

Tracking city CO₂ emissions from space using a high resolution inverse modeling approach: A case study for Berlin, Germany

Authors' responses to reviewers' comments:

We would like to thank both referees for their careful reviewing and constructive comments, recommendations, and suggestions. Our responses to these comments are as follows:

[RC]: Reviewer's comment

[AR]: Authors' response

[ME]: Manuscript edits & modification

Reviewer #1

[RC] "One criticism is that the accuracy needs in terms of city emissions are not discussed. In the introduction, it is explained that an independent evaluation of the emissions is needed to verify the impact of mitigation efforts. If this is the primary objective of the mission, I would think that an accuracy of a few percent is needed. The introduction also mentions the objective of inventory verification. However, the accuracy that is required for such objective is not mentioned. Finally, the introduction mentions the potential of independent measurements of CO₂ for emission trading. However, I believe that emission trading is at the scale of industries, and not at the scale of the cities. I thus recommend a better description of the needs with a quantification of the accuracy requirement. Clearly, this question has a strong impact on the results. Indeed, the abstract conclude by saying that "CarbonSat is well suited to obtain city scale CO₂ emissions". Since there was no requirement set, one cannot draw such conclusion. "

[AR] We agree that the paper can be improved in this respect and we will modify the Introduction and the Abstract to consider this comment. The text will be modified as follows:

[ME] P2L13 in ACPD: "The reporting of the emissions of CO₂ is currently determined by national and regional agreements and legislation. This is an evolving topic for policy makers. For example, there exists an emission inventory which accounts for total annual U.S. emissions between 1990 and 2014 (EPA, 2016). In the European Union, the monitoring and reporting of greenhouse gas emissions are performed and regulated under the Commission Regulation (EU) No 601/2012 (European Commission, 2012). Similarly, the UK Government has announced, under the Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013, that companies are required to report their annual greenhouse gas emissions in their directors' report (see http://www.legislation.gov.uk/uksi/2013/1970/pdfs/ukxi_20131970_en.pdf). There is also a guideline for national greenhouse inventories prepared by a task force of the IPCC (IPCC 2006). Following the agreement of the UNFCCC COP21 in Paris 2015, it is likely that new guidelines for reporting the emissions of greenhouse gases will be required.

The uncertainties, i.e. the sum of systematic and stochastic error, in the national average of annual fossil fuel CO₂ emissions from the United states is estimated to be 2 to 5 % (EPA, 2016). The corresponding values for countries without well-developed energy sector statistics are even higher, giving rise to uncertainties of about 10 to 20 % at the national level (Gregg et al., 2008). When disaggregating these national emissions at fine scales (e.g. city scale) based

on conventional accounting methods, the associated uncertainties are expected to be significantly higher compared to those of national averages (Oda and Maksyutov, 2011). Hence reliable emission estimates are not often available at a scale relevant for urban emissions and the associated uncertainties. “

P2L20 in ACPD: “In order to assess accurately the contribution of a city or other emission hot spot to the CO₂ or other GHG emission, accurate knowledge of the surface fluxes at high spatial and temporal resolutions are needed. Ideally the accuracy of the estimated flux needs to be high for unambiguous attribution of source strength. The uncertainty of these estimations is required to be reduced to the extent that is feasible. In ESA (2015) it is noted (see their Sect. 4.1.2) that accuracies better than 10% would be useful for providing important additional information for cities where inventories exist, and accuracies better than 20% would contribute knowledge for cities where inventories do not exist.”

“The results show that these potential space-based top-down flux estimates have high accuracy; hence this study contributes to the definition of achievable targets for emission fluxes at the city scale.”

Abstract

P1L33 in ACPD: “Overall, we conclude that a satellite mission such as CarbonSat has high potential to obtain city-scale CO₂ emissions as needed to enhance our current understanding of anthropogenic carbon fluxes, and that CarbonSat-like satellites should be an important component of a future global carbon emission monitoring system.”

[RC] I have been confused with the size of the state vector λ . On the one hand, I understood (p7, line 20 and below) that only the scaling to prior emissions of the Berlin region was retrieved (the spatial and temporal variations of the emissions are assumed) so that there is a single element in the state vector. On the other hand, on page 8 (around 5), it is stated that the prior uncertainty is a matrix with no correlation, which clearly indicates that there are several elements in the state vector. Please clarify.

[AR] The text will be modified to clarify this.

[ME] P7L21 in ACPD: “The other element is a constant, i.e. $\lambda_0 = 0$, for the entire scene per overpass to account for variations of the background XCO₂ (see Eq. 4) and to treat the background variations independently of the city emissions as done in Buchwitz et al. (2013b).”

[RC] Page 9 around line 30. I could not understand why a 500 km swath instrument leads to 25 valid observations during the year while the same with a swath width of 240 km leads to much more than half of this number. I would have expected that, as the swath is reduced by a factor of slightly more than 2, the number of valid observation be reduced by a factor of significantly more than 2. Please discuss. Is this number typical of what can expect for cities with similar cloud cover as Berlin, or is the CarbonSat orbit centered over Berlin which makes it a favourable case ?

[AR] It is mentioned in Page 7 line 16 in ACPD that there are 41 days of potential overpasses over Berlin for the 500 km swath width, based on the quality filtering scheme as described in Buchwitz et al. (2013a). However, as stated in page 9 line 27 in ACPD, we have applied an additional quality filter based on the posteriori random error of the retrieved emission, i.e., results are not shown for the days where retrieved emission random error exceeds 25%. This resulted in the number of valid days

reducing to 25 for the 500 km swath width. The additional quality filter did not reduce much the valid observations for the 240 km swath width. The text will be modified as follows:

[ME] P9L30 in ACPD: “Applying this additional quality criterion has further reduced the number of potential overpasses for the 500 km swath width.”

[AR] This number (“N”) is something typical we would expect for cities with similar cloud and SZA.

[ME] P9L31 in ACPD “The value obtained here for “N” useful overpasses is expected to be typical for other cities with similar cloud coverage and latitude.”

[RC] “A quite high scaling factor” (between CO₂ and aerosol optical depth). Please explain the reasoning to state that it is a high scaling factor. Indeed, eventually a CarbonSat-like mission may be used to monitor cities that are not as “clean” as Berlin. For such cities, the scaling factor chosen may be an underestimate.

Will be modified in the text as follows:

[ME] P9L34 in ACPD “To quantify the urban aerosol enhancement over a region around two power plants in Germany and to study their impact on emission estimates, Krings et al. (2011) followed the above criteria and used an AOD scaling factor of 0.05 per 1% (4 ppm) of local ΔXCO_2 . As compared to their study, we have used a much higher scaling factor of 0.2, i.e., the AOD change, ΔAOD at 550 nm is 0.2 per 4 ppm of local anthropogenic ΔXCO_2 .”

[RC] I was surprised by the discussion of the “clean pixel method”. In the present state, it is very hard to understand and it comes at odds with the rest of the paper. I strongly suggest to remove this section.

[AR] The discussion about the clear-pixel method has been removed.

[RC] 14C is mentioned. I do not think that anyone believe that 14C can be measured from space. The present paper is about spaceborne observation and I think it is misleading to mention 14C here. As for the other tracers (CO, NO_x), the authors do know that, when adding this source of information, one also adds an unknown variable (relative fraction of emissions). Thus, I think it is misleading to suggest that the concomitant measurement of these gases would allow a distinction of the biogenic and anthropogenic contributions.

[AR] The mention of $\delta^{14}C$ has been removed as it cannot be measured from space with available technology. However, we disagree with the following statement about the additional tracers. There are a number of studies (including satellite-based) which showed the potential of using multiple species as tracers for anthropogenic CO₂ emissions from fossil fuel combustion. We see the importance of potential synergies to utilize these measurements to separate the anthropogenic and biospheric parts of the signals. The text will be modified to include the citations of these previous studies.

[ME] P13L5 in ACPD “To assess the relative contribution of biogenic and anthropogenic sources, one can utilize additional co-emitted tracers such as CO and NO_x (Newman et al. 2013; Silva et al., 2013; Berezin et al., 2013 and Reuter et al. 2014). In the time frame of a

potential CarbonSat mission, Sentinel-5 will be providing data on CO and tropospheric NO₂ (Ingmann et al., 2012), which when combined with CarbonSat data is expected to provide information for the attribution of air masses originating from fossil fuel combustion. Depending on the extent of the variability and the possible uncertainties, we can also rely on the biospheric and global model simulations to differentiate different source-sink contributions.”

[RC] I could not understand the argument in the sentence “By showing that the systemic error of the retrieved fluxes. . .”. Please rephrase

[AR] done

[ME] P15L22 in ACPD “By showing that the systemic error of the retrieved fluxes is lower than the difference between the prior fluxes and the true fluxes in most cases, the results from the inversion experiment build confidence in our uncertainty estimations and ensure that the optimization is done correctly.”

[RC] Finally, I have a recommendation for discussion: The simulations are made at 10 km resolution and the authors do not mention a significant loss of information from the original 2 km of the CarbonSat instrument. One should then wonder what is the added value of the high spatial resolution of CarbonSat. It seems that 10 km resolution is good enough to observe the plume from the Berlin city

[AR] The point is already discussed in the manuscript. Please see page 8, line 28 and page 9, line 14 in ACPD. However, in order to make this point into the Discussion Sect., we modified the manuscript as follows:

[ME] P13L43 in ACPD “When using observations at CarbonSat’s 2 km spatial resolution, as mentioned in Sec. 4.1, it is likely that the magnitude and variability of local anthropogenic XCO₂ enhancement would be higher than our estimation that is based on simulations at 10 km spatial resolution. One of the main advantages of CarbonSat’s resolution is its ability to provide a large number of cloud-free observations, and this study identified the potential observations over Berlin by utilizing CarbonSat’s 2 km spatial resolution. “

[RC] Figure 2 is mentioned but not discussed. It does not bring anything to the paper and I thus strongly suggest to remove it.

[AR] done

[RC] In Figure 3, it seems that the original data (10 km resolution) went through spatial smoothing. I would like to see the pixels. Also, a zoom over the Berlin region would be appropriate

[AR] done

[RC] Figure 4 : Please use Y axis that start at zero. The current presentation is somewhat misleading. The difference between IER and EDGAR are surprisingly large. I wonder whether there are arguments to favor one versus the other.

[AR] The y-axis range has been modified. Our analysis shows large difference in emission intensities between IER and EDGAR over Berlin; however, it is difficult to

say which one is more appropriate than the other. We will add following line in revised version of the manuscript:

[ME] P5L15 in ACPD “However, based on available sources of information it is difficult to conclude which inventory is more accurate.”

