

A: See **R#10**.

R#16: Given what you wrote, there is a constant extrapolation to the surface. However, it doesn’t look that way for the second little intrusion at x=0. Also, there are a couple of local maxima in between dots. You need to explain the interpolation scheme and this would a good place to compare against kriging or other interpolation schemes. Also, if the dots are measurements, please color-code them with the actual measurement values at that x-y position. This will help evaluate the interpolated fields better. A last point: Why is this continuous on x and y? Wouldn’t it make sense to sketch out the actual grid boxes here as well?

A: Fully done. Figure 5 was confusing and misleading by several reasons: (i) Showing concentrations instead of the locally measured fluxes is potentially misleading; (ii) the old figure added interpolation & smoothing from the graphics package. The new style of cross sections in Figures 4 through 8 is much clearer.

R#17: Page 12, line 3: Please add citation for proxy method (this is not common knowledge).

A: We added the references to the revised manuscript:
See Frankenberg et al. (2005) and Schepers et al. (2012) for more information on the proxy method and Krings et al. (2011) for its application on MAMAP measurements.

R#18: Figure 6: I think the figure itself doesn’t tell more than the text, could be skipped.

A: Figure 6 has been removed.

R#19: Same with Fig. 7

A: Figure 7 has been removed.

R#20: Page 14, line 17: So in essence, you don’t really need the wind speeds in this case as the error is rather small?! Ideally, you won’t always need both aircraft. If there is confidence in modeled winds, it would be a good sign for future remote-sensing only campaigns.

A: We agree with the reviewer. Not having to use additionally measured in-situ wind speed would indeed be a huge step forward. A wider and systematic analysis on the accuracy of the model wind would indeed be very interesting but can of course not be accomplished within the present study. We added a short discussion about this in the manuscript:

The analysis in this study shows an overestimated model wind speed of about 6% (or about 0.4 m/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 (0.7 m/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.*

R#21: Figure 10: Weird x-spacing (value 493?). also better to use same x-scale for both subplots.

A: We removed the Figure completely as suggested in R#2.

R#22: Page 21, line 10 +/-: Wouldn't you ideally fit a Gaussian model with a vertical windspeed profile? This would rather directly model the total column AND the wind-profile.

A: Currently our model is set up to work with an average wind speed. The direct utilisation of the vertical wind profile u(z) for the inversion would be an interesting experiment. This would basically mean fitting the measurements to the sum of a vertically piecewise (or even continuously) changing Gaussian plume model (ignoring second order lateral and temporal wind speed variations in a first step). However, as Reviewer 1 pointed out that the discussion is "somewhat unnecessary", since in this case the vertical gradient in wind speed is not very strong we do not want to extend the discussion on this here and leave that to future work on the topic.

R#23: How high is your Gaussian profile extending to the vertical? That might be a plot to add (or is it just 2D in x and y?).

A: The Gaussian model to determine the average wind speed is 2D in along wind and vertical directions (x and z). With the same reason as above we do not want to add an additional plot for this but briefly discuss the vertical distribution as a function of downwind distance:
At the first remote sensing leg 700m downwind of power plant Niederaußem, the plume reaches about 1 km height, and at 2 km downwind distance the CO2 is already well mixed according to the plume model which represents a temporal average.

R#24: Page 23, line 19: "This is because they to a good extent cancel out..." "largely" cancel out?

A: Done.

R#25: Page 26, line 20: I think they gase don't need to be inert, just have lifetimes much longer than the time between emission and measurement. I would guess even NOx emissions could work with an "inert" assumptions on this very local scale.

A: Agreed. Has been clarified in the text of the revised version.

R#26: As a last general point: Please try to re-structure somewhat to bring out the key messages in a more concise way (and illustrate better how your in-situ inversions differ from others).

A: We agree and improved on the description as indicated in our answers above and to Reviewer 1.

R#27: At the end, provide a more generic overview of both flux estimates and

its pro/cons and path forward. This could extend to a discussion using high-resolution mapping like the cited AVIRIS-NG papers (which should be cited at page 26, line 9). Last but not least my sincere apologies for the late review.

A: Done. The discussion will be along the lines of what we also wrote as answer to Reviewer 1:

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. 11. 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.

References:

Frankenberg, C., Meirink, J. F., van Weele, M., Platt, U., and Wagner, T.: Assessing Methane Emissions from Global Space-Borne Observations, Science, 308, 1010–1014, doi:10.1126/science.1106644, 2005.

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.

Schepers, D., Guerlet, S., and Butz, A.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10 307, doi:10.1029/2012JD017549, 2012.

Comment by L. Golston

We would like to thank L. Golston for the additional comments. In the following we carefully address the comments point by point. First we repeat the comment (C#...) and then give our response (A).

C#1: The authors describe two methods for quantifying point source emissions, one based on in-situ mass balancing and one on remote sensing coupled with an inverse Gaussian plume model, along with field data comparing the two methods. I would like to add three comments on the description of the in-situ method, which I hope will help strengthen that section.
It was surprising that Gordon et al. 2015 (AMT 8:3745-3765) and Cambaliza et al. 2014 (ACP 14:9029-9050) are not discussed or even cited, given that they both investigate in detail the uncertainties of the in-situ aircraft mass balance methodology.
Gordon et al. 2015 discusses issues of interpolation, extrapolation, turbulent fluxes, and issues related to determining the background concentration in the context of determining emissions from an elevated source as was done here, while Cambalizia et al. 2014 also considers interpolation, boundary layer entrainment, and other effects. On page 4 transparency of interpolations and extrapolations is claimed as a benefit here, but seem less well developed than in either of those papers. The description of both on Page 8, is actually specifically not transparent and neither seems to be included in the error budget of Table 4. Part of the reason why the in-situ method is discussed in such detail seems to be because a variation on the mass balance method is presented, however the differences and benefits are not clearly distinguished in Section 4.1 or the results.

A: Two papers where Cambaliza was a co-author were referenced. However, the additional references from Cambaliza et al. (2014), and Gordon et al. (2015) were studied now in detail and were helpful. In the revised manuscript, we discuss especially the link to points that Gordon et al. have mentioned.

C#2: I also recalled that Figs 3, 4, and part of 2 are identical to Figure 3 in Hacker et al. 2016 (Animal Production Science 56:190-203), who cited the report from the authors of the current paper, Bovensmann et al. (2014). Since the figures are now also in Hacker et al., I think that the original Bovensmann et al. (2014) should be referenced here to avoid confusion.

A: The new Figures 2 and 3 for the measurements and 4 to 8 for the gridding are much clearer now.

C#3: Finally, it would be helpful to know whether the turbulent (5 Hz) could be resolved as indicated in the caption for Figure 3 or if there was attenuation, and how the inclusion of the turbulent flux compares to not including it.

A: This is now explicitly done by providing our standard fluxes (averages plus inter- and extrapolations of local mass x wind) plus 'flux 2', which was calculated after the averaging of the mass- and wind-field by our method, and by Kriging.
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.

Airborne remote sensing and in-situ measurements of atmospheric CO_2 to quantify point source emissions

Thomas Krings[1], Bruno Neininger[2,3], Konstantin Gerilowski[1], Sven Krautwurst[1], Michael Buchwitz[1], John P. Burrows[1], Carsten Lindemann[4], Thomas Ruhtz[4], Dirk Schüttemeyer[5], and Heinrich Bovensmann[1]

[1]Institute of Environmental Physics, University of Bremen, FB 1, P.O. Box 330440, D-28334 Bremen, Germany
[2]METAIR AG, Airfield Hausen am Albis, CH-8915 Hausen am Albis, Switzerland
[3]Zurich University of Applied Sciences, CH-8400 Winterthur, Switzerland
[4]Institute for Space Sciences, Free University of Berlin, Carl-Heinrich-Becker-Weg 6-10, D-12165 Berlin, Germany
[5]ESA / ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands

Correspondence to: Thomas Krings (Thomas.Krings@iup.physik.uni-bremen.de)

Abstract. Reliable techniques to infer greenhouse gas emission rates from localised sources require accurate measurement and inversion approaches. In this study airborne remote sensing observations of CO_2 by the MAMAP instrument and airborne in-situ measurements are used to infer emission estimates of carbon dioxide released from a cluster of coal fired power plants. The study area is complex due to sources being located in close proximity and overlapping associated carbon dioxide plumes. For the analysis of in-situ data, a mass balance approach is described and applied. Whereas for the remote sensing observations an inverse Gaussian plume model is used in addition to a mass balance technique. A comparison between methods shows that results for all methods agree within ~~a few percent~~ 10% or better for cases where in-situ measurements were made for the complete vertical plume extent. ~~Even though the power plants are partly in close proximity and the associated carbon dioxide plumes are overlapping it is possible to derive emission rates from remote sensing data~~ The computed emissions for individual power plants ~~that agree well~~ are in agreement with results derived from emission factors and energy production data for the time of the overflight.

1 Introduction

Knowledge of emissions of the greenhouse ~~gases~~ gas carbon dioxide (CO_2) ~~and methane (CH_4)~~ originating from localised sources is often inadequate (Ciais et al., 2015; NRC, 2010).

Even for well monitored localised CO_2 emitters, there are significant differences between inventories calculated using different but plausible methods. For example, Ackerman and Sundquist (2008) found differences of more than 20% between emissions from U.S. power plants. Differences in inventories of about a factor of two were found for CO_2 emissions from flaring in the oil and gas production (Ciais et al., 2015)~~,~~.

~~Similarly, the magnitude of fugitive emissions of natural gas, which comprises predominantly , is not clear and results from the lack of measurements . Recent studies using different methods show that there are significant disagreements between process based bottom up inventories on the one hand and top down emission estimates based on atmospheric measurements on~~

~~the other hand . This is particularly relevant because losses of from the use of natural gas from traditional sources or fracking negate the benefits in reducing global warming when using natural gas as an energy source in the transition from coal based to alternative renewable energy sources . These benefits result from its lower specific carbon dioxide emissions per Joule of energy compared to burning coal~~.

Top down estimates of localised sources are generally obtained using airborne or ground based in-situ measurements. Recently also the use of airborne remote sensing has demonstrated the ability to accurately estimate emissions. All methods have their distinctive advantages and disadvantages. Ground based in-situ measurements are fairly low-cost. However, they do generally not sample the complete atmospheric boundary layer, which is necessary for an accurate emission estimate. Airborne in-situ allows accurate concentration and wind speed measurements from which emissions can be derived ~~, but optimally require a vertically~~ (e.g. Karion et al., 2013; Cambaliza et al., 2014; Caulton et al., 2014; Gordon et al., 2015; Lavoie et al., 2015). However, they require a dense flight pattern~~. In addition assumptionsabout the lowest atmospheric layer~~ , and assumptions, for example, about the layer below the lowest flight track have to be made~~that can usually not be accessed by aircraft. Furthermore~~. Furthermore, air space restrictions can interfere with required flight patterns. ~~Remote sensing~~ Airborne remote sensing of atmospheric column concentrations of CO_2 allows sounding of the complete boundary layer and offers the opportunity to survey large areas in short time spans (Krings et al., 2011, 2013; Thompson et al., 2015; Frankenberg et al., 2016; Thorpe et al., 2014; Tratt et al., 2014). In contrast to in-situ measurements, however, clear sky conditions are mostly required as they measure backscattered solar electromagnetic radiation, except for those instruments operating in the thermal infrared (e.g. Tratt et al., 2014). ~~Furthermore, wind~~ However, interpretation of thermal infrared measurements depends on the thermal contrast (e.g. Young, 2002). Wind information has to be additionally gathered for example from model data (see, for example, Krings et al., 2011) or additional in-situ wind measurements (see, for example, Krings et al., 2013).

Using the example of CO_2 emissions from a cluster of power plants in western Germany, top down results from airborne in-situ and remote sensing are evaluated and compared to each other, as well as to an independent bottom up estimate computed from energy production and emission factors. Additional complexity is added to the top down inverse problem since sources are partly located in close proximity to each other with overlapping CO_2 plumes. Generally, it is necessary to achieve a correct source attribution in the presence of multiple neighbouring sources for validation purposes. ~~This provides insights into origin and specific processes that lead to the generation of greenhouse gases.~~

2 Measurement campaign and ~~instrumentation~~target area

An airborne measurement campaign combining remote sensing measurements of column-averaged concentrations of CO_2 and CH_4, denoted XCO_2 and XCH_4 with in-situ measurements of atmospheric CO_2 and CH_4 concentration (mass per volume) as well as wind speed and direction in the boundary layer took place in Germany in August 2012. This campaign was carried out in the framework of ESA's Earth Explorer 8 activities for the candidate mission CarbonSat (Bovensmann et al., 2010; Buchwitz et al., 2013). The CarbonSat concept founded the basis for the CO_2 monitoring mission now under preparation within the European Copernicus program.

For the remote sensing of the greenhouse gases CO_2 and CH_4, MAMAP (Gerilowski et al., 2011) was flown above the boundary layer on the Cessna T207A aircraft of the Free University of Berlin. MAMAP is an airborne 2 channel NIR-SWIR grating spectrometer system for accurate measurements of gradients in column-averaged methane and carbon dioxide concentrations. It was jointly developed by the Institute of Environmental Physics / Remote Sensing (IUP/IFE), University of Bremen (Germany) and the Helmholtz Centre Potsdam, German Research Centre for Geosciences (GFZ). It was demonstrated that the instrument is able to detect and retrieve the total dry column of the greenhouse gases CH_4 and CO_2 with a precision of 0.3–0.4% ($1\,\sigma$) at local scales ($\approx 10\,\mathrm{km}$), and that MAMAP is an appropriate tool for detection and inversion of localised greenhouse gas emissions from aircraft (Gerilowski et al., 2011; Krings et al., 2011, 2013; Krautwurst et al., 2017).

For the in-situ airborne measurements, the small research aircraft ~~METAIR TTC-ECO-DIMO~~ METAIR-DIMO was flown in the boundary layer to perform measurements in the ~~plume~~ plumes emitted from the ~~target, to perform~~ targets, to gather background concentration measurements and to perform wind measurements needed for the interpretation of the total column measurements of MAMAP. The aircraft is equipped with underwing-pods, carrying up to 50 kg scientific payload each. The standard equipment measures the meteorological parameters wind (three-dimensional components in ~~turbulent~~ 10 Hz resolution), fast temperature and fast, redundant humidity. A two-channel aerosol counter (MetOne for >0.3 µm and >0.5 µm) can characterise the structure of the boundary layer. The chemical measurements are for CO_2 (redundant) and CH_4 corrected for H_2O interference (dilution and spectroscopic; ~~details~~ for details, see Hiller et al. (2014)) as well as CO, O_3, NO_2, NO_x, NO_y, O_x. The ~~methane monitor is a "Los Gatos DLT-100 Fast Methane Analyser" which was purchased by ETH Zurich and modified in a joint project. The~~ CO_2 is measured with three different time resolutions and accuracies, resulting in an overall accuracy and precision of 0.4 ppm. The individual contributions are: (i) a fast (10 Hz) measurement with a short-term precision (e.g. while crossing a plume) of about 0.2 ppm, with a limited absolute accuracy of about 5 ppm, using a modified LiCor LI-7500. (ii) A more accurate continuous, but slower (0.3 Hz) reference with a modified LI-6262, with an accuracy of better than 0.5 ppm. The highest accuracy is from flask-samples~~(typically 9 per flight)~~, analysed at MPI Jena. This method is described in Hiller et al. (2014).

The surveying strategy was to simultaneously probe the atmospheric boundary layer with in-situ ~~measurements while~~ and remote sensing measurements where the MAMAP remote sensing measurements were performed via the ~~separate aircraft above~~ second aircraft above the boundary layer.