[RC] Figure 6: There are features I do not understand: Around day 75, two successive prior emission values show differences by a factor of two. Based on Figure 4, I cannot understand how the weekly, seasonal or daily cycles can explain a difference by a factor of two (assuming the observation is around 11 when there are little hourly variations). Please investigate

[AR] Those two successive prior emissions (days 75 and 80) correspond to 15.03.2008 (Saturday) and 20.03.2008 (Thursday). Since emissions are generally low during weekend, we can expect an emission difference (about a factor of 1.3 to 1.5) between day 75 and 80. However, we found some issues in the plotting routine of the code used for creating this figure – (1) a shift in the weekdays was not applied to IER 2000 dataset to utilize it for the year 2008 and to preserve the temporal pattern differences between weekdays and weekends (please see Sec. 2.1.1), and (2) slightly different spatial area was used for calculating total EDGAR fluxes over the target region. Fig. 6 in ACPD (will be Fig. 5 in the revised version of the manuscript) will be modified. Although this plotting routine is independent of WRF simulations (it was done separately and had been taken care of) or inversion routines (note that scaling factors are used), this bug has very minor effect on our results of REs and SEs where prior fluxes are used for the estimation. Hence Fig. 7, Fig(s). 10 to 13 in ACPD (will be Fig. 6, Fig(s). 9 to 12 in the revised manuscript) and Table 1 will be modified. Revised figures are shown at the bottom of this document:

In addition to this, some more details will be given in the revised manuscript as follows:

[ME] P9L4 in ACPD: “Note that those fluxes "seen" by CarbonSat can vary significantly from one overpass to the next, as the temporal variations show a strong diurnal cycle (see also Fig. 3), and the time elapsed to transport the plume to where it is observed by CarbonSat changes with wind speed.”

Reviewer #2

[RC] P3 L10: "The goal swath width is 500 km, but a smaller swath width will likely be implemented to limit cost (ESA, 2015). " Here and elsewhere in the text, check the consistency with the fact that CarbonSat was not selected.

[AR] Done. Will be modified as follows.

[ME] P3L10 in ACPD: “The goal swath width for the proposed CarbonSat mission was 500 km, but smaller swath widths were also considered to limit cost (ESA, 2015).”

[RC] P3 L17: "...Buchwitz et al. (2013a)... " I could not find Buchwitz et al. (2013b) in references (which is also cited further in the manuscript). I supposed you mean the paper: "Carbon Monitoring Satellite (CarbonSat): assessment of scattering related atmospheric CO₂ and CH₄ retrieval errors and first results on implications for inferring city CO₂ emissions" Buchwitz et al. AMTD, 2013 Please, confirm.

[AR] Thank you pointing this out. Buchwitz et al. (2013a) is the AMT paper and Buchwitz et al. (2013b) is the one which is given above. The reference section has been corrected accordingly.

[RC] P4 L10: "41 vertical levels" Please, indicate the model top at hPa

[AR] Done. Will be modified as follows.

[ME] P4L10 in ACPD: "and the model top is 1.0 hPa"

[RC] P4 L26: "An overview of the flux optimization is shown in Fig. 2." I think the reference to Pillai et. al. 2012 is enough. As well as for P11 L19-20, "As can be seen in Fig. 2". I suggest to remove Fig. 2

[AR] Done.

[RC] P5 L11: "Figure 4 shows..." I suggest to keep consistency between figures and use a) b) c) etc. for different panels throughout the paper.

[AR] Done.

[RC] As remark, I suggest to add more clarifications in sections 2.2 and 3.2. At this shape it's hard to get into details of the inversion system. It would be useful to add dimensions for every component of the system.....

[AR] Done. Will be modified as follows.

[ME] P7L26 in ACPD: "... and the dimension of K is $n \times m$ where n corresponds to the numbers of elements in the state vector and m is the number of XCO_2 observations."

[RC] P6 L16: "Eq. (4)" -> "Eq. (6)"

[AR] Done. Thank you.

[RC] P8 L8: "Any error correlations are neglected, hence Sprior is set to be a diagonal matrix" - is the measurement error covariance matrix also diagonal? If so, add few words about this assumption, especially for CarbonSat-like XCO_2 observations.

[AR] Will be modified. Please see our response to Referee #1

[RC] P10 L6: "...typically differs..." - Typically for Berlin region or in general?

[AR] in general

[RC] P10 L16-17: "In general, we find that the two different swath widths have a negligible impact on the daily SE of the retrieved emissions" - Please, rephrase this sentence as in conclusion section.

[AR] Will be modified as follows:

[ME] P10L17 in ACPD: "...although decreasing the swath width reduces the "N" useful

overpasses.”

[RC] P12 L25-26: "Furthermore, the systematic errors of the retrieved emission fluxes for both swath widths are found to be lower than the systematic error of the prior fluxes (estimated based on “true” fluxes) except for a very few cases,...” Please, rephrase this sentence.

[AR] Will be modified as follows:

[ME] P12L25 in ACPD: “... swath widths are found to be lower than the difference between the prior fluxes and the “true” fluxes except for a very few cases...”

[RC] P12 L12: "...in the target region is notably different." - Here need to add ref. to the figure 6 in the end of the sentence. Otherwise this figure is not mentioned in the paper at all.

[AR] Will be modified as follows: (please note that Fig. 6 will be Fig.5 in the revised version of the manuscript)

[ME] P9L4 in ACPD: “Figure 5 shows an overview of prior fluxes used for these inversions.”
P12L12 in ACPD: “notably different (see Fig. 5).”

[RC] P15 L22-23: "By showing that the systemic error of the retrieved fluxes is lower than that of the prior fluxes (estimated based on true fluxes) in most of the cases," – please, consider to rephrase this sentence

[AR] Will be modified. Please see our response to Referee #1.

[RC] Also, as a comment to section 4.3 I think there might be effect of ignoring transport model uncertainty giving less weight to the prior fluxes.

[AR] Yes, we do agree that there will be an impact of transport uncertainty, but the scope of this study is to estimate the retrieved flux uncertainties that are caused only by CarbonSat’s related errors. This is already mentioned in the manuscript. Please see page 14, line 1 in ACPD.

As for section 5 - Discussion, I agree with Referee #1 about introduction and discussion of the "clean pixel method" here. From my point of view it disturbs the logic of the paper and I suggest to remove this paragraph.

[AR] Removed. Please see our response to Referee #1.

Revised Table 1

Table 1: Overview of different scenarios, SCE, which are used to investigate the systematic errors of the retrieved emissions. The absolute mean and standard deviations are estimated for two swath widths (SW-500: 500 km and SW-240: 240 km) for all “N” useful overpasses and are expressed in both $\text{MtCO}_2 \text{ yr}^{-1}$ and in %. Err-L, Err-H, and Err-B indicate errors attributed to CarbonSat measurement, high-resolution aerosol related errors, and biogenic modeling errors respectively. Err-Emi indicates whether inversion experiment uses different prior emission fluxes (see Sec. 4.3)

| SCE | Err-L | Err-H | Err-B | Err-Emi | Prior flux | True flux | SE (SW-500) (mean \pm std) % | SE (SW-500) (mean \pm std) $\text{MtCO}_2 \text{ yr}^{-1}$ | SE (SW-240) (mean \pm std) % | SE (SW-240) (mean \pm std) $\text{MtCO}_2 \text{ yr}^{-1}$ |
|-----|-------|-------|-------|---------|------------|-----------|--------------------------------------|--|--------------------------------------|--|
| S01 | ✓ | | | | EDGAR | EDGAR | -5.3 ± 6.1 | -2.5 ± 2.8 | -6.1 ± 5.7 | -2.8 ± 2.6 |
| S02 | | ✓ | | | EDGAR | EDGAR | -7.5 ± 3.4 | -3.6 ± 2.1 | -7.5 ± 2.9 | -3.1 ± 1.6 |
| S03 | | | ✓ | | EDGAR | EDGAR | 18.5 ± 23.3 | 7.5 ± 9.5 | 20.5 ± 23.7 | 8.2 ± 9.4 |
| S04 | ✓ | ✓ | | | EDGAR | EDGAR | -12.7 ± 7.4 | -6.1 ± 3.8 | -12.8 ± 7.1 | -5.9 ± 3.6 |
| S05 | ✓ | ✓ | ✓ | | EDGAR | EDGAR | 5.8 ± 24.4 | 1.5 ± 10.7 | 7.7 ± 25.3 | 2.3 ± 10.6 |
| S06 | ✓ | | ✓ | | EDGAR | EDGAR | 13.2 ± 23.7 | 5.1 ± 10.0 | 14.4 ± 24.5 | 5.4 ± 10.0 |
| S07 | | | | ✓ | IER | EDGAR | 10.1 ± 15.5 | 4.5 ± 6.9 | 10.1 ± 15.9 | 4.5 ± 7.1 |
| S08 | ✓ | | | ✓ | IER | EDGAR | 6.3 ± 16.2 | 2.8 ± 7.1 | 5.9 ± 18.2 | 2.6 ± 7.9 |
| S09 | ✓ | ✓ | | ✓ | IER | EDGAR | -0.5 ± 16.0 | -0.2 ± 6.9 | -0.9 ± 18.2 | -0.4 ± 7.6 |
| S10 | ✓ | | ✓ | ✓ | IER | EDGAR | 24.3 ± 27.7 | 10.8 ± 11.0 | 24.3 ± 29.1 | 10.8 ± 11.4 |
| S11 | ✓ | ✓ | ✓ | ✓ | IER | EDGAR | 17.6 ± 27.9 | 7.8 ± 11.1 | 17.6 ± 29.6 | 7.8 ± 11.6 |

Revised Figures:

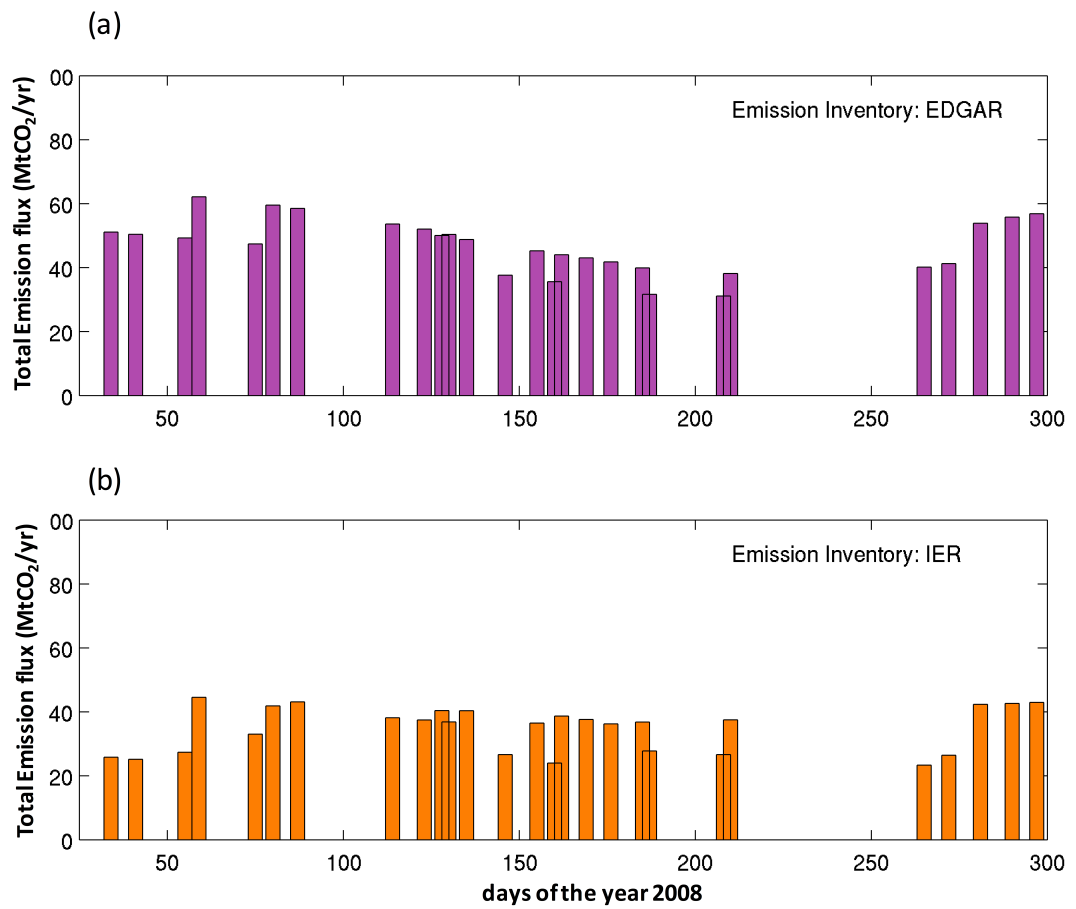


Figure 5: Anthropogenic flux over the target region based on (a) EDGAR inventory, and (b) IER inventory for all of CarbonSat's "useful overpasses" corresponding to 500km swath width for the year 2008.

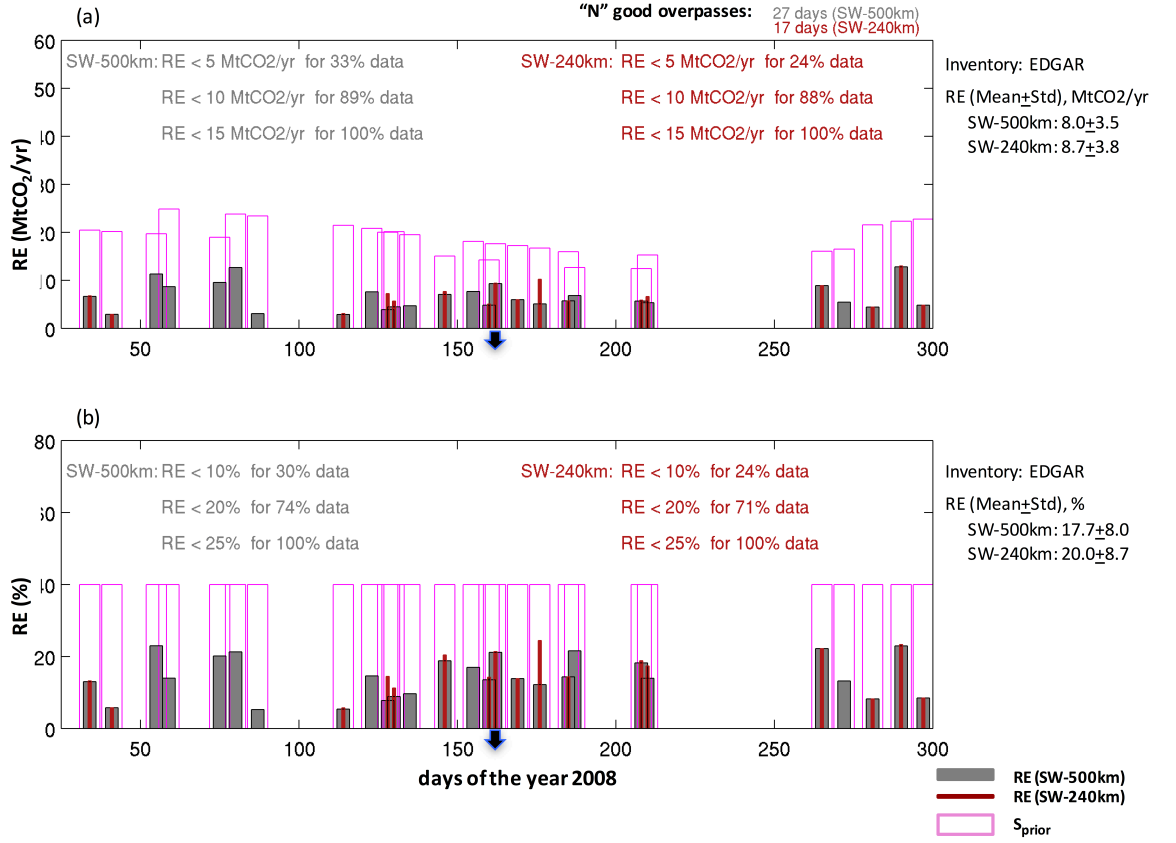


Figure 6: Precision (random errors (RE)) of the retrieved emission fluxes obtained by the inverse optimization using one year of CarbonSat simulated observations. Results of two different swath widths (SW) – 500 km (grey) and 240 km (red) – are shown. S_{prior} values are indicated with magenta bordered bars for visualizing the reduction in uncertainty. The top and bottom panels show RE in MtCO₂ yr⁻¹ and in % respectively. An overview of the statistical distribution of RE, separately for 500 km (grey) and 240 km (red) swath widths, is given inside the panel. The overall mean ± standard deviation is given outside the respective panels. The lower and upper limits of the X-axis (days of the year) is restricted accordingly as there are no good CarbonSat simulated observation during winter months. The arrow marker in the X-axis indicates a particular day (24th June 2008) shown in Figs. 4, 7, and 8.

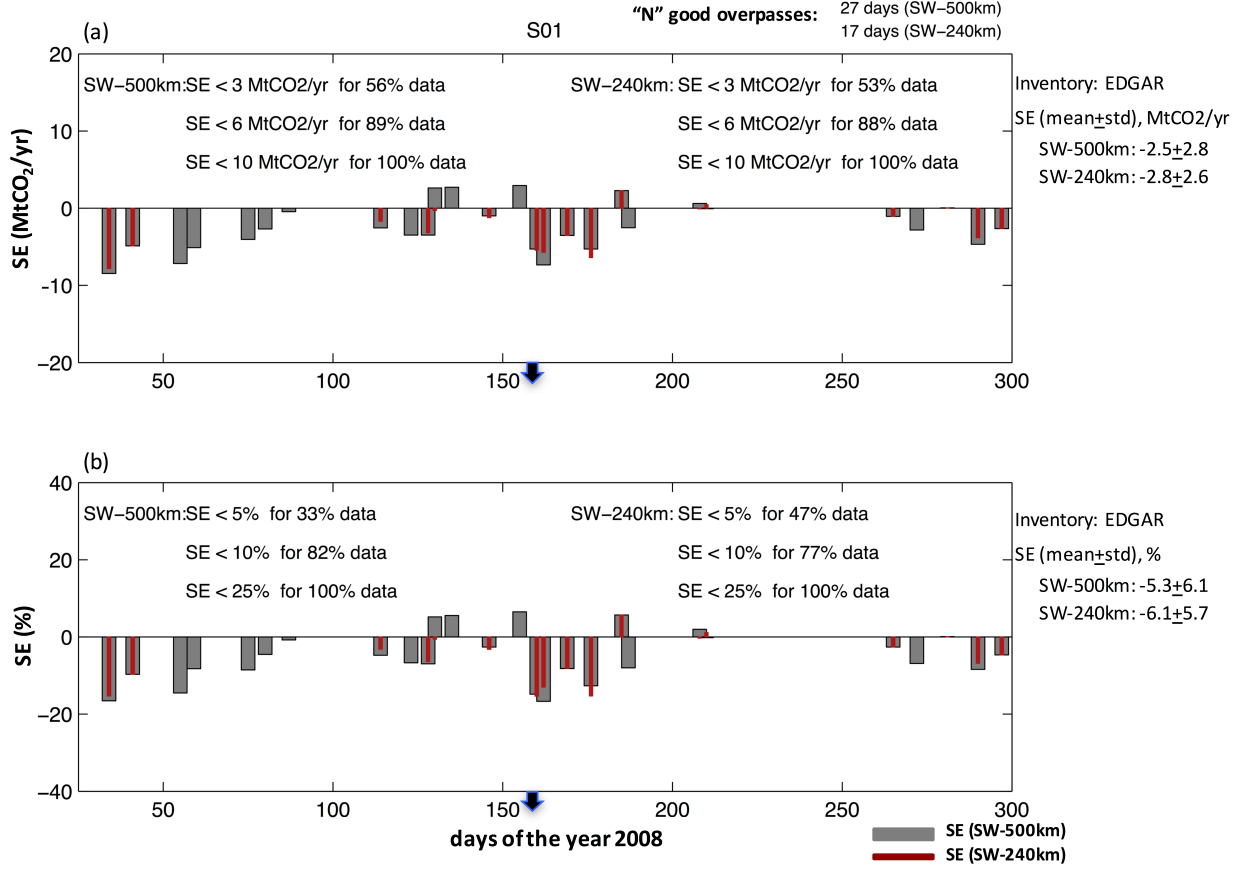


Figure 9: Systematic errors (SE) of the retrieved emission fluxes for S01, obtained by the inverse optimization using one year of CarbonSat simulated observations. Results of two different swath widths (SW) – 500 km (grey) and 240 km (red) – are shown. (a) and (b) show SE in MtCO₂ yr⁻¹ and in % respectively. An overview of the statistical distribution of SE, separately for 500 km (grey) and 240 km (red) swath widths, is given inside the panel. The overall mean +/- standard deviation is given outside the respective panels.