3 ~~Target area~~

While the complete campaign covered also other CO_2 and CH_4 emitting targets (Bovensmann et al., 2014) this study focuses on measurements obtained in an area with several lignite fired power plants in western Germany ~~close to the city of Dsseldorf~~ (see Fig. 1) on 15 August 2012. The power plants are Niederaußem, Neurath (old and new blocks) and Frimmersdorf. The remote sensing flights were performed at about 11:50 – 13:40 UTC (that is 13:50 – 15:40 local time, CEST). The in-situ survey over the same area was conducted between about 12:15 and 14:20 UTC. Wind was blowing ~~approximately~~ from South-East (145°- 148°) so that the CO_2 plumes of individual power plants were overlapping. ~~Variation of the surface elevation in the~~

~~immediate vicinity of the power plants was rather low in particular for the area in between power plants. The location of the open cast mines, visible as surface depressions, have shifted since the topography measurements shown in Fig.1 as can be seen from more recent imagery available, for example, via GoogleEarth. The open cast mine west of Frimmersdorf has been moved further to the West, wheras that south of Niederaußem has been closed and refilled~~

The in-situ measurements concentrate on transects at several altitude layers around the two Neurath power plants and around the extended area including Frimmersdorf power plant.

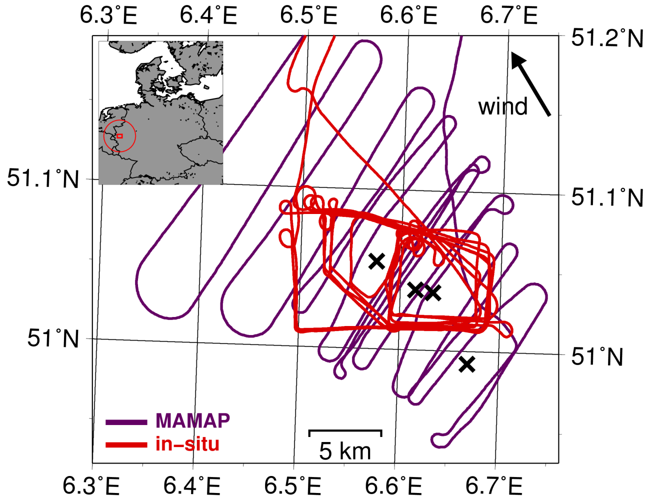

Figure 1. Map of the target area in western Germany. The crosses denote the four lignite fired power plants in the area. ~~Topographic data have been obtained~~ From upwind to downwind, i.e. from South-East to North-West, the ~~Shuttle Radar Topography Mission~~ power plants are Niederaußem, Neurath (~~SRTM~~new)~~version 2.1~~, Neurath (old) ~~, a collaborative effort from NASA, NGA as well as the German~~ and ~~Italian Space Agencies~~Frimmersdorf.

3 Methodology

3.1 Flux estimates from in-situ measurements

Calculating fluxes of trace gases through an imaginary vertical plane is ~~trivial~~ simple when the concentration field and the wind field are known for a sufficient time during quasi stationary conditions. However, in reality, such perfect measurements are not possible. Not the accuracy or precision of the measurements are a prime concern, but, unknown parts in the fields (inter- and extrapolations), and~~ - most important -~~, most important, remaining instationarities both by short-term fluctuations (hitting a part of a plume or not), and by varying source strengths and changing meteorological conditions. Cambaliza et al. (2014), Gordon et al. (2015), Lavoie et al. (2015) and Caulton et al. (2014) ~~discuss~~ are discussing comparable airborne in-situ observations downwind of regions and individual sources ~~. Their basic methods are~~ on different scales. The basic method to derive emissions form atmospheric CO_2 or CH_4 measurements is described in Mays et al. (2009). The instrumentation and methods that were used in the work presented here are quite similar~~. However, by comparing the two approaches that were developed independently, three main differences can be identified~~, with the following characteristics with reference to Gordon et al. (2015): (i) ~~The methodof inter- and extrapolation of concentrations and winds are more transparent and straight forward here, where "Kriging" as applied in the referenced work is not necessarily best suitable for this type of data set. (ii) Here, also turbulent fluxes are included, both in the horizontal, and in the vertical. (iii) Due to a slower aircraft with an optimized wind measuring system, the accuracy of the wind field is slightly better. According to~~ A 'single screen' approach was chosen (as opposed to a box method); (ii) due to the small distance to the sources, terms like chemical conversion or storage in the volume are irrelevant; (iii) the CO_2 emissions were from hot stacks at short distance which reduces the problem of extrapolation to the surface; (iv) the plume concentrations were large (see Fig. 2); (v) the time resolution of the instrumentation and the relatively low speed of the aircraft (typically 40 $\mathrm{m\,s^{-1}}$) allowed to include the turbulent horizontal fluxes (at 5 Hz for wind, concentrations and density, the spacial resolution was better than 10 m).

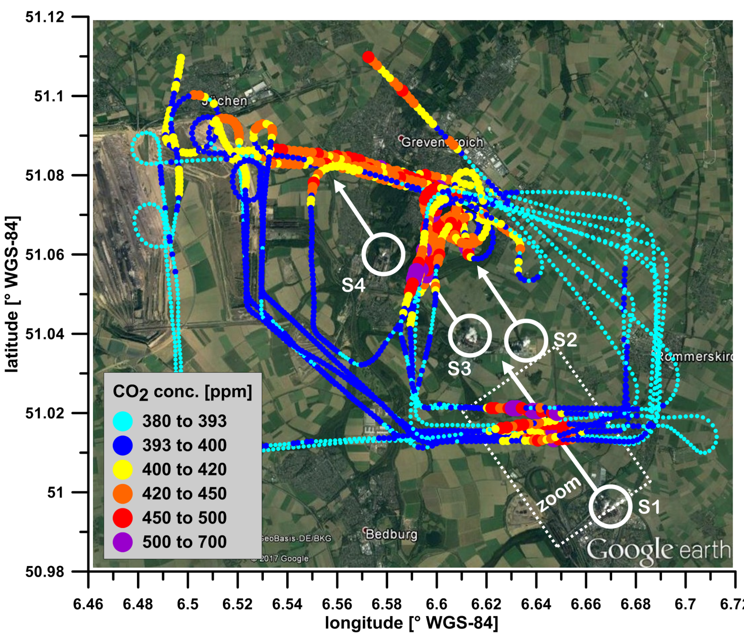

Figure 2. The in-situ flight pattern around and between the four sources (S1: Niederaußem; S2: Neurath (new); S3: Neurath (old); S4: Frimmersdorf), with the wind (white darts, from 146°). The CO_2 concentrations (1 dot per 3-s-average, i.e. about every 120 m) are color coded. In the North, the flight pattern was along an air space that could not be entered. Also on top, the cross sections were limited to 3500 ft AMSL (1050 m AMSL).

The single screen approach was chosen for practical reasons, because flying around a source means to spend most of the time in background concentrations, having a higher risk to miss quasi-stationary conditions. Some circumferential tracks (Fig. 2) confirmed the background concentrations found on the edges of the single screens. In contrast to these benefits we suffered from a non-ideal flight pattern caused by two reasons: (a) The air space restrictions both horizontally and vertically were complex, and (b) this was the first day in this yet unknown region, i.e. we were still in the process of optimizing, preparing for other days with other wind directions (south to southwest). However, this was one of only a few days of the campaign, where the remote sensing method could be directly compared with in-situ, because on many other days the cloud conditions were less ideal for remote sensing.

As in other work mentioned above, the measured ~~wind vectors were of limited accuracy. It was decided to use modeled wind fields instead. In later work of this group , the wind measurements might have been improved~~data had to be inter- and extrapolated to a grid covering the whole cross sections. Because the data density on most cross sections was relatively high (see Fig. 3), and the plumes were not at all regularly shaped (rather an ensemble of puffs than a Gaussian plume), a method for the inter- and extrapolation was developed that is mainly using the grid cells of the cross sections in which measurements can be found. This structure can be seen in Fig. 4.

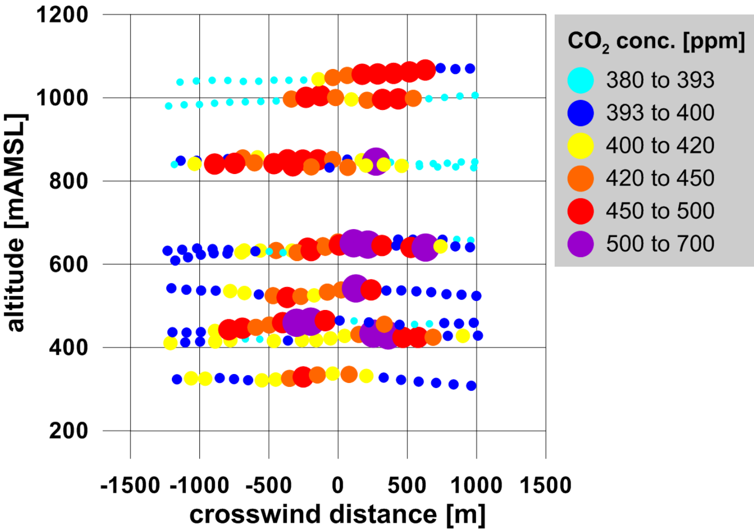

Figure 3. Details of the cross section through the plume of Niederaußem (S1). The projected CO_2 concentrations (1 dot per 3-s-average, i.e. about every 120 m) are color coded. It is important to average the fluxes along the wind, i.e. all values have to be projected to an imaginary cross section perpendicular to the wind, and not parallel to the flight pattern (see Fig. 2 and main text). The background concentration was determined to be at about 392 to 393 ppm, which is a small uncertainty compared to the maximum plume concentrations of about +200 ppm).

Before describing the interpolation process, the main difference in the method compared with other approaches is emphasised. In all the work referenced above, the fluxes were calculated by multiplying an interpolated wind field with an interpolated field of the mass distribution. This was done as well ('flux 2', as shown in Fig. 5). However, ~~the two methods and the discussions about remaining uncertainties are very similar. Referring to the importance of turbulent fluxes, gives a concise overview~~

~~about the difficulties of complete closures of fluxes. Measuring fluxes means to be aware of contributions that might not be accurately captured, and to minimize them by suitable methods and measuring strategies, which includes the choice of suitable meteorological conditions. In some convective situations, also the vertical flux above the source has to be considered. When dealing with distinct sources in a limited area, where deposition, storage and other terms can be neglected, we have the following situation:~~ when the pointwise measurements of wind and concentration (or mass, when including density as a function of locally measured pressure, temperature and humidity) are known, the mass flux can be calculated locally, for each point of the measurements. With the mentioned common sampling rate of 5 Hz (original data 10 Hz or faster; all sensors good enough for 5 Hz), this means, that every 10 m, there is a measurement of CO_2 flux perpendicular to the chosen cross section, including the turbulent contribution up to this frequency or length scale. When averaging several of such measurements, this results in a direct average mass flux for a chosen grid cell in the vicinity of the flight track. The most extreme coarse approach is to average over the whole flight near the cross section. These results are listed as 'bulk average' in the supplementary material. It is astonishing, that these results, which make no assumptions about the shape of the plumes whatsoever, and do not need any method for interpolation, are mostly near or even within the range of the more sophisticated results based on reconstructed fields. Even more extreme is 'bulk average 2', where the averages of all the mass measurements (above background) were multiplied with the averaged winds perpendicular to the whole cross sections. This finding is to some degree relativizing the discussion about subtleties in interpolation methods. Nevertheless, in the following, we explain our linear inter- and extrapolation method, and compare it with Kriging.

~~Since the wind field is not homogeneous in the vertical, the conceptual model described in Fig.~~

The cells shown in Figs. ~~**??** can either be applied on different layers~~ 4 and 7 with bold, larger numbers for the fluxes are resulting from several measurements in these cells. The inter- and extrapolation was done following three rules: (i) single gaps in columns or layers were filled by linear interpolation (first horizontal, then vertical, and again horizontal); (ii) the lowest layer with measurements was completed by taking the value of the second lowest cell above if there were valid measurements in it; (iii) larger gaps were filled differently for two types of parameters: (a) wind and air mass (density) was filled with the averages in the layer; (b) concentrations, ~~or the "wind vectors" and~~ CO_2 masses and fluxes (all above background) were filled with the background. These rules are conserving the values in the cells with measurements and tend to underestimate the fluxes through cells without measurements, resulting in conservative estimates. These rules apply for cells in the range between the highest and the lowest level with measurements.

In all cases, the ~~"concentrations" are already multiplied and averaged in the vertical. The second interpretation is more adequate, because the first one raises the question of what happens if a parcel of air changes its altitude, and hence the layer. The general statement is~~lowest flight track was about 150 mAGL (legal limit without special permission), and the highest track was limited by the air space, which was still in the plume on this day. This means, that the ~~inflow and outflow of background concentrations is balanced (sum = zero~~layers between 1000 mAMSL (or 1100 mAMSL for the plume Niederaußem) and the top below the stable layer at 1300 mAMSL (centered altitude of the cell), and ~~only the additional fluxes added to the background by known (and unknown) sources in the "box" are of interest.~~

z [mAMSL]	Niederaussem: CO_2 flux [kg/s] above background in the grid cells (sums and linear interpolations & extrapolations)																					
1300	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
1200	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.2*	*2.4*	*1.9*	*2.7*	*2.3*	*1.9*	*3.4*	*3.2*	*5.7*	*7.1*	*6.6*	*0.6*	*0.1*	*0.1*	*0.1*
1100	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.5*	*4.7*	*3.8*	*5.3*	**4.6**	**3.9**	**6.8**	**6.5**	**11.4**	**14.2**	**13.1**	**1.2**	**0.3**	**0.3**	**0.1**
1000	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.5**	**4.7**	**3.8**	**5.3**	**3.9**	**0.5**	**4.0**	**11.3**	**9.7**	**2.9**	**2.4**	**0.1**	**0.0**	**0.0**	**0.0**
900	*0.3*	*0.1*	*2.0*	**0.3**	**0.3**	**5.0**	**3.6**	**1.3**	**6.3**	**11.2**	**15.9**	**15.4**	**5.1**	**0.1**	**0.0**	*7.8*	*5.8*	*1.5*	*1.2*	*0.1*	*0.0*	*0.0*	*0.0*
800	**0.6**	**0.2**	**4.1**	**10.2**	**9.4**	**7.2**	**5.0**	**4.6**	**19.2**	**5.6**	**7.7**	**2.0**	**1.2**	**8.1**	**9.8**	**4.3**	**1.8**	**0.1**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
700	*0.5*	*0.4*	*2.4*	*5.4*	*4.9*	*4.2*	*3.6*	*3.3*	*10.6*	*5.5*	*9.5*	*0.3*	**6.4**	**16.1**	**7.7**	**0.2**	**0.2**	**0.2**	**0.1**	**0.2**	**0.2**	**0.1**	**0.1**
600	**0.4**	**0.6**	**0.8**	**0.5**	**0.3**	**1.2**	**2.2**	**2.0**	**2.1**	**5.4**	**11.3**	**5.2**	**8.9**	**23.7**	**14.0**	**7.7**	**1.3**	**7.0**	**26.5**	**3.0**	**1.1**	**0.1**	**0.2**
500	**0.5**	**0.2**	**0.5**	**1.0**	**0.9**	**1.8**	**0.9**	**2.6**	**8.4**	**9.7**	**8.2**	**7.9**	**1.4**	**7.7**	**5.8**	**2.8**	**0.9**	**0.3**	**0.3**	**0.3**	**0.4**	**0.4**	**0.4**
400	**1.0**	**0.7**	**0.6**	**1.0**	**5.8**	**4.8**	**2.5**	**0.8**	**1.3**	**1.0**	**1.9**	**2.6**	**1.7**	**3.0**	**18.6**	**31.3**	**18.3**	**16.5**	**11.2**	**6.2**	**0.8**	**1.5**	**0.3**
300	**0.5**	**1.1**	**1.3**	**1.0**	**0.3**	**0.4**	**1.3**	**1.0**	**3.0**	**7.2**	**12.5**	*2.6*	*1.7*	*3.0*	**1.6**	**0.7**	**0.3**	**0.4**	**0.4**	**0.4**	**0.3**	**0.3**	**0.3**
200	*0.4*	*0.7*	*0.9*	*0.7*	*0.2*	*0.3*	*0.9*	*0.7*	*2.0*	*4.8*	*8.3*	*1.7*	*1.2*	*2.0*	*1.0*	*0.5*	*0.2*	*0.3*	*0.2*	*0.3*	*0.2*	*0.2*	*0.2*
100	*0.1*	*0.2*	*0.2*	*0.1*	*0.0*	*0.0*	*0.1*	*0.1*	*0.3*	*0.7*	*1.4*	*0.3*	*0.2*	*0.4*	*0.2*	*0.1*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
0	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
x [m] >	-1200	-1100	-1000	-900	-800	-700	-600	-500	-400	-300	-200	-100	0	100	200	300	400	500	600	700	800	900	1000