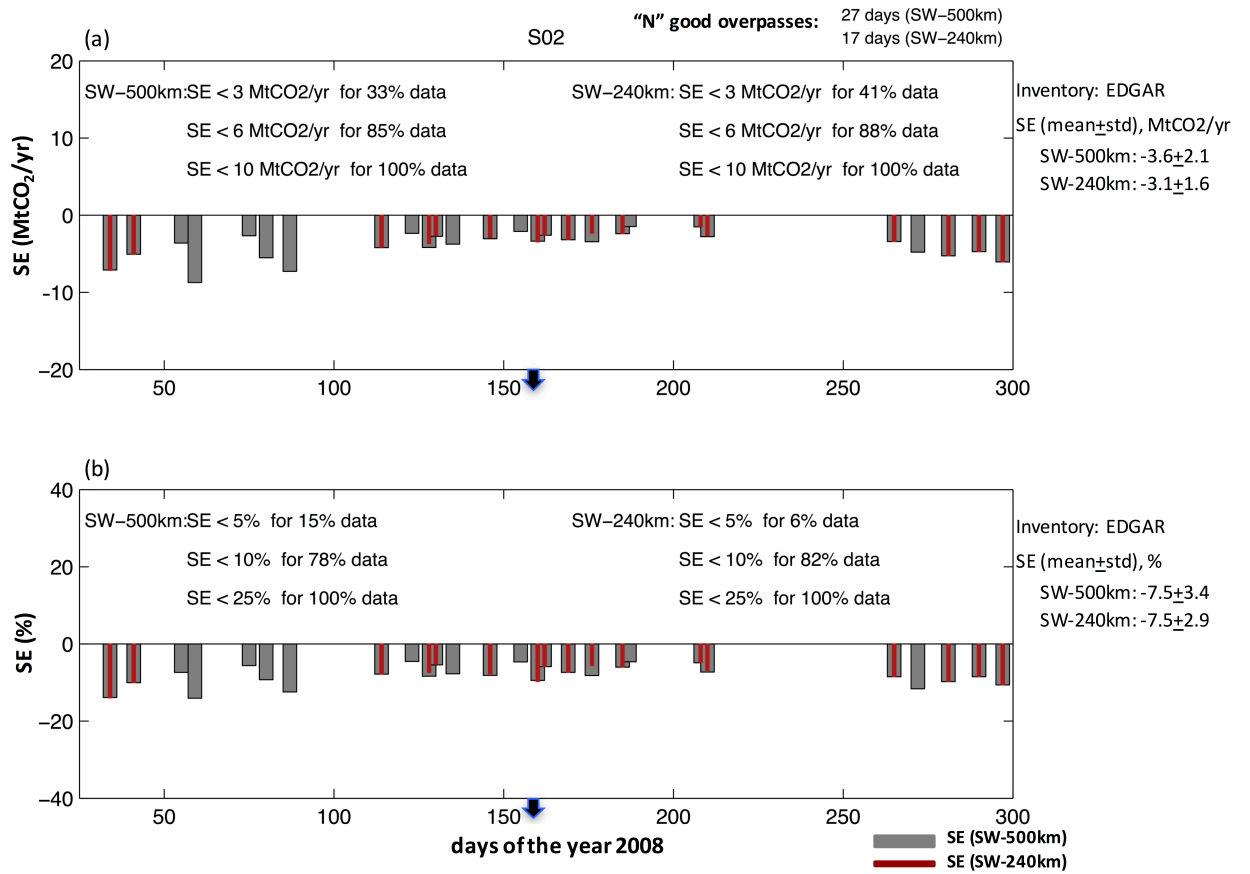


Figure 10: Same as Fig.9, but for S02, quantifying the impact of the worst-case assumption used for aerosol-related biases.

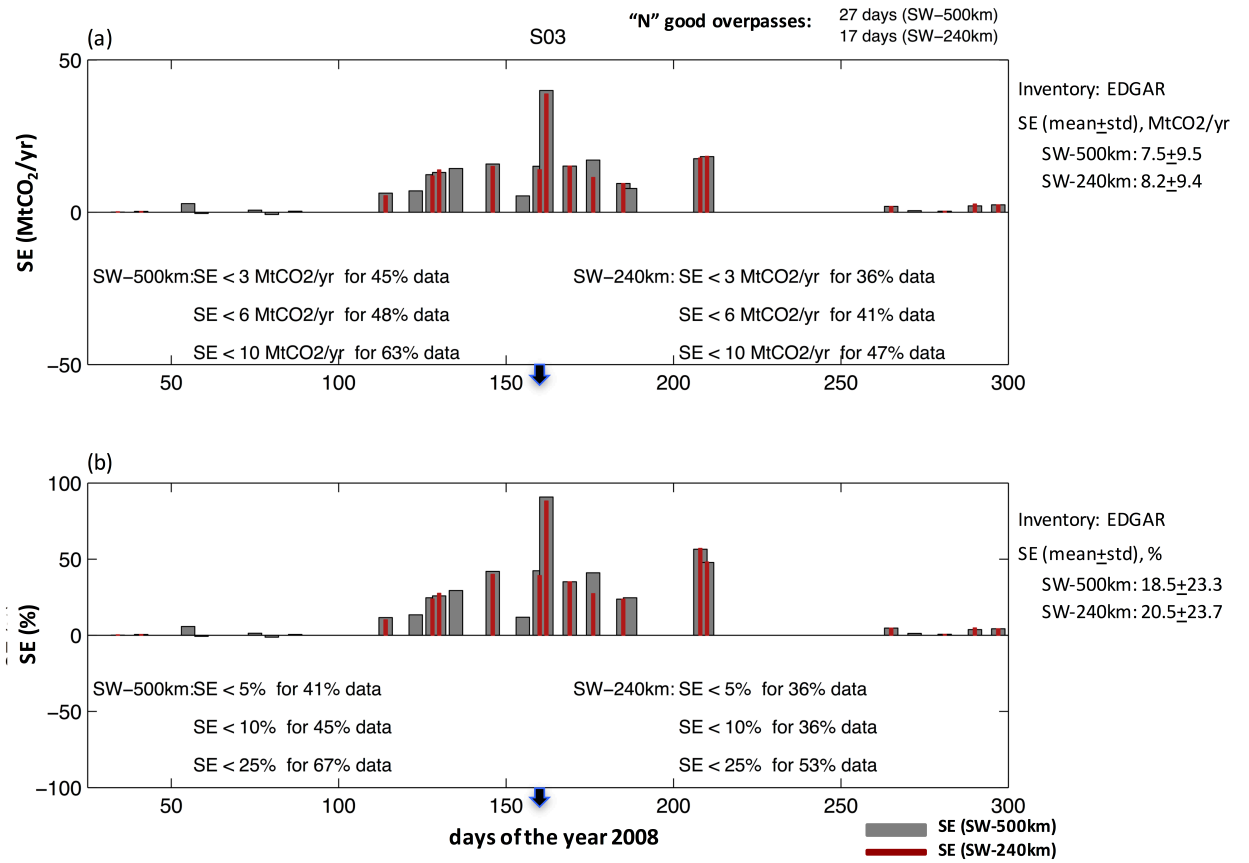


Figure 11: Same as Fig.9, but for S03, quantifying the impact of the worst-case modeling related errors by assuming that biogenic XCO₂ variations cannot be modeled at all.

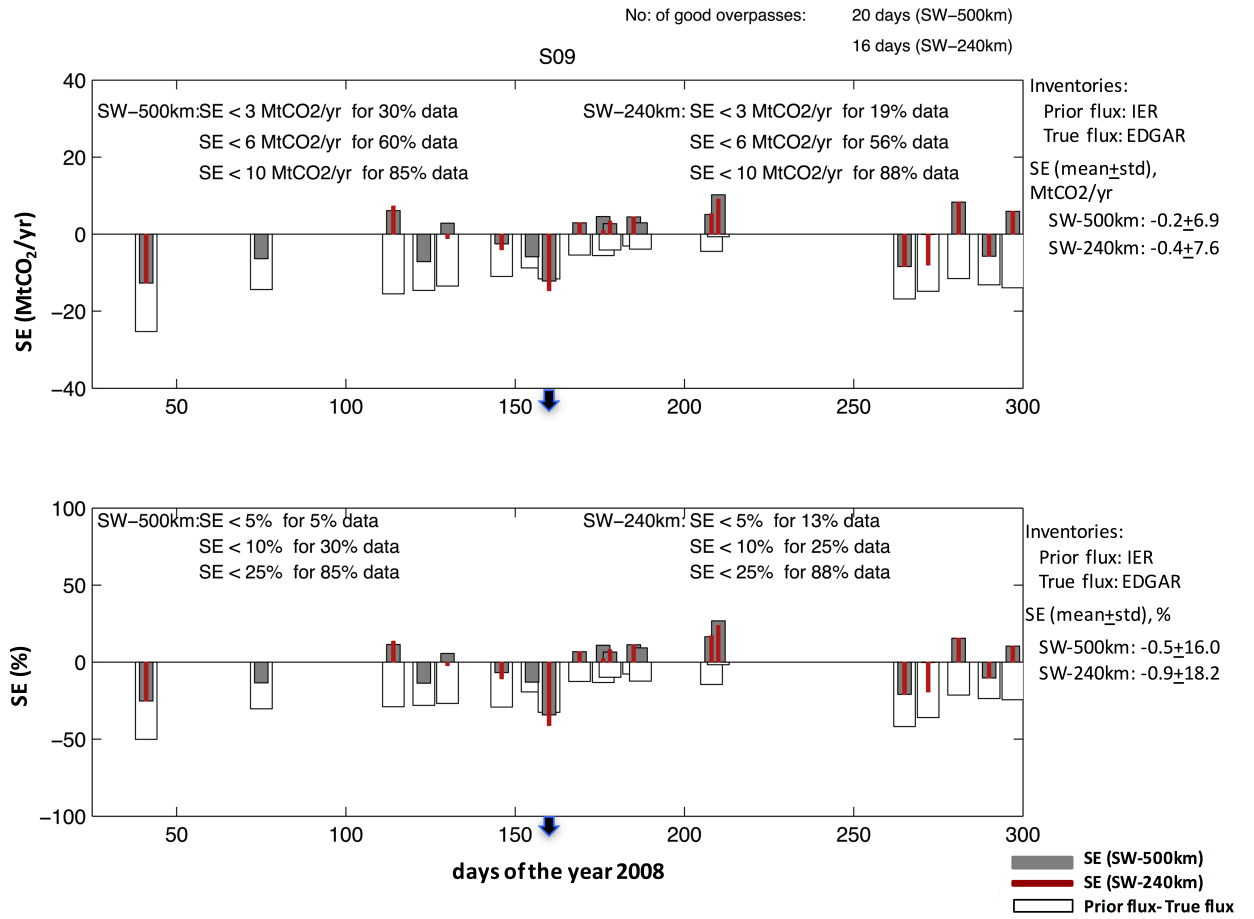


Figure 12: Similar to Fig. 9, but for the inversion experiment S09 using IER (a priori) and EDGAR (true) emission fluxes.