Figure 4. Example of results for the cells on a cross section. Shown is the direct CO_2 flux in the plume of Niederaußem (S1, same plume as in Fig. 3). Since the grid cells are not points, but areas filling $100 \times 100\,\mathrm{m}^2$ (or, more precise, voxels of $100 \times 100 \times 1\,\mathrm{m}^3$), the values are not displayed as contour plots, but, as pixels. The x- and z-coordinates are centered in the cells. The standard inter- and extrapolation was done with the methods described in the text. Cells with bold numbers [kg/s] are containing measurements (20 points in the average), while the inter- and extrapolated cells are shown with smaller italic numbers.

z [mAMSL]	Niederaussem: CO_2 flux type 2 (wind x mass) [kg/s] above background in the grid cells																					
1300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	2.7	2.4	3.9	2.3	2.0	3.4	3.2	5.6	7.0	6.7	0.6	0.1	0.1	0.1
1100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	5.3	4.7	7.8	4.6	4.0	6.8	6.4	11.1	14.1	13.3	1.2	0.3	0.3	0.1
1000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	5.3	4.7	7.8	2.6	0.6	4.0	11.3	9.7	3.0	2.5	0.1	0.0	0.0	0.0
900	0.3	0.1	1.8	0.3	0.3	4.9	3.6	1.3	6.7	11.2	16.0	15.3	4.9	0.1	0.0	8.0	6.2	1.7	1.4	0.1	0.0	0.0	0.0
800	0.7	0.2	3.7	10.2	9.5	7.2	4.0	4.5	19.2	9.6	7.4	2.0	2.8	7.8	8.4	3.6	2.0	0.1	0.0	0.0	0.0	0.0	0.0
700	0.6	0.5	2.1	5.0	5.1	4.6	3.5	3.5	10.8	7.2	10.3	1.1	6.1	16.4	7.6	0.2	0.2	0.2	0.1	0.2	0.2	0.1	0.1
600	0.4	0.6	0.8	0.5	0.4	1.2	2.1	1.9	2.2	5.2	12.0	4.9	8.8	23.7	13.5	8.3	1.3	6.6	26.4	2.9	1.0	0.1	0.2
500	0.5	0.2	0.5	1.0	0.9	1.8	1.1	2.9	8.3	9.0	6.1	6.7	2.0	9.4	6.0	3.3	1.0	0.3	0.3	0.3	0.4	0.4	0.4
400	1.0	0.7	0.6	1.0	5.6	5.1	2.9	0.8	1.3	1.0	2.0	2.6	1.7	2.8	18.3	31.4	20.2	14.3	11.3	6.2	0.8	1.5	0.3
300	0.5	1.1	1.3	1.0	0.3	0.4	1.3	1.0	2.9	7.4	13.2	1.6	4.0	4.8	1.6	0.7	0.3	0.4	0.4	0.4	0.3	0.3	0.3
200	0.4	0.7	0.9	0.7	0.2	0.3	0.8	0.7	2.0	4.9	8.8	1.1	2.7	3.2	1.0	0.5	0.2	0.3	0.2	0.3	0.2	0.2	0.2
100	0.1	0.2	0.2	0.1	0.0	0.0	0.1	0.1	0.3	0.8	1.5	0.2	0.5	0.6	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
x [m] >	-1200	-1100	-1000	-900	-800	-700	-600	-500	-400	-300	-200	-100	0	100	200	300	400	500	600	700	800	900	1000

Figure 5. As Fig. 4 but for CO_2 'flux 2', i.e. the product of averaged wind and mass in the individual cells. The difference between these two methods is discussed in the text and is one of the sensitivity cases as summarized in Table 3, and detailed in the supplementary.

z [mAMSL]	total CO_2 flux [kg/s] above background in the grid cells (by Kriging, without extrapolations)																						
1300																							
1200																							
1100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	1.4	3.0	3.5	4.0	5.6	7.9	10.3	12.8	9.6	2.5	0.3	0.2	0.1
1000	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	5.0	6.8	5.1	3.4	2.0	4.7	8.2	9.1	8.3	5.0	1.3	0.1	0.1	0.1
900	0.3	0.1	1.7	2.5	1.6	3.3	2.3	1.8	7.3	8.0	9.7	7.7	2.5	1.7	7.3	7.2	4.2	2.3	0.8	0.1	0.0	0.0	0.0
800	0.5	0.3	2.9	3.9	2.8	4.6	3.1	3.3	8.7	7.8	10.1	7.1	4.1	6.3	8.3	4.0	1.1	0.6	1.9	0.3	0.1	0.0	0.1
700	0.5	0.4	2.0	2.5	2.0	3.0	2.6	3.2	5.4	6.7	9.9	5.5	5.9	11.7	8.9	3.5	0.9	2.0	7.2	0.9	0.4	0.1	0.1
600	0.5	0.5	0.7	0.6	0.7	1.2	1.5	2.4	3.4	5.2	7.3	3.5	5.6	13.4	11.2	4.5	0.7	2.9	7.7	1.5	0.6	0.3	0.3
500	0.6	0.5	0.7	1.0	3.6	3.9	1.8	3.0	6.3	7.0	6.6	5.9	2.7	5.2	8.6	8.5	2.9	1.6	1.7	1.1	0.5	0.6	0.4
400	0.8	0.8	0.8	1.0	3.2	2.7	1.7	1.9	3.8	5.8	5.2	3.5	0.7	1.7	7.3	12.2	6.5	4.8	4.9	2.3	0.7	0.7	0.4
300	0.7	0.9	1.2	0.9	0.5	0.5	1.1	1.6	3.6	6.4	3.8	0.0	0.1	0.4	0.9	1.2	0.9	0.4	0.7	0.7	0.3	0.4	0.3
200																							
100																							
0																							
x [m] >	-1200	-1100	-1000	-900	-800	-700	-600	-500	-400	-300	-200	-100	0	100	200	300	400	500	600	700	800	900	1000

Figure 6. As Fig. 4 but for the CO_2 flux calculated by Kriging of the original data instead of averaging in cells and applying our method for linear inter- and extrapolations. The Kriging was only done for the part of the cross section, where the layers contained data. The comparison with our standard method (see supplementary) was done by adding the percentage of the extrapolated flux (up to the stable layer and down to the surface) of the standard method (+10% in this case).

~~A two-dimensional area limited by two cylinders of different radius, by a rectangular box, or by an irregular boundary, with a plume leaving the area. In all these cases, the net flux out of the defined region is concentrated within the part of the boundary, where the plume is crossing. The radial lines along the plume are indicating that instead of the length of a boundary, also angles from a polar coordinate system could be used. The detail on the right is showing the incremental calculation of any fluxes through any shape of boundary: the flux is the product of area A, the density ρ, the concentration c, and the perpendicular wind component v_p, with the wind vector v_w, and the crosswind component v_c.~~

~~Instead of a box or cylinder around the source, we put a virtual box downwind of~~ the ~~source(s) according to Fig.~~layers below 250 ~~?? and ??.~~

~~Instead of a vertical cross-section with infinitesimal thickness, a "wall" with a defined thickness d, where the cross section was flown, was observed with the in-situ measurements. The individual fluxes are calculated from individual wind vectors and concentrations in turbulent resolution (5 Hz). In order to get a spatial distribution of the fluxes this "wall" is divided in grid cells (see Fig,??).~~

~~The vertical cross section of Fig.??. Grid cells containing data are coloured (red denotes high concentrations, blue denotes low concentrations) and cells with no data are left blank~~

~~This "wall" has a certain depth, and was observed for a given amount of time. The fluxes through each "brick" of that wall (a grid with vertical and horizontal spacing according to Fig.~~mAMSL (lower boundary of the cells at nominal 300 ~~??)~~ ~~are calculated in a straight forward manner: Each measurement of an instantaneous flux (concentration times perpendicular~~

Cluster: CO_2 flux [kg/s] above background in the grid cells (sums and linear interpolations & extrapolations)

z [mAMSL] / x [m] >	-2800	-2600	-2400	-2200	-2000	-1800	-1600	-1400	-1200	-1000	-800	-600	-400	-200	0	200	400	600	800	1000	1200	1400	1600	1800	2000	2200	2400
1300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1200	0.0	0.0	0.0	0.0	0.0	0.0	2.4	2.0	3.7	1.1	1.5	1.1	2.7	4.9	4.1	8.3	5.6	1.8	1.0	0.1	6.1	3.3	2.5	2.7	1.2	1.0	0.4
1100	0.0	0.0	0.0	0.0	0.0	0.0	4.8	3.9	7.4	2.2	3.0	2.3	5.5	9.7	8.3	16.7	11.1	3.7	2.0	0.2	12.2	6.6	4.9	5.4	2.5	2.0	0.7
1000	0.0	0.0	0.1	0.0	0.0	0.0	7.2	5.9	11.2	3.3	4.4	3.4	8.2	14.6	12.4	25.0	16.7	5.5	3.0	0.3	18.3	9.8	7.4	8.0	3.7	3.0	1.1
900	0.0	0.0	4.1	8.0	5.9	12.2	12.2	8.1	13.3	9.1	9.6	8.5	12.4	17.1	13.1	18.5	11.7	8.9	11.8	12.2	24.7	18.8	0.0	0.0	8.0	3.4	0.0
800	0.0	0.0	8.2	16.0	13.3	9.6	17.3	10.4	15.4	14.9	14.8	13.5	16.5	19.6	13.7	12.0	6.7	12.2	20.6	24.0	31.2	2.9	4.0	5.2	12.4	3.8	0.0
700	0.0	0.0	0.0	11.1	9.8	14.5	14.0	10.0	14.2	16.3	13.0	11.4	10.9	18.8	15.9	6.9	4.7	24.7	10.4	12.0	15.6	1.4	2.1	2.6	2.9	3.3	4.1
600	0.0	0.0	0.0	6.1	6.3	19.5	10.8	9.5	13.0	17.8	11.1	9.3	5.2	17.9	18.2	1.9	2.8	1.6	0.2	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
500	1.1	0.0	0.0	0.0	8.9	13.1	11.0	2.4	0.6	3.3	2.6	3.4	10.2	8.6	7.0	4.2	12.5	16.1	10.0	9.0	1.8	0.7	0.2	2.8	0.0	0.0	0.0
400	1.0	1.3	0.0	0.0	0.5	0.9	1.0	1.8	8.3	14.1	5.3	6.5	7.6	9.7	6.1	10.0	8.8	1.9	0.7	1.4	4.7	2.5	0.2	0.3	0.1	0.0	0.0
300	1.0	1.1	1.1	0.9	1.2	3.9	15.2	1.8	8.3	14.1	5.3	4.3	9.9	3.9	0.8	0.3	0.3	0.3	0.3	0.2	0.2	0.1	0.3	0.4	0.1	0.0	0.0
200	0.7	0.7	0.7	0.6	0.8	2.6	7.0	1.2	5.6	9.4	3.5	2.9	6.6	2.6	0.6	0.2	0.2	0.2	0.2	0.1	0.1	0.0	0.2	0.3	0.0	0.0	0.0
100	0.2	0.2	0.2	0.2	0.3	0.7	0.0	0.3	1.9	3.2	1.2	1.0	2.2	0.9	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Frimmersdorf
1.23 Mt/a
39.10 kg/s

Figure 7. The same as Fig. 4, however for the mixed plume of the whole cluster on the northern transect. The weak plume from Frimmersdorf, which is only 4 km away from this cross section could be isolated below 450 mAMSL. The confirmation was done with an analysis with higher resolution. The layer between 350 and 450 m (centered at 400 m) has enough measurement points allowing to separate the small lower plume from the main plume from Niederaußem (probably the whole width of about 4.5 km) and the two plumes from Neurath old and new. See Fig. 8 for the separation on a cross section upwind of Frimmersdorf, closer to the three other plumes.

~~component of the wind) is regarded asa sample, and averaged in each grid cell. It is important to note that this includes both the "mean advective fluxes", and the "turbulent fluxes ". For the case where turbulent fluxes occur (both horizontally and vertically from neighbouring grid cells), this approach includes such events. Potentially we would obtain similar results when multiplying average concentrations with average winds per grid cell. However,~~ m) had to be extrapolated. Below the lowest layer with measurements, the fluxes were interpolated linearly to zero at the surface (using the SRTM topography for the flat terrain with an average elevation of 48). Since we are primarily dealing with direct flux measurements, this is the only field we have to extrapolate. However, for calculating 'flux 2' as the product of wind and mass, we extrapolated the concentrations to background at the surface and kept the wind constant. This was due to the observation, that the wind on flight altitude was usually the same as reported at the nearest airport on 10 mAGL. The areas of the cells for mass and fluxes were adapted accordingly in cells with topography. On top, the same procedure was applied: fluxes and concentrations (above background) were extrapolated to zero, with constant wind. In principal, more sophisticated extrapolation schemes could be applied as, for example, discussed in Gordon et al. (2015). Ultimately however, these profiles are not known, and their contribution here is low (less than other uncertainties).

Except for the very small plume of Frimmersdorf, the ~~fully resolved approach is more appropriate (i.e.~~contribution of the extrapolated fluxes was less than 15%. Dealing with individual local fluxes instead of mass and wind separately has an additional advantage for the extrapolation, since only one field has to be extrapolated. In cases, where the changes in wind and concentration are expected to compensate each other (e.g. ~~more complete), because possible co-variances are included. If lower concentrations are associated with lower wind speeds,~~ increasing concentrations for sources near the surface, while the wind is decreasing), also the range of plausible values for the extrapolation has a smaller uncertainty. Whenever highly resolved data of good quality is available, we see no necessity of treating wind and concentrations (mass) separately. Neither for the cells with measurements, nor for those that need to be inter- or extrapolated.

Because the standard method for the inter- and extrapolation of measurements to a cross section in the literature referenced above is Kriging, this was applied as well (graphics program Surfer from Golden Software, including gridding with with several options for Kriging and other interpolation methods). An example of such a result is shown in Fig. 6, and ~~vice versa, a real (turbulent)flux is present.However, when concentrations and winds are averaged first, this turbulent flux would not be detected. With the remote sensing approach discussed in , the turbulent fluxes are indirectly included when fitting the column concentrations to a Gaussian plume model. Note that the in-situ method described here does not assume the shape of the plume. It utilizes knowledge of the statistics about parcels of air crossing an imaginary border. The assumptions and the consequences when they are violated are discussed below. Average concentrations per grid cell were calculated in order to plot these fields for the cross-sections. Concentration fields were calculated as absolute concentrations, and as net concentrations, with the background concentrations subtracted. How the background concentrations were calculated is described below.~~ all results are listed in the supplementary. By qualitative reasoning it seems, that the results are less consistent than with the rules for limited linear interpolation described above. Gordon et al. (2015) has shown that Kriging is best for simulated smooth plumes (Gaussian or similar). However, here highly irregular shapes had to be dealt with. Independent of which settings were chosen for the Kriging (the variogram for the parameter was calculated in any case and both block fits and point fits were tried) it seemed that the fields became too smooth. Kriging was applied both to the directly measured fluxes and the 'flux 2' resulting from mass- and wind-fields after Kriging. We do not claim to have found the optimum method, but, we think, that the difference between the methods is smaller than other uncertainties and that basically, we do not know which concentrations, winds and fluxes are present between the cells with measurements. However, with the rules described above, we make sure that the values in the cells with enough measurements are not affected.

Referring to the importance of turbulent fluxes, Foken et al. (2009) gives a concise overview about the difficulties of complete closures of fluxes. The method with the direct local calculation of fluxes at each measuring point is including the turbulent fluxes in the mean wind direction. In some convective situations, also the vertical turbulent flux above the source can be considered, which was not a necessity in this case. In principle also the cross-wind turbulent diffusion could be calculated. However, this does not contribute to the flux from the source, and would only deliver an estimate for the plume broadening (lateral entrainment and detrainment). The results in the detailed table in the supplementary (difference between '200 $\times$ 100 fb' and 'flux 2') are showing, that the contribution of the turbulent fluxes is not positive, and in all the cases only a few percent. Two reasons are possible in combination: (i) uncertainties in the calculations; (ii) when the turbulent flux is mainly responsible for dilution (entrainment), the turbulence is reducing the net flux.

The positions of the cross-sections were selected ~~based on the flight patterns (~~(filtered from the whole flight) with minimum and maximum distance ~~), and~~ to the source, the mean wind direction, and the lateral distance from the centerline. The angle of the cross-section was adjusted for a cross-wind component of less than $0.1\,\mathrm{m\,s^{-1}}$ ~~or less~~, and the width of the cross section should include enough background concentrations.

~~Figures~~The background concentration for this case study was relatively easy to find, and the results are not sensitive to it because the plume concentrations were high above the background (200 ~~?? and ?? are~~ ppm and more; see Fig. 2). Originally, the standard method was to find the background on each layer by finding the minimum concentration. However, a constant background concentration of 392 ppm was selected in order to avoid artefacts in widely contaminated layers. There is another reason against taking the minimum for each layer, resulting typically in decreasing background concentrations with altitude (not so much in this case): The emissions were injected at low altitudes, into the background concentrations that were present there. When such a plume is rising into lower background concentrations, taking the local background would lead to an overestimation of the emission. Therefore it is better to take the background concentration on the altitude of the sources for the whole cross section.

Fig. 2 can also be used to explain the steps of the processing. Based on the flight track on the map, the minimum and maximum distance from the source was defined. ~~The difference of these distancesis the thickness of the wall, shown as "d" in Fig.~~Within these distances, the measurements are projected onto a vertical plane perpendicular to the mean wind vector. The fluxes from Niederaußem were calculated using measurements between 2 and 5 ~~?? and ??~~km distance, on a cross section in 3.5 km distance (this distance is not relevant, because the projection is parallel). The orientation of the ~~wall is~~ cross section was adjusted until the amount of the average crosswind component ~~is~~ was less than $0.1\,\mathrm{m\,s^{-1}}$ ~~or less~~. Larger crosswind components would ~~not be a problem (see Sect.??). However, it makes sense to adjust~~ cause artefacts by the same reason as would be the case when the cross section would be aligned with the flight track, applying wind components perpendicular to this plane. If all the flight tracks would perfectly overlap, the orientation of the cross section would not be important, and both options would be possible (exactly perpendicular to the ~~average winddirection. In a second step, the lateral boundaries are chosen. They should be clearly outside of the plume, but not too far away since this would increase the uncertainty about the background concentration. The wider~~ wind, or along the flight track). However, when different tracks in different distances are involved, the maximum concentrations on the different traverses would not be aligned and would not be averaged in the same grid cell of the cross section~~the more likely secondary sources are included in the flux calculation (see below, when discussing the background) .~~ . Instead, they would contribute to neighboring cells, increasing these fluxes. Since the flight track in other sectors was quite complex, observing this rule was very important. Another precaution was applied: Even if the wind measurement should be accurate also in steep turns, parts with more than five degrees bank angle (roll) were excluded. Since most parts of the plumes were crossed on straight flight legs, this did not reduce the available data considerably.

~~The horizontal and vertical resolutions $hres$ and $zres$ were either 100or in case of sparse data 200. All calculations were done in vertical columns of $hres \times hres$, i.e. all the measurements within d - which is in in the order of a few hundred meters - were projected onto a wall of thickness $hres$. In the vertical, the grid spacing was $zres$. Therefore, the calculations were done in grid cells with the dimension $hres \times hres \times zres$, using data in the volumes $hres \times d \times zres$. In Fig.??, all grid cells containing~~

~~data are coloured and cells with no data are left blank. The algorithm that interpolates~~ For the separation of the individual contributions of sources that were emitting in the same cross section, two methods were applied: In the case of Frimmersdorf, the small plume was identified on the cross section in the lee of all the sources. The parameters height and width of the cross section ~~begins on the top level, where measurements might be sparse. In the example shown, the cells E7 to H7 and J7 to K7 are interpolated linearly, where A7 to B7 would be kept at the value of C7. Missing values in E3 and K3 are interpolated vertically from neighbouring grid cells E2/E4 and K2/K4 respectively.The same for A5, M3, M4 and M6.~~

~~The linear interpolations and extrapolations were sufficient for filling the grids (mainly for graphical reasons), because our focus was on measured data, and the results should not depend too much from interpolations and extrapolations.~~

~~Usually there was a gap between the lowest flight track, and the surface (taken from SRTM). The maximum of the digital terrain model (DTM) below the flight track was taken as the terrain elevation below the column. The different parameters were treated differently.~~

~~Concentrations were kept constant in the grid cells below the lowest flight track. This might result in an overestimation for~~ ~~CO_2~~~~. However, this is not very relevant, and we just do not have more accurate information. Compare also the comment concerning fluxes.~~

~~Masses for column concentrations are proportional to the gap, i.e.in cell C0 in Fig.~~could then be adjusted to cover this plume only. This part on the larger cross section is marked in Fig. ~~??, about 55~~7. However, on the cross section 'cluster 3' (see Fig. 8), the three plumes (Niederaußem, Neurath new and old) were overlapping in a way that a direct separation was not possible. Therefore, two regions on the cross section were defined, where most likely the sources Neurath old and new were dominating. The source Niederaußem, 9 km upwind of this cross-section, was most likely contaminating this whole cross-section, mainly above 450 m, laterally within the limits marked with dashed lines. The percentage of contributions in the overlapping parts was varied in a wide range between 20 and 80% ~~of the mass of cell C1 is taken~~(from Neurath old or new vs. Niederaußem).

~~The wind is zero below the DTM. However, the fluxes above the surface are taken from the layer above. For~~ ~~CO_2~~ ~~this is often not relevant, since the plumes for example from power plants do not reach the surface close to the source, and it is possible to fly in the background concentrations below the plume.~~

~~Extrapolation~~ Using the difference between the whole cluster (measured directly on the most downwind cross-section), and 'cluster 3' (all power plants except Frimmersdorf) measured directly plus the small plume of Frimmersdorf it was possible to estimate the flux that was not captured on the cross-section 'cluster 3'. The underestimation is due to a rather cautious extrapolation above the highest flight track~~: In cases, where the "curtain" flown did not reach the top of the plume extrapolation above the highest track is applied.~~

~~The background concentrations were assumed to be~~ . Of course also the direct measurement of the whole cluster could be underestimated by the same reason, but by adjusting the two, the ~~minimum concentrations in each layer (1 to 7 in Fig.??). It is clear that the background is not just one concentration for all altitudes. However, it is clear that this is a sensitive parameter. Taking the minimum per layer means that any enhancement above this lowest possible background concentration is considered to be a flux from the source(s) under study. However, concentrations exceeding the background could also be caused by sources far away, or convective mixing in the boundary layer. Taking the minimum tends to overestimate the local fluxes.~~

z [mAMSL]	Cluster 3: CO_2 flux [kg/s] above background in the grid cells (sums and linear interpolations & extrapolations)																					
1300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1200	0	0	0	0	0	0	1	2	0	5	5	2	1	7	4	4	2	2	1	1	1	2	0
1100	0	0	0	0	0	1	2	4	1	10	10	3	2	14	8	8	5	3	2	2	2	4	0
1000	0	0	0	0	0	1	4	6	1	16	14	5	3	21	11	11	7	5	3	3	2	5	0
900	0	0	0	0	0	1	2	5	1	12	0	0	0	29	32	1	0	2	1	0	0	0	0
800	0	0	0	0	0	0	0	5	0	8	0	0	0	38	10	8	0	0	3	11	5	0	0
700	0	3	1	0	0	1	17	13	11	9	16	23	21	19	5	4	0	0	0	0	0	2	0
600	1	7	2	0	0	3	33	20	22	9	11	9	4	0	0	0	0	0	0	2	0	0	0
500	0	1	1	1	2	12	30	17	13	7	6	4	3	1	1	0	3	0	0	0	0	1	0
400	0	2	4	12	35	37	27	13	4	4	1	0	1	3	11	5	0	0	0	1	0	0	0
300	0	2	4	12	35	37	27	13	4	4	1	0	0	0	0	0	0	1	1	0	0	0	0
200	0	1	3	8	23	24	18	9	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	1	3	9	9	7	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x [m] >	-2200	-2000	-1800	-1600	-1400	-1200	-1000	-800	-600	-400	-200	0	200	400	600	800	1000	1200	1400	1600	1800	2000	2200

overlap	Neurath old	total	Neurath new	overlap
6.9 Mt/a	20.3 Mt/a	38.7 Mt/a	12.3 Mt/a	9.9 Mt/a
217.5 kg/s	642.3 kg/s	1226.6 kg/s	388.9 kg/s	312.8 kg/s

These sums are before correcting with overlap and missing flux in the whole cross section!

Figure 8. The same as Figs. 4 and 7, on the cross section downwind of the total cluster without Frimmersdorf (S4). The two domains and the overlap with the plume from Niederaußem (S1) in the background was used to separate the almost collinear plumes of Niederaußem and Neurath old (S2) and new (S3). The uncertainty was quantified by varying the parameters in the overlapping cells between 20 and 80% (background from Niederaußem versus Neurath old and new).

~~However, in this case study, where the sources were close together,~~ budget is coherent, as a lower limit of the real emissions. Of course all these operations are not exact. Therefore the error estimates were done quite conservatively by taking the minima and maxima of the results of all the assumptions (e.g. the percentage mentioned above), resulting in error estimates of 22% for Neurath old and 50% for Neurath new. The difference is due to the fact that Neurath old is closer to the cross-section flown than Neurath new, and ~~the background concentrations had to be taken very close to the plumes, the contrary is true: There is a tendency to overestimate the locally contaminated background. On the other hand, the enhanced concentrations in the plumes were very pronounced in this case. These and other contributions to uncertainties are discussed below~~was therefore better distinguishable on the cross-section.

3.2 Discussion about main uncertainties in-situ

3.2.1 Measurement errors

The wind components (u, v, w) have an accuracy of $0.5\,\mathrm{m\,s^{-1}}$ each, while the CO_2 concentrations have an accuracy of better than 1 ppm~~, and the of better than 5~~. However, as the background concentrations are subtracted, the absolute concentrations are not important, resulting in uncertainties in terms of precision (stability of the sensitivity within an hour) instead of absolute accuracies~~, which leaves us with~~. This is resulting in maximum uncertainties of 0.5 ppm for CO_2 (using the two CO_2 instruments in combination with flask samples)~~and 2for~~. The uncertainties of the wind measurements remain in the order of $0.5\,\mathrm{m\,s^{-1}}$ per component (as for the perpendicular component on the cross section). The main uncertainty is the crosswind component relative to the aircraft. However, when flying back and forth through any plume on a similar altitude, then this error is averaged out. ~~Nevertheless, we make a coarse error assessment with these numbers~~Since this was not always possible, we are taking the worst case $0.5\,\mathrm{m\,s^{-1}}$. Then the relative error based on the wind measurement is 10% at wind speeds in the order of $5\,\mathrm{m\,s^{-1}}$, increasing with weaker winds, and decreasing with higher wind speeds. For CO_2, a plume with moderate 50 ppm above background only adds another percent~~, and the uncertainty of the excessive is in the order of one permille~~. In the case discussed here, wind speeds were around $8\,\mathrm{m\,s^{-1}}$, with excessive CO_2 concentrations of more than 100 ppm (see ~~Fig~~Figs. ~~??~~2 and 3), leading to the conclusion that the error for horizontal fluxes due to the uncertainties of the primary measurements is clearly below 10%. This is in agreement with the assessment found in Gordon et al. (2015). In summary: The measurement itself - if state-of-the-art - is not the main source of errors.

3.2.2 Other sources for errors

The accuracy and precision discussed above are important, however, not the main criterion for reliable flux estimates. ~~It is just the "conditio sine qua non".~~ Only under unrealistically ideal conditions, where winds and concentrations over the whole cross sections could be captured completely and instantaneously, and could be repeated many times, the remaining uncertainties would arise from the (systematic) measurement errors. These are, as discussed above, quite small. However, as already mentioned in Sect. 3.1, and discussed in similar work referenced there, other reasons for uncertainties can be dominating, which are ~~hard~~ harder to quantify. The approach used here was to vary the assumptions and parameters used in the calculations in a wide range. This sensitivity analysis allowed comparing the methodological variations with the average ~~fluxes~~results.

The main uncertainties are dependent ~~from~~ on the meteorological conditions and the flight patterns. An ideal flight pattern is covering the plume as completely as possible. For the CO_2 plumes, the minimum height of 50~~or even more~~ m or 150 m above ground is generally ~~not a problem when measuring sufficiently~~ sufficient when measuring close to the ~~sources, because emissions were originating from high chimneys. For the extrapolation to the surface, the concentrations are kept constant (see Fig.~~elevated hot sources. The main source for uncertainties in the flux estimates are the interpolations and extrapolations as discussed in Sect. ~~??), and the wind is linearly reduced to zero at the surface (a logarithmic wind profile would cause a higher risk for an overestimation of the fluxes). Due to reasons discussed above, when explaining the method with the "gridded wall", it is clear that the wall should be as thin as possible, because otherwise, artefacts could occur. In this case, the flight pattern was~~

~~not ideal due to air space restrictions. Also the wind direction was not ideal, because some of the four point sources were nearly collinear. Nevertheless this case was chosen, because the conditions for the remote-sensing (clear sky) were best. Therefore, the uncertainties for the in-situ measurements discussed here are worst-case, and should not be generalised~~3.1.

Another source of uncertainty already mentioned is instationarity, i.e. varying source strengths from day to day, or even by the hour, while the measurements account only for the emissions for the time of the overflight. 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 (see Sect. 5). Another type of instationarity is caused by the atmospheric turbulence on the scale of a few hundred meters, where a maximum concentration (puff instead of continuous plume) can be missed (=underestimation), or captured by coincidence ~~, and therefore overestimated in the average~~(=overestimation). The only way to minimize this is repeating the pattern as often as possible, and is another reason to spend as much time as possible on the single screens (cross sections). Generally spoken, the 4-d inhomogeneities cannot be captured in a "snapshot". ~~This effect can only be minimised with repeated, dense flight tracks.~~

Another issue is the definition of the boundary layer. In an ideal case, cross sections are flown to an altitude, where no ~~excessive concentrations~~ concentrations in excess of the background are detected anymore, i.e. where the plume is confined below. ~~Also this was not ideal at the day of this case study, because the convective atmospheric boundary layer was higher than it was possible to complete the cross sections due to air space restrictions~~Such a plume was captured under ideal conditions with no extrapolation needed at all on August 23 near a single source. Therefore, this source "Weisweiler" is listed in the results Table 3 in Sect. 4.1 as a reference. The astonishing conclusion is, that the variation due to different interpolations is comparable to the non-ideal cases in this study, where extrapolation was necessary. However, the uncertainty of the extrapolations has to be added, as it is done in Table 3.