Tracking city CO₂ emissions from space using a high resolution inverse modeling approach: A case study for Berlin, Germany

Dhanyalekshmi Pillai^{1,2*}, Michael Buchwitz¹, Christoph Gerbig², Thomas Koch², Maximilian Reuter¹, Heinrich Bovensmann¹, Julia Marshall², and John P. Burrows¹

¹ Institute of Environmental Physics, University of Bremen, Bremen, Germany

² Max Planck Institute for Biogeochemistry, Jena, Germany

^{2*} now at: Max Planck Institute for Biogeochemistry, Jena, Germany

Correspondence to: Dhanyalekshmi Pillai (kdhanya@bgc-jena.mpg.de)

Abstract. Currently 52% of the world's population resides in urban areas and as a consequence, approximately 70% of fossil fuel emissions of CO₂ arise from cities. This fact in combination with large uncertainties associated with quantifying urban emissions due to lack of appropriate measurements makes it crucial to obtain new measurements useful to identify and quantify urban emissions. This is required, for example, for the assessment of emission mitigation strategies and their effectiveness. Here we investigate the potential of a satellite mission like Carbon Monitoring Satellite (CarbonSat) which was proposed to the European Space Agency (ESA) – to retrieve the city emissions globally, taking into account a realistic description of the expected retrieval errors, the spatiotemporal distribution of CO₂ fluxes, and atmospheric transport. To achieve this, we use (i) a high-resolution modeling framework consisting of the Weather Research Forecasting model with a greenhouse gas module (WRF-GHG), which is used to simulate the atmospheric observations of column averaged CO₂ dry air mole fractions (XCO₂), and (ii) a Bayesian inversion method to derive anthropogenic CO₂ emissions and their errors from the CarbonSat XCO₂ observations. We focus our analysis on Berlin in Germany using CarbonSat's cloud-free overpasses for one reference year. The dense (wide swath) CarbonSat simulated observations with high-spatial resolution (approx. 2 km x 2 km) permits one to map the city CO₂ emission plume with a peak enhancement of typically 0.8-1.35 ppm relative to the background. By performing a Bayesian inversion, it is shown that the random error (RE) of the retrieved Berlin CO₂ emission for a single overpass is typically less than 8 to 10 MtCO₂ yr⁻¹ (about 15 to 20% of the total city emission). The range of systematic errors (SE) of the retrieved fluxes due to various sources of error (measurement, modeling, and inventories) is also quantified. Depending on the assumptions made, the SE is less than about 6 to 10 MtCO₂ yr⁻¹ for most cases. We find that in particular systematic modeling-related errors can be quite high during the summer months due to substantial XCO₂ variations caused by biogenic CO₂ fluxes at and around the target region. When making the extreme worst-case assumption that biospheric XCO₂ variations cannot be modeled at all (which is overly pessimistic), the SE of the retrieved emission is found to be larger than 10 MtCO₂ yr⁻¹ for about half of the sufficiently cloud-free overpasses, and for some of the overpasses we found that SE may even be on the order of magnitude of the anthropogenic emission. This indicates that biogenic XCO₂ variations cannot be neglected but must be considered during forward and/or inverse modeling. Overall, we conclude that a satellite mission such as CarbonSat has high potential to obtain city-scale CO₂ emissions as needed to enhance our current understanding of anthropogenic carbon fluxes, and that CarbonSat-like satellites should be an important component of a future global carbon emission monitoring system.

1. Introduction

One of the main objectives of any climate policy initiative is to limit atmospheric greenhouse gas emissions resulting from anthropogenic activity to a level that minimizes adverse modification of the climate system. An essential component in attaining this goal is the accurate quantification of emissions at national and state levels in order to independently verify the implemented climate change mitigation and adaptation measures. In the context of CO₂, cities are significant contributors of emissions, giving rise to approximately 70 % of the total anthropogenic

emissions (Canadell et al., 2010). However, there exist large uncertainties associated with quantifying urban emissions. This makes it difficult to assess the efficacy of any emission management schemes at urban scales.

While mitigation efforts are being taken in some cities around the globe, they lack objective, observation-based methods to verify their outcomes (Pacala et al., 2010). Some observation-based attempts have been made with a focus on deriving city-scale emissions in a variety of urban environments (Bergeron and Strachan, 2011; Levin et al., 2011; Mays et al., 2009; Wang et al., 2010; Zimnoch et al., 2010). However, none of these approaches is able to account for CO₂ emissions from urban areas with the accuracy required for verification, nor are they easily adaptable to other locations. As a result our current emission estimates are purely based on inventories (bottom-up approach), which have large uncertainties due to many unresolved processes related to spatial and temporal heterogeneity of emission fluxes and local transport phenomena (Amstel et al., 1999; Gregg et al., 2008; Marland, 2008; White et al., 2011). Recent revelations about the inaccuracy of the knowledge of motor vehicle emissions emphasize this point.

The reporting of the emissions of CO₂ is currently determined by national and regional agreements and legislation. This is an evolving topic for policy makers. For example, there exists an emission inventory which accounts for total annual U.S. emissions between 1990 and 2014 (EPA, 2016). In the European Union, the monitoring and reporting of greenhouse gas emissions are performed and regulated under the Commission Regulation (EU) No 601/2012 (European Commission, 2012). Similarly, the UK Government has announced, under the Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013, that companies are required to report their annual greenhouse gas emissions in their directors' report (see http://www.legislation.gov.uk/ukxi/2013/1970/pdfs/ukxi_20131970_en.pdf). There is also a guideline for national greenhouse inventories prepared by a task force of the IPCC (IPCC 2006). Following the agreement of the UNFCCC COP21 in Paris 2015, it is likely that new guidelines for reporting the emissions of greenhouse gases will be required.

The uncertainties, i.e. the sum of systematic and stochastic error, in the national average of annual fossil fuel CO₂ emissions from the United states is estimated to be 2 to 5 % (EPA, 2016). The corresponding values for countries without well-developed energy sector statistics are even higher, giving rise to uncertainties of about 10 to 20 % at the national level (Gregg et al., 2008). When disaggregating these national emissions at fine scales (e.g. city scale) based on conventional accounting methods, the associated uncertainties are expected to be significantly higher compared to those of national averages (Oda and Maksyutov, 2011). Hence reliable emission estimates are not often available at a scale relevant for urban emissions and the associated uncertainties. This is problematic in terms of judging the effectiveness of emission reduction schemes or designing new management strategies for emission trading. Furthermore, uncertainties in emission estimates impose important limitations on regional carbon budget estimations derived by most atmospheric inverse frameworks (top-down approach), in which anthropogenic emission fluxes are assumed to be well-known (Corbin et al., 2010; Göckede et al., 2010; Gurney et al., 2002, 2005).

In order to assess accurately the contribution of a city or other emission hot spot to the CO₂ or other GHG emission, accurate knowledge of the surface fluxes at high spatial and temporal resolutions are needed. Ideally the accuracy of the estimated flux needs to be high for unambiguous attribution of source strength. The uncertainty of these estimations is required to be reduced to the extent that is feasible. In ESA (2015) it is noted (see their Sect. 4.1.2) that accuracies better than 10% would be useful for providing important additional information for cities where inventories exist, and accuracies better than 20% would contribute knowledge for cities where inventories do not exist.

The key limitations to constrain emission fluxes at urban scales via inverse modeling are the unavailability of direct, continuous and high frequency atmospheric CO₂ measurements representing CO₂ enhancement in urban domains, as well as the inability of current inverse modeling systems to capture the fine-scale variability caused by the atmospheric transport and emission processes at a scale relevant for urban emissions (e.g. Bréon et al., 2015). An

assessment study based on ground-based measurements indicated potential drawbacks of using CO₂ surface measurements for emission verification, and strongly recommended the use of sufficiently accurate column averaged CO₂ dry air mole fractions, denoted as XCO₂, measured from the ground and/or space as the best approach to detect and quantify emissions and emission trends from urban regions (McKain et al., 2012). An effective observation-based scheme is able to disentangle anthropogenic emissions from CO₂ fluxes originating from biosphere-atmosphere exchange.

Despite its importance, none of the existing satellites has been specifically designed and focused on observing XCO₂ at urban scales. However the first attempt to detect and quantify anthropogenic urban area CO₂ emissions from space was initiated with the launch of SCIAMACHY onboard ENVISAT (2002-2012) (Burrows et al., 1995; Bovensmann et al., 1999), which had a variety of atmospheric trace gas targets and applications. This has been followed by TANSO onboard GOSAT (launched in 2009) (Kuze et al., 2009)

Analysis of SCIAMACHY XCO₂ retrievals revealed that regionally elevated atmospheric XCO₂ over highly populated regions correlates well with anthropogenic CO₂ emissions in terms of relative emission increase per year (Schneising et al., 2008, 2013). However, these analyses are limited to large and intense emission regions owing to the coarse spatial resolution (~ 60 km x 30 km) of the SCIAMACHY measurements. Reuter et al. (2014) also presents results related to anthropogenic CO₂ emissions for large areas using an assessment of SCIAMACHY XCO₂ and NO₂ retrievals.

By using GOSAT observations, Kort et al. (2013) reported significant enhancements of XCO₂ over megacities (3.2±1.5 ppm for Los Angeles and 2.4±1.2 ppm for Mumbai), and argued that these enhancements can be exploited to track anthropogenic emission trends over megacities. However, constraining fossil fuel CO₂ emissions by using GOSAT XCO₂ retrievals is limited by the sparseness of the GOSAT data (Keppel-Aleks et al., 2013). Another satellite mission, OCO-2, has been launched in 2014, with the aim of measuring global XCO₂ with the precision, resolution, and coverage needed to characterize CO₂ sources and sinks at regional scales (≥ 1000km) (Crisp et al., 2004). In addition to these, there have been some recent attempts to utilize ground-based measurements of XCO₂ to constrain emissions from cities such as Los Angeles (Wong et al., 2015) and Berlin (Hase et al., 2015).

In an effort to overcome these limitations and to achieve XCO₂ observations with the precision and accuracy, spatiotemporal coverage, resolution, and sensitivity to near-surface concentration variations that are required to derive emissions at urban scales, a satellite mission was proposed to the European Space Agency (ESA): Carbon Monitoring Satellite (CarbonSat) (Bovensmann et al., 2010). CarbonSat aimed to measure XCO₂ and XCH₄ at a high spatial resolution (approx. 2 km × 2 km), with good spatial coverage via continuous imaging across a wide swath. The goal swath width for the proposed CarbonSat mission was 500 km, but smaller swath widths were also considered to limit cost (ESA, 2015).

In this study, we investigated two potential measurement swath widths: 500 km (goal requirement) and 240 km (breakthrough requirement). As a result of its relatively wide swath and high spatial resolution, CarbonSat is designed to disentangle natural and anthropogenic sources of CO₂ and CH₄ from localized sources such as cities, power plants, methane seeps, and landfills, by utilizing its unique greenhouse gas imaging capability achieved by its high spatiotemporal coverage and resolution. More details on the mission and the current instrument concept are given in Buchwitz et al. (2013a) and in ESA, 2015.