Since a deductive error estimate as it is possible for the basic measurements is not possible for the overall flux ~~result~~results, a sensitivity analysis was applied to all cases. In this case study, the five individual cross sections with fluxes were calculated using ~~nine~~ six sets of parameters ~~. The lower limits for the fluxes were found by applying no extrapolation at all. Then the extrapolation to the surface was added, which only contributed a few percent of the fluxes in four of the five cross sections, and finally an extrapolation up to 1400a.m.s.l. (above mean sea lavel, estimated top of the convective boundary layer based on profiles shown in Fig.~~for our approach, and for Kriging, and with three extreme (unrealistic) assumptions. The percentage of the contributions from the extrapolations is listed in Table ~~**??** and **??**). These three cases were calculated with vertical and horizontal grid resolutions of 50, 100 and 200~~3 as well. They ranged between 10 and 20% for the directly measured plumes, with one exception (Frimmersdorf), where the ~~100were standard (best adapted to the typical vertical distance between flight tracks). This led to a sensitivity study where the medians per cross section were taken as the result and the second lowest and the second highest values (something like 10- and 90-percent percentiles) were considered as the "error bars". This led to uncertainties in the order of ±28~~weak plume increased the extrapolated contribution to 45%. A conservative approach for the total error is to add half of these contributions (assuming a 50% uncertainty in the extrapolations) to the differences found by the sensitivity analysis. The overall uncertainty for the total emissions of the cluster would then be 18%, 14% for ~~the worst cross section (big gap between the lowest track, and the surface), and ±10for the best. They are combined when sums or differences~~

~~of cross sections are calculated when trying to separate the different sources. However, the total was measured in one single cross section which included the three major sources , i.e.the same total as in the results from the remote-sensing~~Niederaußem, while ranging between 32 and 63% for other individual sources (Table 3), which were separated indirectly.

When combining sources by adding or subtracting (e.g. subtracting Frimmersdorf and Niederaußem in order to get an estimate for old and new Neurath), these uncertainties are adding in an unknown way (some systematic errors do not add, but compensate, others are adding as components (root of sums of squares), or have to be added for a possible maximum error. Therefore, when separating sources as described in Sect. 3.1, the extreme values within the parameter space were taken.

3.3 Fluxes from remote sensing ~~greenhous~~ greenhouse gas information

3.3.1 Measurement data

The processing of the MAMAP remote sensing data is based on the methods described by Krings et al. (2011, 2013). A modified version of the Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm (Buchwitz et al., 2000) is used to obtain vertical column information of CH_4 or CO_2. It relies on a least squares fit of the logarithmic simulated radiance spectrum to the measurements after correction for dark signal and pixel-to-pixel gain. Additionally a look-up-table approach has been implemented accounting for varying solar zenith angle (SZA) and surface elevation. The conversion from total columns to column averaged dry air mole fractions (XCH_4, XCO_2) is performed using the proxy method, assuming that locally CH_4 is sufficiently constant to compute XCO_2 (or vice versa for XCH_4). See Frankenberg et al. (2005) and Schepers et al. (2012) for more information on the proxy method and Krings et al. (2011) for its application on MAMAP measurements. This method is suitable for point sources as is the case in this study.

~~Emission rate estimates are then obtained using an inverse Gaussian plume model fitting flux and atmospheric stability. In a second approach mass balance estimates are computed leading to two independent inversion methodologies with the exception of wind information which is taken from the routine analysis of the numerical weather prediction model COSMO-DE and the in-situ turbulence probe of the DIMO aircraft which is used for both methods.~~

~~For more details regarding the inversion approaches see .~~

4 ~~Results~~

3.1 ~~In-situ~~

~~Error analysis for in-situ flux estimates based on different extrapolation scenarios and varying grid resolution (see Sect.3.2.2). The emission rate for Neurath new and old blocks was derived from the difference between Niederaußem and total. Low estimate Median High estimate Niederaußem 20.6 24.0 26.7Neurath (new and old blocks) 23.1 25.0 34.5Total 43.7 49.0 61.2~~

~~The results of the in-situ flux analysis are shown in Table3. The median of the nine methods of calculation as described above for the total emission flux is 49.0, from which 79.1was measured directly, which means that 20.9is determined from the linear extrapolation to the surface. The lower estimate (without any extrapolation, which is unrealistic) is at 43.7, and the "best~~

~~estimate" including an extrapolation to the top of the well mixed boundary layer at 1400 m a.m.s.l. is at 61.2. The low, median and high estimates for the single source Niederaußem are 20.6, 24.0 and 26.7respectively, whereas the two collinear sources Neurath new and old, which could not be separated in a reliable manner, led to an estimate of 25.0(23.1 to 34.5).~~

~~Two days later, where the conditions were more favorable in terms of wind direction and better defined top of the mixed boundary layer, slightly higher values and narrower error margins were found, while the reported emissions were 10lower, indicating that the sub-optimum conditions on August 15 led to an underestimation, however, with realistic minima and maxima according to the sensitivity analysis.~~

~~Interpolated and extrapolated CO_2 concentrations measured in-situ downwind of Niederaußem power plant.~~

3.1 ~~Remote sensing~~

3.0.1 ~~Measurement data~~

~~The column averaged dry air mole fractions of CO_2 were retrieved using CH_4 as a light path proxy: $XCO_2(CH_4)$.~~ The corresponding background profiles for the ~~linearization~~ linearisation points use the U.S. standard atmosphere (U.S. Committee on Extension to the Standard Atmosphere, 1976) scaled to actual values. In this case, background XCO_2 was ~~set~~ determined to 390 ppm based on in-situ data upwind of the power plants in the boundary layer and results from the SECM model (Reuter et al., 2012) above. The methane background XCH_4 was estimated to about 1.805 ppm also based on in-situ measurements in the upwind area of the boundary layer scaling the standard atmosphere. An assumed uncertainty of the ratio of the background columns of 1% accounts for possible deviations from these values. The spectroscopic data base used for the computation of absorption cross sections was HITRAN ~~2012.~~ 2012 (Rothman et al., 2013).

Aircraft altitude during the measurements was almost constant at about 1590 mAMSL (+/- 25 m), which was also selected for the reference radiative transfer. Assuming a constant aircraft velocity of 200 km/h, the ground scene size is about 22 m x 54 m (cross track $\times$ along track) for the installed optical front telescope. Thereby the along track ground scene size describes the full width at half maximum for the sensitivity along the flight track. During the measurement, the solar zenith angle varied from about 37.5° to 45.3°.

The radiative transfer model was interpolated using a two dimensional look-up table (LUT) based on solar zenith angle and surface elevation. For that the SRTM digital elevation model (Shuttle Radar Topography Mission (SRTM) version 2.1, http://dds.cr.usgs.gov/srtm/version2_1/, a collaborative effort from NASA, NGA as well as the German and Italian Space Agencies) was used. Due to the changing measurement geometry, the conversion factor to correct for the altitude sensitivity effect (Krings et al., 2011) has to be determined for each measurement independently using also a LUT. This correction takes into account that light passes twice below the aircraft where the observed plumes are located. On average, the conversion factor for the present measurements is about 0.49.

3.0.1 Quality filtering

~~Fit quality of the retrievals ordered by value of the root mean square (RMS) value of the difference between fit and measurement. The green horizontal line denotes the filter threshold.~~

~~Averages of the profile scaling factor ratios of CO_2/CH_4 versus the maximum signal. The measurements displayed in grey left of the green vertical line denoting the signal threshold of 13000counts are excluded from the data. The increased values with respect to the main distribution are due to the actual CO_2 plumes. The ratios have not been corrected for the altitude sensitivity (Sect.3.3.1) and do therefore not denote XCO_2.~~

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 ~~(see Fig.??)~~ 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).

An additional filter has been applied dependent on the signal strength to avoid measurements close to saturation (more than ≈90% detector filling) and in the lower signal to noise range, e.g., over water which has a lower surface spectral reflectance in the short-wave infrared. Filtering of the data accounts for not only SNR but also whether linear full well is achieved. For the full well ADC range selected 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 selected maximum threshold value affects only 4 single measurements during the whole measurement period.

Measurements with a detector filling of less than about 20% (13000 counts) appear to have a slight signal dependency ~~(Fig.??)~~ and were neglected for the inversion process. Furthermore to ensure nadir viewing geometry the deviation from the vertical was not allowed to exceed 5°.

The ~~XCH_4(CO~~XCO_2(CH_4) precision after filtering is approximately 0.29% determined from the standard deviation of the data outside the plume area.

Fig. 9 shows the XCO_2(CH_4) data acquired over the coal fired power plants without and with the filtering applied. Clearly visible are the overlapping CO_2 plumes originating at the individual power plant locations and advected in downwind direction towards North-West in agreement with the wind field as computed by the ~~COSMO-DE model~~ the routine analysis of the numerical weather prediction model COSMO-DE (Doms, 2011).

3.0.2 Atmospheric conditions and wind information

A fundamental parameter for the inversion is wind speed. To compute an average wind speed throughout the CO_2 plume from model and in-situ data also the boundary layer depth is important.

~~From left to right, in-situ data of temperature, virtual potential temperature, aerosol particles larger than 0.3and aerosol particles larger than 0.5. For temperature and virtual potential temperature also the COSMO-DE model result is shown and for the temperature additionally ground based in-situ data from Dsseldorf airport which is approximately 30NNE of Neurath power plant.~~

~~Airborne in-situ measurements of CO_2 in the whole measurement area (left) and upwind of the plume (right).~~

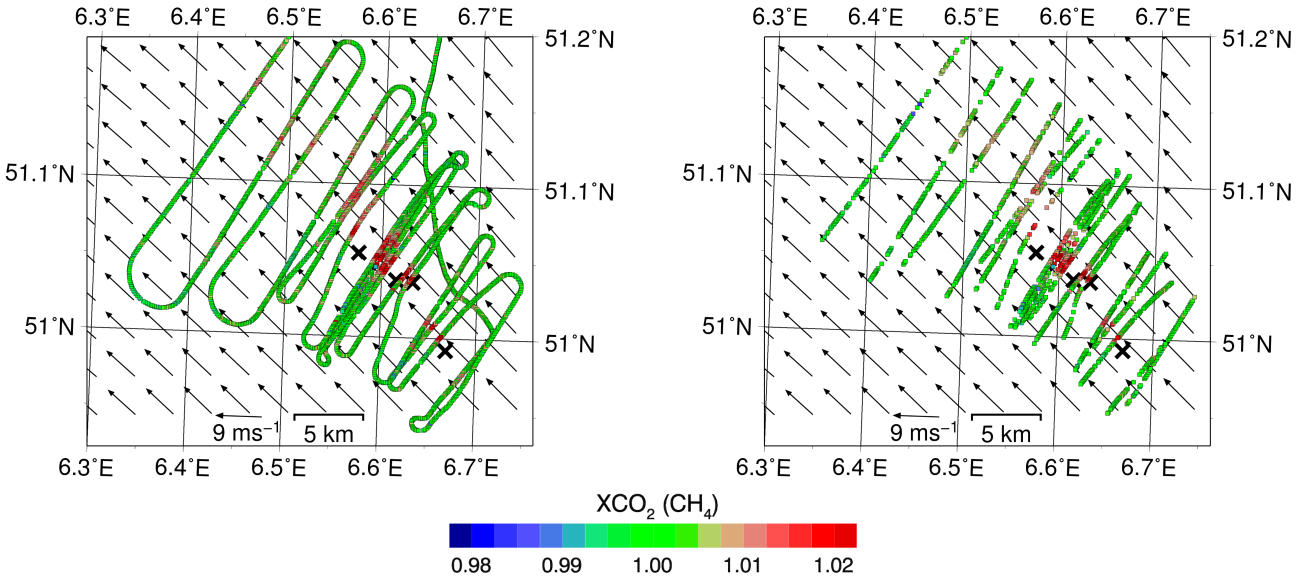

Figure 9. Qualitative MAMAP remote sensing XCO_2 data unfiltered (left) and ~~the cross sections~~ filtered as described in the main text (right). The crosses denote the power plant locations (see Fig. 1). Arrows denote wind vectors from the COSMO-DE model at an altitude of about 350 ~~above sea level~~ mAMSL (model layer 45) at 13:00 UTC.

~~Fig.?? shows the in-situ and COSMO-DE result for temperature~~ The aerosol and virtual potential temperature ~~for the model grid point about 1.5North of Neurath power plant in the centre of the main plume area. In a first approximation the boundary layer depth can be identified by a sharp increase in virtual potential temperature. While in the COSMO-DE model data there appears to be a weakly developed boundary layer evolving from about 400altitude at 11:00to about 800at about 14:00there is no indication in the in-situ virtual potential temperature data~~ profiles give no indication that the transition to the free troposphere is located in the lower 1100 m. ~~The aerosol distribution is also not conclusive in this respect. This is because a sharp decrease at the transition to the free troposphere might be expected.~~ This is furthermore confirmed by the fact that there are enhanced CO_2 amounts throughout the probed altitude layers~~(see Fig.??)~~. Consequently it can only be concluded that the boundary layer depth is likely larger than 1100 m. For the analysis a boundary layer depth of 1500 m was assumed (with uncertainty estimates for cases of 1200 m and 1800 m, see Sect. 3.1).

~~Modelled and measured wind direction are similar in lower altitudes but agree less well at higher altitudes and later times. However, the wind direction is derived from the remote sensing data directly without using the model information.~~

Figure 10 shows measured wind speed and direction for the ~~same modelgrid point~~day analysed in this study compared to results from the COSMO-DE model. The measured in-situ wind speed averaged over 60 s ranges from about 7 $m\,s^{-1}$ to 11 $m\,s^{-1}$ for the time and altitude range of the overflight and is rather constant with altitude. Altitudes covered by in-situ measurements range from about 300 m to 1100 ~~a. m.s. l~~mAMSL.

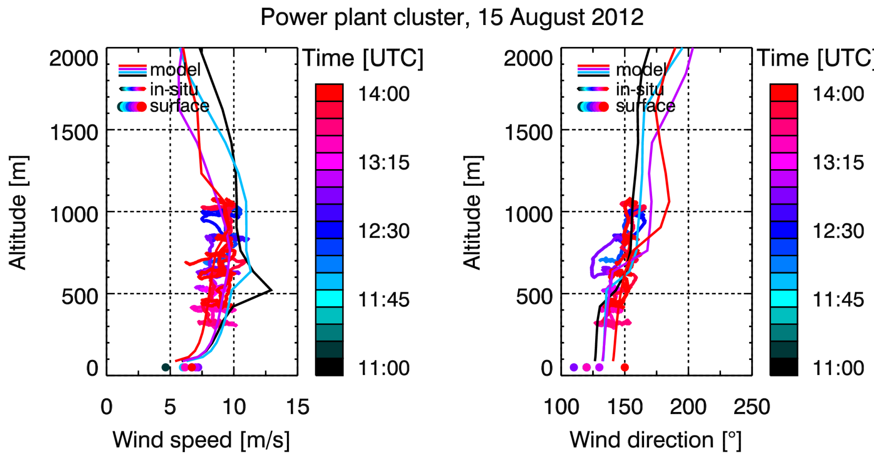

Figure 10. Wind speed (left) and direction (right) as measured from the in-situ turbulence probe as well as COSMO-DE model data and surface in-situ information from Düsseldorf airport about 30 km north-north-east of Neurath power plant (obtained from weather underground, http://www.wunderground.com).

Modelled and measured wind direction are similar in lower altitudes but agree less well at higher altitudes and later times (Fig. 10). However, the wind direction is derived from the remote sensing data directly without using the model information.