The goal of the present study is to assess an instrument like CarbonSat's capability to quantify emission patterns of moderate to strong localized sources, taking into account a realistic description of the retrieval errors as given in Buchwitz et al. (2013a), the spatiotemporal distributions of CO₂ emissions, and atmospheric transport. Here we present results focusing on Berlin in Germany, being a large city, but not a megacity. According to the classification of Globalization and World Cities (GaWC) for the year 2012 (<http://www.lboro.ac.uk/gawc/gawcworlds.html>), Berlin is categorized as a “Beta level” city that provides a moderate economic contribution to the world economy.

Berlin is located in the northeast of Germany (see Fig. 1) and is relatively isolated, i.e. it is not a part of a large agglomeration of several cities. This permits us to clearly identify the anthropogenic CO₂ emission plume of Berlin from a single CarbonSat “XCO₂ image”. We use a high-resolution modeling framework, comprising the Weather Research Forecasting (WRF) model combined with a greenhouse gas module (WRF-GHG, Beck et al., 2011) and the Vegetation Photosynthesis Respiration Model (VPRM) to simulate CO₂ mixing ratios for a domain centered on Berlin. An analysis is carried out for CarbonSat’s cloud-free overpasses for one reference year by applying a simple Bayesian inversion scheme to estimate the emission budget with associated uncertainty. A preliminary analysis using a least-squares-fitting algorithm was reported in Buchwitz et al. (2013b), but here we present more detailed analysis, which differs from the previous study as follows: the present study (1) uses high-resolution model simulations for each cloud-free CarbonSat overpass over Berlin for the simulated year 2008, (2) prescribes the updated emission inventory including hourly variations, (3) utilizes a Bayesian inversion approach, and (4) examines more scenarios to extend the error analysis study.

2. WRF-GHG inverse modeling system

A high-resolution inverse modeling system, utilizing atmospheric XCO₂ measurements at high spatial and temporal resolution, is used to retrieve the CO₂ emissions at an urban scale. It comprises two components: the WRF-GHG model linking atmospheric transport and the fluxes to realistically represent the distribution of atmospheric CO₂ mixing ratios, and a Bayesian inversion technique to optimize the fluxes. One primary objective is to quantify the uncertainties in the retrieved anthropogenic CO₂ emission fluxes resulting from typical and reasonable estimates of the systematic and random error of the XCO₂ measurements for an instrument like CarbonSat for the spatial resolution of 2km x 2km and the uncertainty in *a priori* knowledge of the surface flux of CO₂. For this, we used WRF-GHG forward simulations as the “true” representation of the atmospheric CO₂ concentrations and the associated fluxes as the “true fluxes” to be retrieved. Hence the deviation in the retrieved fluxes (via inverse optimization) relative to the “true fluxes” is caused by the CarbonSat simulated observation errors and the modeling errors (including the use of different emission inventories) depending on different scenarios analyzed. Each component of the inverse modeling system is described in the following.

2.1. WRF-GHG forward model simulations

The present study uses the WRF-GHG (version WRFv3.4) forward simulations of CO₂ concentrations at high spatial (10 km × 10 km) and temporal (1 hour) resolutions for all of CarbonSat’s overpasses over Berlin in the year 2008. The WRF-GHG modeling system has already been used in several regional studies and has shown remarkable performance in capturing fine-scale spatial variability of CO₂ mixing ratios (e.g. Ahmadov et al., 2007, 2009; Pillai et al., 2010, 2011, 2012). The model domain describes a region (spatial extent of ~ 900 km × 900 km) centered over Berlin (Fig.1) and the simulations use 41 vertical levels (the thickness of the lowest layer is about 18 m and the model top is 1.0 hPa). Simulations are conducted separately for each day for a period of 30 hours, including a meteorological spin-up time of 6 hours starting at 18 UTC the previous day.

The initial and lateral boundary conditions of the meteorological variables, the sea surface temperature (SST) and the soil initialization fields for each run are prescribed from the European Centre for Medium-Range Weather Forecasts (ECMWF) model analysis data (<http://www.ecmwf.int>) with a spatial resolution of about 25 km and 6-hourly temporal intervals. As initial atmospheric CO₂ fields and the lateral boundary concentrations, simulations use global CO₂ concentration simulations by the atmospheric Tracer transport Model 3 (TM3) with a spatial resolution of 4° × 5°, 19 vertical levels and a temporal resolution of 3 hours (Heimann and Körner, 2003). TM3 simulations used for this study are generated by a forward transport simulation of fluxes that have been optimized using a global network of CO₂ observing stations (Rödenbeck, 2005). Biospheric fluxes within the regional domain are calculated online in WRF-GHG with a diagnostic biospheric model, the Vegetation and Photosynthesis and Respiration Model (VPRM), utilizing remote sensing products and meteorological data at high temporal and spatial resolutions (Mahadevan et al., 2008). To obtain more realistic estimates of biospheric fluxes, a set of parameters in the VPRM,

specific for each vegetation class, have been optimized against eddy flux observations obtained during the CarboEurope IP experiment at various sites (21 measurements sites) under different vegetation types within Europe (Pillai et al., 2012). Regional oceanic fluxes are neglected here since their contribution is insignificant in the context of the present study.

2.1.1. Fossil fuel emission fluxes

The anthropogenic CO₂ emission fluxes are based on the EDGAR (Emission Database for Global Atmospheric Research, version 4.1, year 2008) global inventory with a spatial resolution of 0.1° x 0.1°. EDGAR is an annually varying database, but we apply time factors in order to provide hourly emissions. The time factors for seasonal, daily, and diurnal variations are based on the step-function time profiles published on the former EDGAR website: http://themasites.pbl.nl/images/temporal-variation-TROTREP_POET_doc_v2_tcm61-47632.xls (see Kretschmer et al. (2014); Steinbach et al. (2011) for further details). WRF-GHG simulations using these EDGAR emissions are treated as the real distribution of atmospheric CO₂ (hereafter referred to as “true CO₂ conc.”), and the associated EDGAR fluxes as “true fluxes”.

In order to examine the impact of the spatio-temporal distribution of fossil fuel emission structures on atmospheric CO₂ and to quantify the associated uncertainties in the optimized fluxes, we use different emission data as the prior emissions, namely those compiled by the Institut für Energiewirtschaft und Rationelle Energieanwendung (IER inventory), University of Stuttgart, (<http://carboeurope.ier.uni-stuttgart.de>) for the year 2000, at a spatio-temporal resolutions of 10 km and 1 hour. Temporal variations in the IER inventory include traffic rush hours, difference in power demand between weekdays and weekends, domestic heating, and air conditioning (Pregger et al., 2007). While utilizing the IER year 2000 database to represent the simulation year (2008), we apply scaling factors in a manner similar to that in Pillai et al. (2011) to preserve the temporal emission pattern differences between weekdays and weekends. Simulations using the IER database are used as the current knowledge about the atmospheric concentration for the inverse optimization described in Sec. 4.3.

Both these emission fluxes are re-gridded to WRF-GHG’s 10 km Lambert Conformal Conic projection grid, conserving the total mass of emissions. These hourly fluxes are added separately to the first model layer, and transported separately as tagged tracers. Figure 2 shows a spatial map of the averaged EDGAR and IER emission fluxes as well as their differences for the Berlin region. Strong emissions associated with the city can be seen well in both inventories. In general, both emission inventories show good consistency in terms of spatial emission structures; however significant differences in emission intensities (magnitude) between the inventories, especially for large cities and power plants, are common (e.g. Fig. 2c). These differences are larger for emissions resulting from power plants than for those from cities (not shown). Figure 3 shows the temporal variability of urban-scale emission fluxes in hourly, weekly and monthly averaged time scales for a region around Berlin (~ 100 km × 100 km). For Berlin emissions, considerable differences in temporal variations are found between both inventories, with maximum values of 22.5, 18.5, and 24.0 MtCO₂yr⁻¹ for hourly, weekly and monthly averaged timescales respectively. As compared to the IER inventory, the EDGAR inventory shows consistently larger emissions for Berlin, however this is not the case for some other cities in Europe. **Based on available sources of information, it is difficult to conclude which inventory is more accurate.** The seasonal variability exhibited by EDGAR Berlin emissions is substantially larger than that of the IER inventory. Larger emissions are seen in the EDGAR inventory in winter months, with values approximately a factor of 1.5 higher than those in summer months. This results from the increased demand of domestic heating in winter. In terms of the seasonal variability of the Berlin city emissions, the IER inventory shows a relatively small difference in winter-summer emission patterns (temporal) as compared to EDGAR, and shows overall larger emissions in winter. Both inventories show lower emissions during weekends, consistent with the reduced demand of transportation and power consumption. The hourly averaged Berlin emissions provided by both inventories display peak values during 7 to 9 am and 5 to 7 pm (local times), reflecting morning

and evening rush hours in terms of city traffic. Interestingly the IER Berlin emissions show “delayed” morning rush hours on weekends, with a maximum value around 11 am (local time).

The significant difference between these inventories in both temporal and spatial scales implies that our current knowledge of urban-scale emissions is inadequate, even for Central Europe, which is relatively well characterized in terms of emissions compared to many other parts of the world. Note that a part of these emission differences is likely due to the different data compilation years of the IER and EDGAR inventories. This “knowledge gap” is also important in inverse-modeling-based estimations of the source-sink distribution of CO₂, in which fossil fuel fluxes are generally assumed to be known. How critical the effect of this assumption is depends on the impact of these differences in emissions (emission uncertainties) on modeled atmospheric mixing ratios, as well as on the transport errors that are included in the model-data mismatch error in the inverse modeling framework. The impact of emission uncertainties is further discussed in Sec. 4.1.

2.2. Inverse optimization technique

The inverse optimization utilizes observational constraints to adjust a subset of parameters λ out of model parameters \mathbf{p} in the surface flux model $\mathbf{f}_m(\mathbf{p})$ in order to obtain a modeled concentration consistent with the observations. Hence the anthropogenic atmospheric concentration \mathbf{c} (column averaged dry air mole fraction) at different locations and times can be represented as:

$$\mathbf{c} - \mathbf{c}_{bg} = \mathbf{F} \mathbf{f}_m(\lambda) + \boldsymbol{\varepsilon}_{error} \quad (1)$$

Here, the matrix \mathbf{F} links the atmospheric concentration to a vector $\mathbf{f}_m(\lambda)$ whose dimension is equal to the total number of surface flux elements, multiplied by total time steps. The vector \mathbf{c}_{bg} is the background column averaged dry air mole fraction i.e. the concentration due to the advection of upstream tracer concentrations. For the inversion, $\mathbf{f}_m(\lambda)$ is assumed to be linearly dependent on λ and is expressed as:

$$\mathbf{f}_m(\lambda) = \boldsymbol{\Phi} \lambda \quad (2)$$

where λ represents a vector of daily scaling factors of surface fluxes, and $\boldsymbol{\Phi}$ represents the surface flux field over the model domain.