Table 1. Comparison of modeled and measured wind speed for several altitude layers.

Altitude range ~~a.m.s.l.~~ [mAMSL]	Model wind speed $[\mathrm{m\,s^{-1}}]$	In-situ wind speed $[\mathrm{m\,s^{-1}}]$	Wind speed difference (in-situ - model) $[\mathrm{m\,s^{-1}}]$	Relative wind speed difference (in-situ - model)/model [%]
291 – 440	8.99	8.53	-0.46	-5.1
440 – 588	9.33	8.47	-0.86	-9.2
588 – 737	9.48	8.53	-0.95	-10.0
737 – 885	9.84	9.11	-0.74	-7.5
885 – 1034	9.27	9.40	+0.13	+1.4
			Average: $-0.58\,\mathrm{m\,s^{-1}}$	Average: -6.1%

Since in-situ wind information is not always available in time and space where remote sensing measurements are taken, the in-situ wind data is used to calibrate the COSMO-DE model result in the measurement area during the time of the remote sensing overflights.

The precision of the wind model was estimated to about 0.9 $\mathrm{m\,s^{-1}}$ (1 σ) with negligible bias for the case described in Krings et al. (2011). Assuming the same error holds in the present study, this leads to a wind based relative error (1 σ) on the inversion of about 10%.

This error can be reduced when on site wind information is available, for example, from airborne turbulence measurements as they were used in Krings et al. (2013) and that were also performed during the present campaign.

~~3D representation of the in-situ flight track. Also shown are projections on the ground and on a vertical plane for better interpretation.~~

~~Figure?? shows the flight pattern of the in-situ measurements. The measurements concentrate on transects at several altitude layers around the two Neurath power plants and around the extended area including Frimmersdorf power plant.~~

To quantify the difference between measured and modeled wind speed, the probed altitude range has been divided into 5 equally thick layers in which the deviation between the in-situ measurements on the one hand and the associated model data interpolated in time and space on the other hand were computed and subsequently averaged over the altitude layers.

For the available model and measurement data from the target area and time, this yields an average overestimation of about 0.58 $\mathrm{m\,s^{-1}}$, i.e. about 6.1%. This is well within the approximate error of about 0.9 $\mathrm{m\,s^{-1}}$. Similar to Krings et al. (2013) the in-situ wind error of 0.5 $\mathrm{m\,s^{-1}}$ is assigned to the calibrated wind.

The complete results are shown in Table 1. Note that the results do not directly relate to Fig. 10 which only shows the model wind speed at one specific COSMO-DE grid point while model wind data from the whole measurement area enters the computations in Table 1. The standard deviation of the model wind speed over the measurement area for the model layer shown in Fig. 10 is about 5.8%.

3.0.3 Inversion for emission data

Emission rate estimates are then obtained using an inverse Gaussian plume model fitting flux and atmospheric stability. In a second approach mass balance estimates are computed leading to two independent inversion methodologies with the exception of wind information which is taken from COSMO-DE model and the in-situ turbulence probe of the DIMO aircraft and that is used for both methods.

Preparation and performance of the Gaussian plume inversion and the integral inversion method is very much in line with Krings et al (2011, 2013). Since the inversion proved to be extremely stable no a priori information on emission rate or stability were required simplifying the cost function to be minimised in the iterative inversion process.

The data were gridded to pixels of 35 m~~x~~ $\times$ 35 m having approximately the same area as the MAMAP ground scene size. The impact of different pixel sizes for the gridding is assessed in Sect. 3.1. No additional smoothing was applied. Note that the gridding was not used for the mass balance method.

Prior to the inversion the data were normalised dividing by the regional background. Since the measured area and time is somewhat larger than in the previous studies, no constant normalisation was selected but a track by track procedure that is also able to account, for example, for linear gradients that are unrelated to the source. Thereby data from each cross section (see Fig. 9) is normalised by a linear function that is determined by the flanks of the track excluding the plume area. If the track does not measure sufficient data outside the plume, then this method results in an underestimation of emissions. Because of that and because gaps due to filtering are rather large from track 10 onwards, only the first 9 downwind tracks with respect to Niederaußem power plant ~~have been included in the analysis so that emission from Frimmersdorf were not determined in this approach~~could be analysed using the standard filtering. To investigate the measurements from further downwind, the signal threshold was relaxed 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 it is ensured that a sufficient set of measurements, even if of lower quality, are available for interpretation. The Gaussian plume method was not applied for these data as that would require to mix data which were subject to different filter criteria.

Obtaining an adequate estimate of the mean wind speed with which the emitted gas is transported is generally challenging when there are ~~serveral~~ several sources which are separated in downwind direction. ~~Naturally, the relative emission rates of the sources will generally not be known before the data analysis. Therefore the weighting of the wind profile according to the significant altitude layers is not trivial.~~ In the present case study the low variability of wind speed with altitude (see Fig. 10)~~however~~ , however, makes the estimation less sensitive to errors in this regard. ~~The~~ Here, the mean wind speed was estimated assuming Niederaußem power plant, the power plant that is located most ~~downwind~~ upwind of all emitters, as ~~the~~ an only source. The emitted CO_2 was then distributed using a vertical Gaussian dispersion with the stability parameter resulting from ~~the inversion of 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. The cross section wind speeds~~ flight legs based on relative concentrations per altitude layer. The resulting wind speeds for each remote sensing flight leg were used individually for the mass balance approach and averaged for the inverse plume model over the relevant area. At the first remote sensing leg 700 m downwind of power plant Niederaußem, the plume reaches about 1 km height, and at 2 km downwind distance the CO_2 is already well mixed according to the plume model which represents a temporal average.

Boundary layer top height ~~(represented as a reflective layer as in)~~ (represented as a reflective layer as in Turner, 1970) and emission height needed for the vertical dispersion were varied around the baseline parameters to estimate the range of errors resulting from these assumptions (see Sect. 3.1).

~~Wind direction was determined from the measured remote sensing data to about 147.5°by both visual inspection and minimizing the stabiliy parameter (see also Sect.3.1). When fitting the stability parameter to the retrieved XCO_2 this yields a=214.0 (±8.8statistical error), i.e.stability class A (very unstable atmospheric conditions,), independent of wind speed.~~

~~Using this stability as input for a vertical Gaussian dispersion model as a function of distance to the source to compute a weighted mean of the wind profile, an average wind speed of 8.2was obtained for the plume area covered by the first 9 downwind tracks of the MAMAP data. For this the in-situ calibrated wind model data was used.~~

~~Applying this wind speed to the Gaussian plume inversion, the result for the average emission rate for the time of the overflight is in total about 63.6(±3.0statistical error) split into 24.0(±4.6), 14.2(±6.3) and 25.4(±5.2) for the power plants Niederaußem, Neurath new and Neurath old, respectively. As mentioned before the evaluated tracks are all located upwind of power plant Frimmersdorf. Therefore no emission rate for Frimmersdorf was derived. The contour lines as an overlay on the retrieved XCO_2 can be seen in Fig.12.~~

~~Gridded MAMAP XCO_2 results rotated so that wind direction points into positive x-direction and contour lines of the inferred plume models for the individual power plants. Total emission rate is 63.6for the time of the overflight. Ground scenes are shown slightly enlarged for better visiblity.~~

~~For the mass balance approach a wind speed was computed for each individual track ranging from about 8to 9. Similar to the inverse plume model, the first 9 downwind tracks were analysed and the associated cross sections are shown in Fig.13. The data were normalised for each flight track individually using a linear fit based on the data outside the plume. This was applied in order to account for local gradients or other offsets. The definitions of the outside plume area are listed in Table4.~~

~~MAMAP XCO_2 cross sections for the tracks used for the emission rate estimates. The area outside the dashed vertical lines denote the data that were used for the normalisation.~~

~~Normalisation distances to the end of the measurement track for each individual remote sensing transect. Track Distance to end of track Comment Upwind, Downwind 3 – 5 2000Baseline normalisation length for shorter tracksDownwind 1 3000Avoid measurements with increase in CH_4 next to plumeDownwind 2 1500Plume not centered on track Downwind 6 – 9 3000Track lengths increased and wider plume further downwind~~

~~The results are shown in Fig.15 for the individual tracks and the average emission in-between power plants. Figure15 also shows that there is basically no influx from upwind into the measurement area.~~

~~Mass balance results based on MAMAP remote sensing data. Vertical lines denote the location of power plants as downwind distance from Niederaußem. Horizontal lines and emission values show average total emissions of the upwind sources.~~

3.0.4 ~~Error assessment~~

3.1 Error assessment

The influence of various parameters on the inversion results was investigated. This was mainly accomplished by evaluating errors introduced by uncertainties in the input parameters for the inversion methods, except for the statistical error on the plume inversion which yields about 3% for the total emission estimate and about 6% maximum error for a single power plant emission ~~.~~

~~The computation of an average wind speed for the entire plume area (for the inverse plume model)~~ and for the ~~single tracks (when using~~ statistical error on the mass balance approach ~~), respectively, is challenging. For this work the wind speed was estimated using a forward model for the plume dispersion assuming a single source at the location of Niederaußem power plant with an emission height of 250including initial plume rise. Furthermore a boundary layer depth of 1500was assumed. However, the boundary layer depth is not well defined~~ derived from variability of inversion results (see Sect. ~~3.0.2)and the assumption of a single source is a simplification. To obtain a very rough estimate for the uncertainty on the resulting emission rates wind speeds were computed also for boundary layer depths of 1200and 1800, and similarly for emission heights of 200and 300.~~

~~For the mass balance approach the impact of the emission height on the emission results is less than 1for the closest tracks and reduces for the transects further from the source. The error due to uncertainties in the boundary layer depth are largest for~~

~~the far tracks since the plume will be more mixed. This results in an error of about -6for the greater depth and an error of +4for the reduced boundary layer depth.~~

~~For the inverse Gaussian plume model, the error on the emission height is always less than 0.5and largest for the emission rate of Niederaußem. The impact of the boundary layer depth is very similar to that of the mass balance approach and within ±6.~~

~~These errors are similar to what can be assumed for the calibrated wind speed model of ±0.5resulting in a percentage error of about ±6.~~

~~In cases where wind speed shows more pronounced variations with height, an average wind speed could be iteratively computed by weighting the sources according to the inversion result after using an a priori distribution of emissions for the initial state~~4.2).

The wind direction of 147.5° for the inversion was determined by visually matching the best wind direction and minimizing the retrieved stability parameter a of the Gaussian plume inversion and hence minimizing the plume width relative to the centre axis (see Fig. 11). This also minimizes the total emission rate. However this minimum does not exactly coincide with the best fit (lowest RMS of differences between model and measurements) which is accomplished for about 145.5°. To account for the fact that the minimum in stability and fit quality is rather wide, an error of about ±4° was assumed. ~~This results in errors on the emission rate of about ±4for the mass balance approach and about +6for total emissions from the inverse plume model.~~

~~Sensitivity of the inverse Gaussian plume model to the wind direction.~~

~~When varying the assumed source width by ±200(baseline was 300, for all power plants the same) the error for the inverse plume model is about ∓3. The grid size was assumed to be 35×35which is not entirely correct since a MAMAP pixel for this work was rectangular (about 22.2×54.4).The grid size was varied between ±20resulting in errors of up to 2but generally lower (see Fig.~~Importantly, Fig. ~~??). Both error sources do not apply for the mass balance approach~~11 shows furthermore that, in principal, wind direction can be fitted directly to the data.

Several quality filters were applied to the data. The filter for fit quality ~~RMS of 0.9(see Fig~~(see Sect. ~~??~~3.0.1) has been set relatively broad to reject only the data of poorest quality. It does not improve the results to relax the filter further or apply a stricter criteria to reject data. The signal filter was set to reject data with a signal below 13000 counts (see Sect. 3.0.1). When relaxing the filter criteria by -2000 counts the total emission estimate is affected by about 6% for the plume inversion and about 1% for the mass balance approach. More strict filtering will significantly reduce the data which particularly impacts the mass balance and the defined plume area and was therefore not applied.

~~To avoid spurious results when the aircraft is turning (as the instruments optics was not mounted on a stabilization platform) an inclination filter was applied rejecting all data for inclination angles deviating more than 5°from the nadir geometry. Relaxing the filter to 10°results in changes for the total emission rate of about 5for the inverse Gaussian plume and about 1for the mass balance approach for the analyzed tracks furthest from the source.~~

~~The uncertainties derived above for the reasonable filter limits enter the final error budget.~~

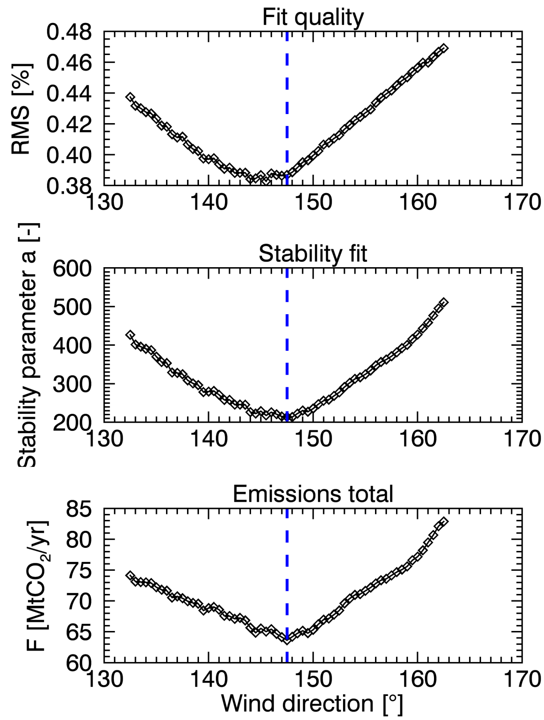

Figure 11. Sensitivity of the inverse Gaussian plume model ~~result~~ to the ~~gridding size. The x-coordinate shows~~ wind direction for flux F, stability parameter a and the ~~edge length of a grid cell~~RMS between measurement and fit.

~~Both inversion methods are impacted by the determination of the background CO_2 profile. For this work the profile was determined to 390(see Sect.3.3.1). As in previous works an uncertainty of approximately 1was assumed leading to an uncertainty of 1on the derived emission rates.~~

The error due to ~~errors~~ uncertainties in the surface elevation model was not investigated ~~. This is because they to a good extent cancel~~ because it largely cancels out for the proxy approach~~and high variations are not to be expected since the open cast mine relocation (compare Sect.??) does not coincide with the measurement tracks of the remote sensing flight~~.

The overall errors (see Table 2) have been computed as root sum square assuming no correlation between errors and yielding about 10% for the mass balance method (not yet including uncertainties based on variability between results from different flight legs as mentioned above) and 15% for the plume inversion approach.

4 ~~Discussion~~Results

4.1 In-situ

The results of the in-situ flux analysis are shown in Table 3. The average of the six methods of calculation (see Sects. 3.1 and 3.2.2) as the best estimate for the total emission flux is 51.6 $\mathrm{MtCO_2 yr^{-1}}$, of which 86% was measured directly.

Table 2. Overview of maximum absolute values of the different error components of the estimated emission rates for the remote sensing results. The uncertainties are derived from a sensitivity analysis and the total error is the root sum square of the individual error components. In case of the mass balance approach an additional uncertainty can be derived from the scatter of inversions for individual tracks (see Fig.15 and Sect. 4.2).

Error source	Mass balance approach	Inverse plume model
wind speed uncertainty (0.5 $\mathrm{m\,s^{-1}}$)	6%	6%
statistical error (maximum)	~~–~~ *	6%
emission height (200–300 mAGL)	1%	0.5%
boundary layer depth (1200–1800 mAMSL)	6%	6%
wind direction (143.5–151.5°)	4%	6%
source width (100–500 m)	–	3%
grid size (15–55 m)	–	2%
signal filter (11000–13000 counts)	1%	6%
inclination filter (5–10°)	1%	5%
~~background profile~~ CO_2 background profile (390 ppm $\pm$ 1%)	1%	1%
Total error	10%	15%

*) statistical errors for the mass balance approach are derived from the resulting emission rates by track in Sect. 4.2

Table 3. Summary of all standard methods applied to the different plumes measured in-situ (for more details see the extended table in the supplementary). For the discussion see text in Sect. 3.1. The 'best estimate' is the average of the sensitivity analyses. The minima and maxima were calculated combining the minima and maxima of either the primary measurements (e.g. for the total cluster, Frimmersdorf, and Niederaußem), or by using the worst-case combinations for sums and differences. The split into Neurath old and new was done as shown in Fig. 8.

Source	best estimate [$\mathrm{MtCO_2yr^{-1}}$]	min [$\mathrm{MtCO_2yr^{-1}}$]	max [$\mathrm{MtCO_2yr^{-1}}$]	error relative to best estimate	fraction of extrapolation	overall uncertainty
total cluster	51.6	50.0	61.3	11%	14%	18%
Frimmersdorf	1.7	1.3	2.7	41%	45%	63%
Neurath old	16.8	13.3	20.8	22%	20%	32%
Neurath new	7.3	3.9	11.3	50%	20%	60%
Niederaußem	25.5	22.0	26.6	9%	10%	14%
Weisweiler (ideal reference case)	18.4	16.3	20.9	13%	0%	13%

Emissions for all individual power plants could be derived. For directly measured fluxes (Niederaußem and total cluster in Table 3), the estimates have an uncertainty in the order of 10–20%. Fluxes derived from differences and sums of the primary fluxes have a much higher uncertainty, especially when the plumes are overlapping (Neurath new and old). Frimmersdorf has a high relative uncertainty due to the fact that 45% of flux had to be derived from extrapolation.

4.2 Remote sensing

Wind direction was determined from the measured remote sensing data to about 147.5° by both visual inspection and minimizing the stability parameter (see also Sect. 3.1). When fitting the stability parameter a to the retrieved XCO_2 this yields a=214.0 (±8.8% statistical error), i.e. stability class A (very unstable atmospheric conditions, Martin (1976); Masters (1998)), independent of wind speed. This is in agreement with the observed convective mixing. However, the stability parameter obtained within the inversion is an effective parameter also subsuming other effects such as increased flew gas temperature and variation of wind direction.

Using this stability as input for a vertical Gaussian dispersion model as a function of distance to the source to compute a weighted mean of the wind profile, an average wind speed of 8.2 $\mathrm{m\,s^{-1}}$ was obtained for the plume area covered by the first 9 downwind tracks of the MAMAP data. For this the in-situ calibrated wind model data was used.

Applying this wind speed to the Gaussian plume inversion, the result for the average emission rate for the time of the overflight is in total about 63.6 $\mathrm{MtCO_2yr^{-1}}$ (±3.0% statistical error) split into 24.0 $\mathrm{MtCO_2yr^{-1}}$ (±4.6%), 14.2 $\mathrm{MtCO_2yr^{-1}}$ (±6.3%) and 25.4 $\mathrm{MtCO_2yr^{-1}}$ (±5.2%) for the power plants Niederaußem, Neurath new and Neurath old, respectively. The overall uncertainty including also other components is about 15% (see Sect. 3.1). The contour lines based on the multiple sources as an overlay on the retrieved XCO_2 can be seen in Fig. 12. Additionally, data and model result per flight leg are shown in Fig. 13. As mentioned before the evaluated tracks are all located upwind of power plant Frimmersdorf. Therefore no emission rate for Frimmersdorf was derived with this approach.

For the mass balance approach a wind speed was computed for each individual track ranging from about 8 $\mathrm{m\,s^{-1}}$ to 9 $\mathrm{m\,s^{-1}}$. Similar to the inverse plume model, the first 9 downwind tracks were analysed and the associated data are shown in Fig. 13. For three tracks further downwind (see) where the usual quality filtering could not be applied, the results have to be interpreted with more caution. At this distance, the plume is considerably wider than on the more upwind tracks so that there is less data available for the normalization.

The data were normalised for each flight track individually using a linear fit based on the data outside the plume (see Sect. 3.0.3). The definitions of the outside plume area are listed in Table 4.

The results are shown in Fig. 15 for the individual tracks and the average emission in-between power plants. Figure 15 also shows that there is basically no CO_2 influx from upwind into the measurement area. Shown in light grey are the tracks further downwind with decreased data quality. The average absolute deviation from the mean was used as an indicator for the precision. The precision is worse for cases where emissions are derived as differences which are subject to error propagation from emission estimates from upwind as well as downwind.

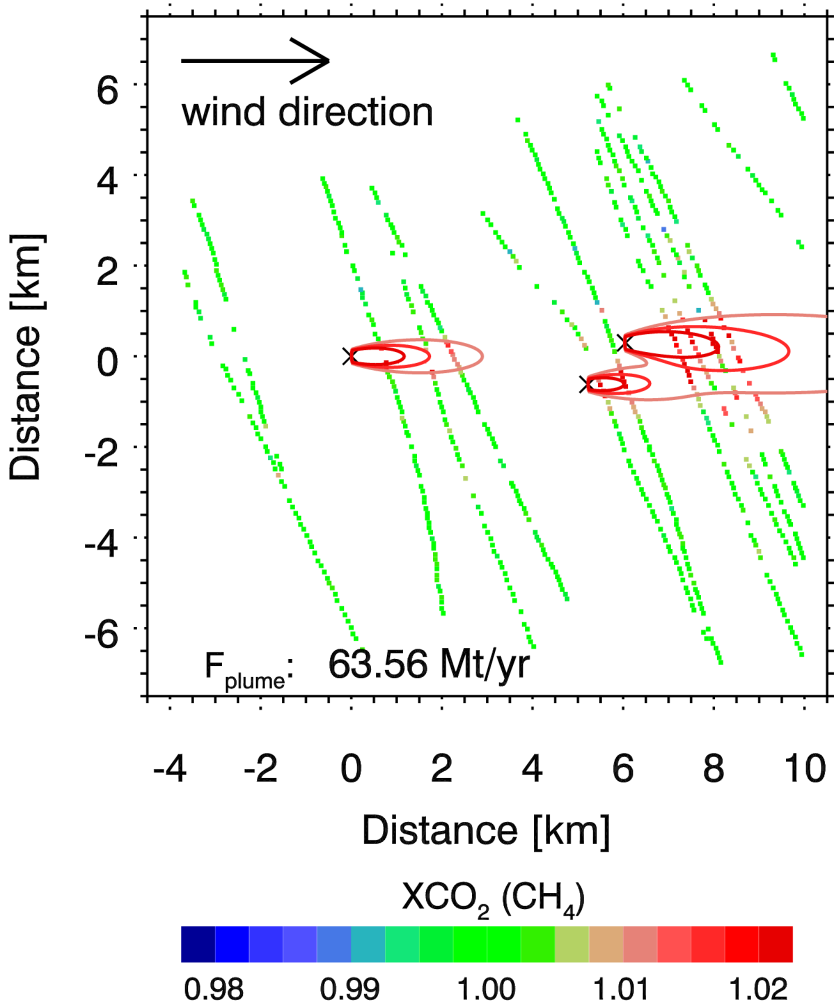

Figure 12. Gridded MAMAP XCO_2 results rotated so that wind direction points into positive x-direction and contour lines of the inferred plume models for the individual power plants. Total emission rate is 63.6 $\mathrm{MtCO_2 yr^{-1}}$ for the time of the overflight. Ground scenes are shown slightly enlarged for better visibility.

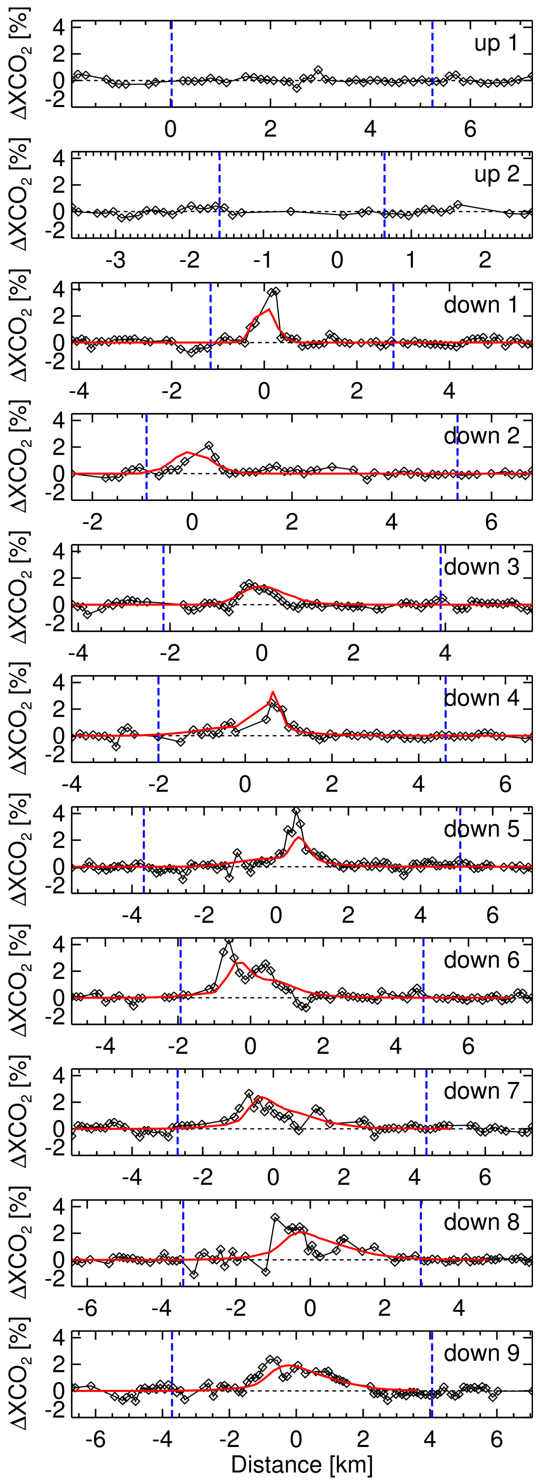

Figure 13. MAMAP XCO_2 transects for the tracks used for the emission rate estimates (black) and modelled Gaussian plume (red). The areas outside the dashed vertical lines denote the data that were used for the normalisation.

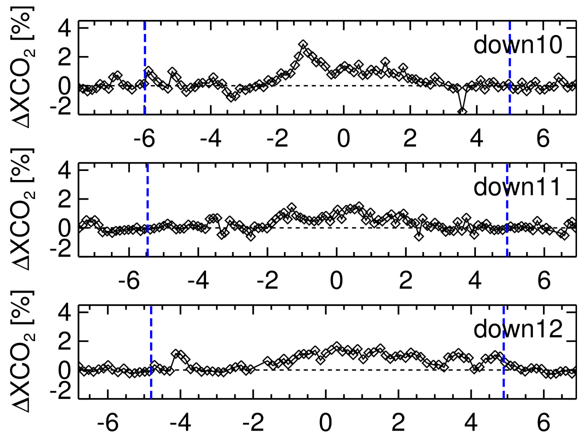

Figure 14. MAMAP XCO_2 measurement transects for the 3 flight legs downwind of power plant Frimmersdorf. The area outside the dashed vertical lines denote the data that were used for the normalisation.

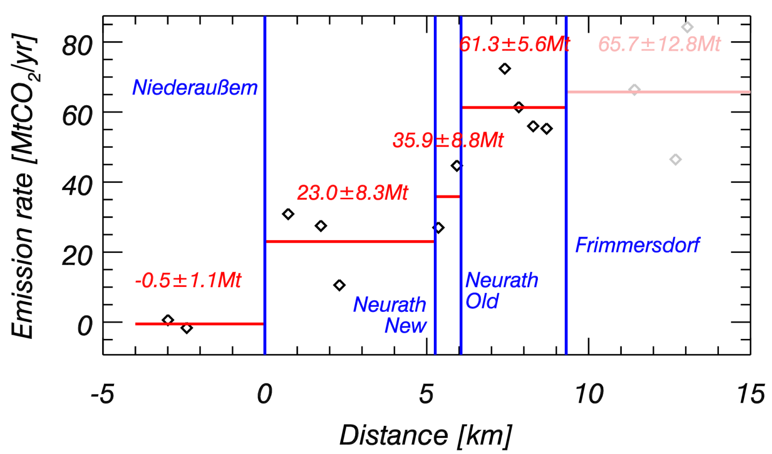

Figure 15. Mass balance results based on MAMAP remote sensing data. Vertical lines denote the location of power plants as downwind distance from Niederaußem. Diamonds show emission rates derived from individual aircraft legs, where grey diamonds indicate reduced data quality (see Sect. 3.0.3). Horizontal lines and emission values show average total emissions of the upwind sources.

Table 4. Normalisation distances to the end of the measurement track for each individual remote sensing transect.

Track	Distance to end of track	Comment
Upwind, Downwind #3 – #5	2000 m	Baseline normalisation length for shorter tracks
Downwind #1	3000 m	Avoid measurements with increase in CH_4 next to plume
Downwind #2	1500 m	Plume not centered on track
Downwind #6 – #9	3000 m	Track lengths increased
Downwind #10 – 12	2000 m	Plume is widened due to distance from the source

5 Discussion of results

The CO_2 emission rate estimates calculated using the different methods for the different power plants are shown in Fig. 16 and comprise the following: MAMAP remote sensing data ~~analyzed~~ analysed with inverse plume model and mass balance approach ~~, respectively~~(Sect. 3.3), in-situ data ~~analyzed using mass balance~~ analysed using the presented mass balance method (Sect. 3.1) and emission rate estimates based on emission factors and energy production data for the time of the overflight. Error bars for the emissions derived from energy production are not shown. 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).

Generally the two inversion approaches for MAMAP agree very well within their uncertainties for the three individual power plants Niederaußem and Neurath (old and new). However, for the mass balance approach, the uncertainties as determined from the variability of emission estimates for individual power plants (see Fig. 15) are larger when differences were computed due to error propagation. This track by track variability is likely due to instationarity of the atmosphere and shows that repeated measurements are vital to obtain an accurate emission estimate. The inverse plume model, which inverts for all power plants simultaneously, is less affected by the instationarity since all available data is considered for all power plants reducing the overall uncertainty.

When comparing with the CO_2 release computed from energy production the agreement is very good for all methods for the emissions from the power plant Niederaußem. For the two Neurath power plants, the remote sensing results indicate less emission from the new units and more from the old units while the overall result is approximately the same. This is then also reflected in the ~~total emissions which are~~ emissions of Niederaußem, Neurath old and new, which is very similar for remote sensing methods and the computed emissions. The ~~total emissions~~ combined emissions of the three power plants are 63.6±15%, 61.3±13% und 63.8 $\mathrm{MtCO_2yr^{-1}}$ for the MAMAP plume inversion, MAMAP mass balance and the computed emissions. The relative difference to the computed emissions is thereby -0.3% (plume inversion) and -3.9% (mass balance).