A linear model is obtained by combining Eq(s). 1 and 2:

$$\mathbf{y} = \mathbf{K} \lambda + \boldsymbol{\varepsilon}_{error} \quad (3)$$

where the measurement vector \mathbf{y} is given by

$$\mathbf{y} = \mathbf{c} - \mathbf{c}_{bg} \quad (4)$$

and, \mathbf{c}_{bg} is obtained by linearizing the model with a reference state $\lambda_0 = 0$ (see Eq. 1).

The Jacobian matrix that represents the sensitivity of the observations \mathbf{y} to the state vector λ is given by

$$\mathbf{K} = \mathbf{F} \boldsymbol{\Phi} \quad (5)$$

The state vector and the Jacobian matrix are further described in Sec. 3.2. *A priori* knowledge of the surface fluxes, λ_{prior} , along with their uncertainties is incorporated in the Bayesian formulation. The term, $\boldsymbol{\varepsilon}_{error}$, is assumed to follow the Gaussian distribution described by the error covariance matrices of the measurements, \mathbf{S}_e and the prior estimate, \mathbf{S}_{prior} . The posterior estimate of λ is obtained by minimizing the cost function, J , which is given as:

$$J(\lambda) = (\mathbf{y} - \mathbf{K} \lambda)^T \mathbf{S}_e^{-1} (\mathbf{y} - \mathbf{K} \lambda) + (\lambda - \lambda_{prior})^T \mathbf{S}_{prior}^{-1} (\lambda - \lambda_{prior}) \quad (6)$$

Analytically solving for the minimum of Eq. (6) gives the optimal estimate of the state vector of the scaling factors $\hat{\lambda}$, as well as the associated error covariance matrix of $\hat{\lambda}$, termed as the posterior uncertainty, $\mathbf{S}_{\hat{\lambda}}$. These are expressed as follows (Rodgers, 2000):

$$\hat{\lambda} = (\mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{K} + \mathbf{S}_{prior}^{-1})^{-1} (\mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{y} + \mathbf{S}_{prior}^{-1} \lambda_{prior}) \quad (7)$$

$$\mathbf{S}_{\hat{\lambda}} = (\mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{K} + \mathbf{S}_{prior}^{-1})^{-1} \quad (8)$$

3. Bayesian Inversion of CarbonSat measurements

3.1. Pseudo observations

The inversion utilizes a one year dataset of CarbonSat simulated observations at a spatial resolution of $2 \text{ km} \times 2 \text{ km}$, generated using the WRF-GHG forward model ($10 \text{ km} \times 10 \text{ km}$) as described in Sect. 2.1 and CarbonSat's retrieval error ($2 \text{ km} \times 2 \text{ km}$), estimated using an error parameterization scheme based on the measurement characteristics as described in Buchwitz et al. (2013a). The error parameterization scheme, described in detail in Buchwitz et al. (2013a), is based on six parameters consisting of solar zenith angle (SZA) and scattering-related parameters such as albedo in the near-infrared (NIR) and the first shortwave-infrared (SWIR-1) bands, cirrus optical depth (COD), cirrus top height (CTH), and aerosol optical depth (AOD) at 550 nm. We use the "Level 2 error dataset" (L2e files), described in Buchwitz et al. (2013a), that contains the random and systematic errors of CarbonSat's XCO₂ retrievals based on the error parameterization scheme. CarbonSat is assumed to follow an orbit similar to NASA's Terra satellite (www.nasa.gov/terra/), but with an equator crossing time of 11:30 a.m. Hence, for specifying the CarbonSat's geolocation, the L2e files utilize the geolocation provided in the Terra Level 1 dataset for the year 2008, but modified to consider the difference in equator crossing time. This dataset contains fields such as geodetic coordinates, ground elevation, and solar and satellite zenith angles etc., determined using the spacecraft attitude and orbit, a digital elevation model, and information derived from various other datasets such as the Filled Land Surface Albedo Product, generated from MOD43B3 (<http://modis-atmos.gsfc.nasa.gov/ALBEDO/>) at a spatial resolution of 1 minute (2 km at equator, and $< 1 \text{ km}$ at the poles), which is used to account for surface albedo. The cirrus parameters are represented using a spatiotemporally smoothed ($8^\circ \times 8^\circ$ and 3 months) dataset of COD and CTH, originally derived from CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) onboard CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations, Winker et al., 2009). Global aerosol data products from the "GEMS project" (<http://gems.ecmwf.int/>) at a spatiotemporal resolution of $1.125^\circ \times 1.125^\circ$ and 12 hourly are used to account for aerosols (AOD). This dataset is based on the assimilation of MODIS data and we use the AOD at 550 nm. As described in Buchwitz et al. (2013a), the L2e dataset only contains those CarbonSat simulated observations which are approximately cloud-free as determined using a cloud mask obtained from MODIS Terra (using the MODIS cloud cover data product (MOD35) at a spatial resolution of about $1 \text{ km} \times 1 \text{ km}$). As the remaining ground pixels may still suffer from cloud contamination (e.g., due to "too high" amounts of thin cirrus) or other disturbances, a quality filtering scheme is applied which is based on retrieved (e.g., COD and AOD) and known quantities (e.g., SZA). The quality filtering scheme is described in Buchwitz et al. (2013a) and we use here only those ground pixels which are considered "good" according to this scheme.

Initially, we have identified all the potentially useful Berlin overpasses, i.e., overpasses where at least some CarbonSat simulated observations are present over Berlin and surroundings for a given CarbonSat orbit. We found that the maximum number of observations is obtained during the summer months due to most favorable observation conditions (less clouds for extended time periods and regions, high SZA, etc.). In total, there are 41 days (orbits) of potentially useful overpasses over Berlin for the year 2008 for a swath width of 500 km. Note that the number of overpasses are smaller in the figures shown later. This is because of an additional quality filtering procedure applied after the inverse optimization that is based on retrieved random errors, as explained later.

3.2. Definition of the state vector and Jacobian matrix

In the present study, the state vector has two elements. The first element λ (the scalable parameter of the emission flux) corresponds to the scaling factor of emission fluxes for a trimmed model domain, i.e., a region around Berlin (spatial extent: approximately 100 km \times 100 km, hereafter referred to as the “target region” (TR). The other element is a constant, i.e. $\lambda_0 = 0$, for the entire scene per overpass to account for variations of the background XCO₂ (see Eq. 4) and to treat the background variations independently of the city emissions as done in Buchwitz et al. (2013b). The temporal resolution of λ is set to be daily, assuming no spatial variations within the target region. The prior value of this scaling factor, λ_{prior} , is set to unity.

The Jacobian matrix \mathbf{K} relates the measurement vector \mathbf{y} to the state vector λ , and has elements that represent the response in mixing ratios to the emission fluxes (see Eq. 5) and the dimension of \mathbf{K} is $n \times m$ where n corresponds to the numbers of elements in the state vector and m is the number of XCO₂ observations. Since we do not have an adjoint model, these sensitivity functions are derived by perturbing each element of the emission flux field Φ over the target region by small increment and applying the forward model (WRF-GHG) to obtain the resulting perturbed concentration field ($\mathbf{C} + \Delta\mathbf{C}$) over the target region. Hence, \mathbf{K} is calculated as follows:

$$\mathbf{K} = \frac{\mathbf{C} + \Delta\mathbf{C} - \mathbf{C}}{\sum_{\text{TR}} \Phi_{\text{perturbed}} - \sum_{\text{TR}} \Phi} \quad (9)$$

The posterior estimate of the scaling factor, $\hat{\lambda}$, is derived by minimizing the cost function, $J(\lambda)$, as given in Eq.7.

3.3. Error covariance matrices

Bayesian inversion utilizes error covariance matrices to account for the measurement error and the prior flux error variances and co-variances. The measurement error covariance matrix, \mathbf{S}_e , is constructed by specifying the XCO₂ random errors (single measurement precision) derived using the error parameterization scheme described in Sect. 3.1. Note that the XCO₂ random error is primarily determined by the instrument signal-to-noise performance (but also to some extent by the retrieval algorithm, see Buchwitz et al. (2013a)) and is typically about 1.2 ppm (for the assumed threshold requirement signal-to-noise ratio performance assumption used by Buchwitz et al., 2013a) except for some especially unfavorable conditions such as low albedo and high SZA scenarios. Transport model uncertainty is neglected here since the objective of current study is to quantify the uncertainty in the retrieved fluxes due to CarbonSat’s retrieval errors only.

The prior flux uncertainty, \mathbf{S}_{prior} , is set uniformly to 40 % of the total emission over the target region to ensure that the difference between the “true” and prior fluxes is appropriately considered. We consider the fact that the increased variability of emissions at the high resolution (as it is used in this study) leads to increased uncertainty due to the lack of information about the emission processes at the required spatial and temporal resolutions. The magnitude of \mathbf{S}_{prior} is specified here based on the approximate difference between the IER and the EDGAR inventories over the target region.

4. Results: Estimation anthropogenic XCO₂ enhancement and retrieved flux uncertainty over Berlin

In this study, we use anthropogenic XCO₂ enhancement, which is defined as the enhancement in XCO₂ resulting from local anthropogenic emissions relative to the background concentration. The tagged tracer option in WRF-GHG stores XCO₂ enhancement resulting from EDGAR emissions separately, and we use this field to represent anthropogenic XCO₂ enhancement. The uncertainty in the retrieved emission attributed by CarbonSat’s retrieval error is a function of the anthropogenic XCO₂ enhancement over Berlin, the number of potential observations in and around Berlin, and the retrieval uncertainty (random and systematic components). In this manner we take into account the influence of these parameters to achieve a robust estimation of the retrieved surface emission uncertainty or error.

4.1. Local anthropogenic XCO₂ enhancement

The XCO₂ enhancements resulting from anthropogenic emissions over Berlin are estimated in order to assess whether these emission enhancements are detectable by an instrument having the performance of CarbonSat i.e. to assess whether the resulting plumes are statistically significant and robust, thereby enabling the changes or trends in anthropogenic emission over the cities.

Figure 4 shows the “true” anthropogenic XCO₂ enhancement on a reference day (24th June 2008), the anthropogenic XCO₂ enhancement based on the IER inventory, and the difference in XCO₂ enhancement due to the difference in emission inventories. From Fig. 2a and Fig. 4a, it can be concluded that, given the availability of a satellite instrument which is able to precisely detect the associated XCO₂ mixing ratio enhancements ranging from 0.80 to 1.35 ppm at a high spatial resolution and adequate spatial coverage, anthropogenic emissions from a city with the size of Berlin and other localized emission sources can be estimated from space with sufficient accuracy. It should be noted that the magnitude of detectable anthropogenic XCO₂ enhancements is likely to be underestimated in our study because the “true” fields of XCO₂ variations are simulated at a 10 km spatial resolution instead of CarbonSat’s resolution (~2 km × 2 km).