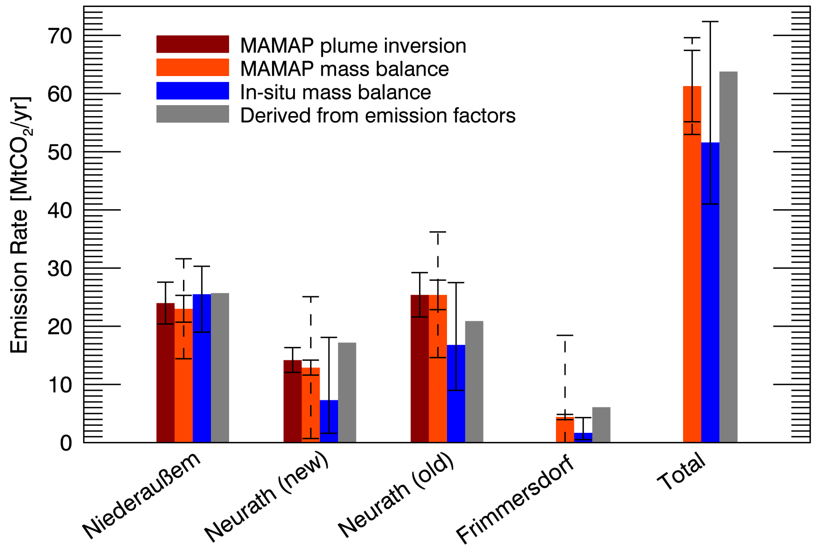

Figure 16. Inversion results compared to results obtained from emission factors and energy production for the time of the overflights. ~~Error bars denote 1σ errors for~~ For the remote sensing ~~, and~~ mass balance approach the ~~10and 90percentile for in-situ~~ smaller error bars denote the uncertainties derived from the sensitivity analysis while the larger include also the precision (see ~~Sect~~Fig. ~~3.2.2~~15) applying the root sum square. The error bars for in-situ are worst-case limits based on the sensitivity analysis and half of the extrapolated emissions below and above of the captured plumes. Emissions from Frimmersdorf were not evaluated with the remote sensing plume inversion.

If no in-situ data had been available, that is if the wind had been derived only from the COSMO-DE model, the errors would have been -6.4% and -10.0% respectively, reflecting the importance of additional wind measurements.

The total emissions derived from in-situ data are ~~somewhat lower for the new and old blocks of power plant Neurath at 25.0(23.1 to 34.5) and for the total at 49.0(43.7 to 61.2)~~in agreement with the emission computed from emission factors but are somewhat lower. The reasons and uncertainties were explained in ~~Sect~~Sects. ~~4.1 and 3.2~~ 3.2 and 4.1. However, the selection of the measurement day for detailed analysis was largely driven by the clear sky requirement for remote sensing and there was no ideal overlap between the optimal measurements from the in situ instruments and suitable remote sensing measurement days. Similar for Neurath and Frimmersdorf, while for Niederaußem, where almost the complete plume could be captured the agreement is exceptionally good.

The remote sensing mass balance results for power plant Frimmersdorf are based on data with less strict filters to obtain a sufficient large data set to compute an emission rate. While the scatter and, hence, the uncertainty is quite large, the mean value indicates an emission rate of 4.4 $\mathrm{MtCO_2yr^{-1}}$ which is close to the emission rate of 6.1 $\mathrm{MtCO_2yr^{-1}}$ based on emission factors. However, the associated error based on the sensitivity analysis and precision is about 8.3 $\mathrm{MtCO_2yr^{-1}}$.

6 Conclusions

This work enhances the comparison between measurement and inversion approaches using in-situ and remote sensing data to obtain emission rates for flue gases from a cluster of point sources with known locations. These sources were partly in close proximity to each other and the plumes of – in this case – CO_2 from coal fired power plants overlapped adding complexity to the inverse problem.

In contrast to the in-situ method, the remote sensing measurements ~~require~~ required clear sky conditions at the time of the measurement. MAMAP measures solar backscattered electromagnetic radiation in the short-wave infrared. To simplify the radiative transfer calculations, cloud free atmospheres are generally 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 restriction impacted the selection of the measurement day for the analysis in this work. This resulted in some days with potentially more favorable conditions for the in-situ method (coverage, flight restrictions, etc.) being disregarded. Nevertheless, the in-situ measurements for the selected day allowed a good estimate of the emission rate when the extrapolation ~~to the upper limit of the mixed~~ up to the limiting stable layer was applied~~(upper end of the error bar in Fig.16)~~.

Both remote sensing point source inversion methods are able to quantify the emissions within the error bars – ~~10and~~ about 15% for the ~~mass balance and~~ plume inversion approach ~~, respectively – assuming that the emissions derived from energy production which have been used for comparison are accurate within a few percent. The uncertainty for the mass balance result is lower because there are fewer input parametersthat affect the overall result~~and about 13% for the mass balance method referring to the combination of the three power plants. The mass balance approach requires less parameters. It is, for example, not essential to know the exact source location and dimensions ~~for the mass balance approach~~ which is an advantage for surveying unknown sources with a non-imaging instrument like MAMAP. ~~An imaging instrument with sufficient spatial resolution will be able to determine the source location from the data directly~~However, the mass balance approach showed a lower precision when only few flight legs per source are available, in particular close to the source and when differences between inversion results are interpreted. To mitigate this effect which is likely based on atmospheric instationarity it is of advantage to gather measurements on multiple flight legs. This results in a higher precision and an improved error estimate.

One critical external input parameter for the analysis of the remote sensing data is wind information, which in this work was derived using model and in-situ data. While the wind direction can be fitted to the data directly, this is not possible for the wind speed which scales linearly with the emission rate.

The in-situ inversion proves to be accurate for power plant Niederaußem where ~~a~~ an almost complete sampling of the plume was possible. Further away from the sources, capturing the complete vertical plume extent in the higher reaching convective boundary layer was not possible due to airspace restrictions. ~~Also the fact, that the background concentrations were derived from concentration minima in this widely contaminated area might have led to a negative bias for the estimated emissions.~~

While the individual results for remote sensing and in-situ yield very similar results provided sufficient sampling, a joint inversion approach may complement the individual methods also when there is no complete plume coverage. The accuracy of in-situ error estimates for cases with better coverage in less restricted places is better than 15% (reference case in Table 3).

The methods presented here are demonstrated for CO_2 emissions from point sources, however, they are directly applicable in the same way to other ~~largely chemically inert gaseous compunds~~ gaseous compounds that disperse in the atmosphere and that have a lifetime longer than between emission and measurement, such as, for example, CH_4 ~~that is also~~ which can also be derived from MAMAP remote sensing observations.

7 ~~Wall alignment for the in-situ flux estimate~~

~~Discussing three possible problems assuming a misalignement of the wall (see~~ This case study illustrates the advantages and disadvantages of the used methods. The remote sensing approach needs clear sky conditions but 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. ~~**??**), a widening of the plume within the wall, or a plume leaving the wall in another cell than it has entered. A_1 is the perfectly aligned wall element for the given wind vector v_w, with no cross wind component, where the plume crosses with a width of w_1. A_2 is misaligned by angle α with a cross wind component v_c, and A_3 is a wall element further downstream, where the plume might be wider (w_3). It is obvious from the graphics that $v_p = v_w \cos(\alpha)$, and $w_1 = w_2 \cos(\alpha)$, i.e.$w_2 = w_1 / \cos(\alpha)$. Since now $v_w \cdot w_1$ and $v_p \cdot w_2$ are equal,~~ 15. 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 ~~the same total mass transport across the wall (air or trace gas), even when the flux (mass per area and time) is reduced when misaligned. When the plume widens, the concentration is diluted accordingly, i.e.the total mass transport remains the same (also in two dimensions) as long as the flow (not the plume!) has no convergence or divergence within the wall. The interesting case is when a plume does not leave the wall in the same grid cell than it has entered it. In this case , both grid cells are associated with the same flux, i.e. the mass transport is doubled. This is also the case when a plume that is smaller than one grid cell is found in both cells (splitted) . Conclusions: (i) the alignment with the mean wind should be adjusted because of this third effect (this has absolute priority, i. e. the wall does not need to be aligned with the flight track, which was - in some cases - due to airspace restrictions - not crossing the plume perpendicularly); (ii) the size of the grid cells must not be too large to ensure correct determination of the background concentrations; (iii) as stated elsewhere, the variation of the grid size is a measure for the sensitivity .~~ generally the need for external 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 0.4 $\mathrm{m\,s^{-1}}$)

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 ($0.7\,\mathrm{m\,s^{-1}}$). 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.

~~Figure illustrating misalignment of the wall for the in-situ flux estimate.~~ 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.

Acknowledgements. The measurement campaign C-MAPExp was funded by the European Space Agency (ESA). The MAMAP activities are funded in parts by the University of Bremen and the State of Bremen. Wind data from the COSMO-DE model were obtained from the German Weather Service (DWD). The reference values for CO_2 emission rates of the power plants were derived from power generation values of each power plant location kindly provided by RWE AG, Essen, Germany. We acknowledge Armin Jordan from the Max Planck Institute for Biogeochemistry in Jena, Germany, for the analysis of airborne grab samples at the BGC GasLab.

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.

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications, Atmos. Meas. Tech., 3, 781–811, doi:10.5194/amt-3-781-2010, 2010.

Bovensmann, H., Krings, T., Gerilowski, K., Neininger, B., Ruhtz, T., and Lindemann, C.: C-MAPExp Final Report - "Scientific and Technical Assistance for the Deployment of a flexible airborne spectrometer system during C-MAPExp", ESA Study, 2014.

Buchwitz, M., Rozanov, V. V., and Burrows, J. P.: A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH_4, CO, CO_2, H_2O, and N_2O total column amounts from SCIAMACHY Envisat-1 nadir radiances, J. Geophys. Res., 105, 15 231–15 245, 2000.

Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V., Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO_2 and CH_4 retrieval errors by error parameterization, Atmospheric Measurement Techniques, 6, 3477–3500, doi:10.5194/amt-6-3477-2013, http://www.atmos-meas-tech.net/6/3477/2013/, 2013.

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

Caulton, D. R., Shepson, P. B., Santoro, R. L., Sparks, J. P., Howarth, R. W., Ingraffea, A. R., Cambaliza, M. O. L., Sweeney, C., Karion, A., Davis, K. J., Stirm, B. H., Montzka, S. A., and Miller, B. R.: Toward a better understanding and quantification of methane emissions from shale gas development, Proceedings of the National Academy of Sciences, 111, 6237–6242, doi:10.1073/pnas.1316546111, http://www.pnas.org/content/111/17/6237.abstract, 2014.

Ciais, P., Crisp, D., van der Gron, H. D., Engelen, R., Janssens-Maenhout, G., Heimann, M., Rayner, P., and Scholze, M.: Towards a European Opterational Observing System to Monitor Fossil CO_2 emissions, Final Report from the expert group, European Comission, 2015.

Doms, G.: A Description of the Nonhydrostatic Regional COSMO-Model, Deutscher Wetterdienst, Technical Report (http://www.cosmo-model.org/), last access: February 2015, 2011.

European Commission: Decision 2007/589/EC, establishing guidelines for the monitoring and reporting of greenhouse gas emissions pursuant to Directive 2003/87/EC of the European Parliament and of the Council, Official Journal of the European Union, L 229, 1–85, 2007.

Foken, T., Aubinet, M., Finnigan, J. J., Leclerc, M. Y., Mauder, M., and Paw U, K. T.: Results Of A Panel Discussion About The Energy Balance Closure Correction For Trace Gases, Bull. Amer. Meteor. Soc., 92, ES13–ES18, doi:10.1175/2011BAMS3130.1, 2009.

Frankenberg, C., Meirink, J. F., van Weele, M., Platt, U., and Wagner, T.: Assessing Methane Emissions from Global Space-Borne Observations, Science, 308, 1010–1014, doi:10.1126/science.1106644, 2005.

Frankenberg, C., Thorpe, A. K., Thompson, D. R., Hulley, G., Kort, E. A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B. D., Aubrey, A. D., Hook, S., and Green, R. O.: Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region, Proceedings of the National Academy of Sciences, 113, 9734–9739, doi:10.1073/pnas.1605617113, http://www.pnas.org/content/113/35/9734.abstract, 2016.

Gerilowski, K., Tretner, A., Krings, T., Buchwitz, M., Bertagnolio, P. P., Belemezov, F., Erzinger, J., Burrows, J. P., and Bovensmann, H.: MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: Instrument description and performance analysis, Atmos. Meas. Tech., 4, 215–243, doi:10.5194/amt-4-215-2011, 2011.

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.

Hiller, R. V., Neininger, B., Brunner, D., Gerbig, C., Bretscher, D., Künzle, T., Buchmann, N., and Eugster, W.: Aircraft-based CH_4 flux estimates for validation of emissions from an agriculturally dominated area in Switzerland, J. Geophys. Res, 119, 4874–4887, doi:10.1002/2013JD020918, 2014.

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, 10.1002/grl.50811, 2013.

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, Atmospheric Measurement Techniques, 10, 3429–3452, doi:10.5194/amt-10-3429-2017, https://www.atmos-meas-tech.net/10/3429/2017/, 2017.

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.

Lavoie, T. N., Shepson, P. B., Cambaliza, M. O. L., Stirm, B. H., Karion, A., Sweeney, C., Yacovitch, T. I., Herndon, S. C., Lan, X., and Lyon, D.: Aircraft-Based Measurements of Point Source Methane Emissions in the Barnett Shale Basin, Environ. Sci. Technol., 49, 7904–7913, doi:10.1021/acs.est.5b00410, 2015.

Martin, D. O.: The Change of Concentration Standard Deviations with Distance, J. Air Poll. Control Assoc., 26, 145–147, doi:10.1080/00022470.1976.10470238, 1976.

Masters, G. M.: Introduction to Environmental Engineering and Science, Prentice-Hall, Inc., 2nd edn., 1998.

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

NRC: National Research Council (NRC) – Committee on Methods for Estimating Greenhouse Gas Emissions, Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements, ISBN 0-309-15212-7, available from http://www.nap.edu/catalog/12883.html, last access: March, 2016, 2010.

Reuter, M., Buchwitz, M., Schneising, O., Hase, F., Heymann, J., Guerlet, S., Cogan, A. J., Bovensmann, H., and Burrows, J. P.: A simple empirical model estimating atmospheric CO_2 background concentrations, Atmospheric Measurement Techniques, 5, 1349–1357, doi:10.5194/amt-5-1349-2012, http://www.atmos-meas-tech.net/5/1349/2012/, 2012.

Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Benner, D. C., Bernath, P., Birk, M., Bizzocchi, L., Boudon, V., Brown, L., Campargue, A., Chance, K., Cohen, E., Coudert, L., Devi, V., Drouin, B., Fayt, A., Flaud, J.-M., Gamache, R., Harrison, J., Hartmann, J.-M., Hill, C., Hodges, J., Jacquemart, D., Jolly, A., Lamouroux, J., Roy, R. L., Li, G., Long, D., Lyulin, O., Mackie, C., Massie, S., Mikhailenko, S., Müller, H., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E., Richard, C., Smith, M., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G., Tyuterev, V., and Wagner, G.: The {HITRAN2012} molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 4–50, doi:https://doi.org/10.1016/j.jqsrt.2013.07.002, http://www.sciencedirect.com/science/article/pii/S0022407313002859, 2013.

Schepers, D., Guerlet, S., and Butz, A.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10 307, doi:10.1029/2012JD017549, 2012.

Thompson, D. R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R. O., Kratwurst, S., Krings, T., Luna, B., and Thorpe, A. K.: Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane, Atmos. Meas. Tech., 8, 4383–4397, doi:10.5194/amt-8-4383-2015, 2015.

Thorpe, A. K., Frankenberg, C., and Roberts, D. A.: Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS, Atmospheric Measurement Techniques, 7, 491–506, doi:10.5194/amt-7-491-2014, 2014.

Tratt, D. M., Buckland, K. N., Hall, J. L., Johnson, P. D., Keim, E. R., Leifer, I., Westberg, K., and Young, S. J.: Airborne visualization and quantification of discrete methane sources in the environment, Remote Sensing of Environment, 154, 74–88, doi:10.1016/j.rse.2014.08.011, 2014.

Turner, D. B.: Workbook of Atmospheric Dispersion Estimates, U.S. Department of Health, Education, and Welfare, 95 pp., 1970.

U.S. Committee on Extension to the Standard Atmosphere: U.S. Standard Atmosphere, 1976, National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, Washington D.C., 241 pp., 1976.

Young, S.: 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.