Noteworthy is that the spatial and temporal difference in EDGAR and IER emission inventories gives rise to a notable XCO₂ mixing ratio difference between 0.4 and 1.0 ppm. For Berlin, this is about 40% of the total “true” XCO₂ enhancement. It should be noted that surface concentrations show larger relative differences than the column dry mole fraction for CO₂, XCO₂, because of their higher sensitivity to the change in surface fluxes. Hence this result indicates the importance of characterizing emission uncertainties, even for the region where fossil emissions are often considered to be “well-quantified” in comparison to the biospheric carbon balance. Neglecting this uncertainty term would lead to significant biases in the net carbon exchange estimations, particularly when assimilating concentration measurements closer to emission sources such as cities.

4.2. Uncertainty of the retrieved Berlin emissions

In this section, we show the results obtained by inverting CarbonSat simulated observations over the target region, taking into account different sources of possible errors including CarbonSat measurement errors and modeling errors. Inversions are performed separately for each potentially useful CarbonSat overpass (see above) to derive the total emission flux and its error over the target region. Note that those fluxes retrieved or “seen” from CarbonSat XCO₂ measurements can vary significantly from one to the next overpass i.e. within weeks, because the time elapsed to transport the CO₂ plume and thus changes in XCO₂ to where it is observed by CarbonSat vary with wind speed and strong temporal variations in emissions (see also Fig. 3). Figure 5 shows an overview of prior fluxes used for these inversions.

The systematic error (SE) of the retrieved emission fluxes, which are specific for each source of errors or combination of errors, are determined separately by defining six scenarios represented by S01 through S06 (Table 1). These scenarios are described in the following subsections, while additional scenarios S07 – S11 are presented and discussed separately in Sect. 4.3. Note that the distance from the center of the target region to one of its boundaries is roughly 50 km, which corresponds to a time of approximately 3 hours for air parcels travelling with a velocity of 4.5 ms⁻¹. This means that the observed local CO₂ emission plume is not only determined by the emission at the time of the overpass, but also during a time interval of several hours before the time of the overpass. This is taken into account when modeling the CO₂ emission plume. For the inversion, it is assumed that the time dependence of the emissions in the time period of up to several hours (3 to 6 hours) before the overpass is at least reasonably well known except for the scenarios S07 to S11. As noted earlier, the “true” XCO₂ variations in this study are based on 10 km spatial resolution instead of 2 km in CarbonSat simulated observations. For the inversion results, we assume negligible representation error arising from these spatial scale mismatches. Based on meteorological conditions, the representation error introduced by decreasing the horizontal resolution from 2 km to

10 km can be approximately 0.5 ppm on average for CO₂ concentrations at the surface (Tolk et al., 2008). However, it is expected that the representation error for XCO₂ between these horizontal scales will be much lower than that for CO₂ concentration at surface (see Pillai et al., 2010).

Before analyzing SE for the different scenarios, we first present the random error (RE) of the retrieved emission. RE is caused by the measurement noise, i.e., by the random part of the measurement error; hence it is independent of the above-mentioned SE scenarios. In the optimal case, the instrument noise is determined by the shot noise of the detector arrays. In practice, there are additional sources of noise such as read out noise, digitization noise etc. Figure 6 shows the random errors of the retrieved emissions over the target region, obtained by inverting the entire one-year data set of simulated CarbonSat XCO₂ retrievals. As explained above, we have investigated two different swath widths, 500 km and 240 km. The results are shown only for the days where the number of CarbonSat simulated observations around the target region is sufficiently dense (covering the emission plume and its surroundings) to obtain a retrieved emission random error of less than 25 %, i.e., we use the a posteriori random error of the retrieved emission as a quality criterion (as also done in Buchwitz et al. (2013b)). Applying this additional quality criterion has further reduced the number of potential overpasses for the 500 km swath width. This number, labeled as “N” useful overpasses, is 27 for a swath width of 500 km and 17 for a swath width of 240 km. The value obtained here for “N” useful overpasses is expected to be typical for other cities with similar cloud coverage and latitude. As can be seen in Fig. 6, decreasing the swath width not only reduces the number of useful overpasses, but also increases the RE of the retrieved fluxes for some overpasses. The RE of the retrieved emission (from a single overpass) is usually found to be less than 20% (approximately 10 MtCO₂ yr⁻¹) of the emission fluxes for both swath widths.

4.2.1. Impact of CarbonSat measurement errors (scenario S01)

Here, we focus on scenario S01, and estimate the uncertainty in the retrieved emission fluxes caused exclusively by CarbonSat measurement errors. For this, we assume that the XCO₂ variability in the target region is dominated by the anthropogenic CO₂ emission and that there is negligible XCO₂ variability due to biogenic fluxes over the target region, or that this biogenic component can be modeled well, and thus can be subtracted from the observations without introducing any modeling-related errors.

The systematic measurement error of the CarbonSat simulated observations over the target region for a typical day (24th June 2008) for S01 is shown in Fig. 7a. This is estimated using the error parameterization scheme of Buchwitz et al. (2013a), as shortly described in Sec. 3.1. The mean systematic measurement error over the target region is about 0.25 ppm for this day. For the scenario S01, the “observed” anthropogenic XCO₂ by CarbonSat is thus the sum of this measurement error (Fig. 7a) and the “true” anthropogenic XCO₂. Fig. 8a shows the observed anthropogenic XCO₂ enhancement for S01 over the target region during the overpass on 24th June 2008. For the comparison, the corresponding “true” anthropogenic XCO₂ enhancement, i.e. without any source of errors, is shown in Fig. 8g. The “true” emission plume, originating almost from the centre of the target region, can be clearly seen with a maximum value of about 0.90 ppm. As can be seen, the observed CarbonSat XCO₂ pattern (Fig. 8a) differs from the “true” XCO₂ pattern (Fig. 8g) by the measurement errors (Fig. 7a); hence the retrieved emission via inversion typically differs from the “true” emission that results in a systematic error (SE) of the retrieved emission. The extent of this systematic error depends on how well the systematic measurement error correlates with the “true” XCO₂ pattern.

Figure 9 shows the systematic errors of the retrieved emissions for CarbonSat overpasses over the target region obtained by inverting the entire one-year data set of simulated CarbonSat XCO₂ retrievals for the scenario S01. Shown are the results for swath widths of 500 km and 240 km for all “N” useful overpasses (days). Overall, the absolute magnitude of the systematic errors of the retrieved emissions for both swath widths for the scenario S01 is found to be less than 10% for most of the overpasses (about 75% of the “N” useful overpasses for the year 2008), which corresponds to about 5.3 MtCO₂ yr⁻¹. For the 500 km swath width in S01, the mean and standard deviation of the SE for all “N” useful overpasses is -2.5 MtCO₂ yr⁻¹ (-5.3%) and 2.8 MtCO₂ yr⁻¹ (6.1%), respectively (see also

Table 1). In general, we find that the two different swath widths have a negligible impact on the daily SE of the retrieved emissions, although decreasing the swath width reduces the “N” useful overpasses.

4.2.2. Impact of CarbonSat measurement errors with worst-case aerosol related biases (scenarios S02 and S04)

Note that in the previous section we have used the CarbonSat systematic XCO₂ retrieval errors as provided by the error parameterization scheme described in Buchwitz et al. (2013a). However, as explained in Buchwitz et al. (2013b), this scheme may underestimate aerosol related biases if the spatially (not aggregated) high-resolution CarbonSat simulated observations are used for applications like the one used here. The reason is that aerosol-related retrieval biases have been computed using quite smooth model aerosol input data sets, which might not be sufficient to represent the aerosol plume over Berlin.

To consider this, an additional error term has been defined which is referred to as “high resolution aerosol error” in this manuscript. In this sub-section we present results for scenario S02, where the measurement error used for S01 described in the previous section has been replaced by the high-resolution aerosol error contribution to the systematic measurement error. We also present results for scenario S04, where the measurement error is the sum of the S01 and S02 errors.

The method of computing the “high resolution aerosol error” is described in detail in Buchwitz et al. (2013b). Here we describe it briefly as follows. A local AOD enhancement has been computed by scaling the observed anthropogenic XCO₂ spatial pattern, i.e., the AOD enhancement is assumed to be perfectly correlated with the CO₂ emission plume of interest (see Fig. 7b and Fig. 8b). To quantify the urban aerosol enhancement over a region around two power plants in Germany and to study their impact on emission estimates, Krings et al. (2011) followed the above criteria and used an AOD scaling factor of 0.05 per 1 (4 ppm) % of local Δ XCO₂. As compared to their study, we have used a much high scaling factor of 0.2, i.e., the AOD change, Δ AOD at 550 nm is 0.2 per 4 ppm of local anthropogenic Δ XCO₂. Overall, these are worst-case assumption that are supposed to result in upper limits of systematic XCO₂ errors due to aerosols and resulting errors of the retrieved emissions. For a more detailed discussion see Buchwitz et al. (2013b).

The resulting SEs of the retrieved emissions for scenario S02 are found to be negative, indicating systematic underestimation of retrieved emissions (see Fig. 10). As can be seen, the absolute magnitudes of errors are slightly higher than those for S01. The mean and standard deviation of SE for S02, considering all “N” useful overpasses and the 500 km swath width, are -3.6 MtCO₂ yr⁻¹ (-7.5%) and 2.1 MtCO₂ yr⁻¹ (3.4%) respectively.

Another scenario, S04, investigates the impact of both high-resolution aerosol related errors (used for S02) and the “default” CarbonSat measurement errors (used for S01) on retrieving anthropogenic emissions. Inversions are performed by utilizing these two sources of error, i.e. the XCO₂ systematic error for S04 is the sum of XCO₂ systematic error specified in S01 and S02 (see Fig(s). 8d and 9d)). As expected, the SE of the retrieved emission for S04 is found to be higher than those of S01 and S02, and their values are close to the linear sum of systematic emission errors for S01 and S02 (see Table 1). As already explained, the definition of S04 likely represents the possible worst-case measurement scenario in particular with respect to aerosol related errors.

4.2.3. Impact of biospheric modeling error (S03, S05, S06)

In this section, we explore the impact of modeling error on retrieving Berlin city emissions. In the last two sections, it is assumed that the spatial variability introduced by the biogenic component of XCO₂ in the target region is well known or sufficiently small that it can be ignored. However, in reality there are notable perturbations caused by the spatial variability of biogenic XCO₂ in the target region that cannot be ignored. As an example, Fig. 7c illustrates the biogenic XCO₂ variability in the target region during a CarbonSat overpass. Most critical in terms of this uncertainty is how well the biogenic XCO₂ pattern is correlated with the anthropogenic XCO₂ pattern. In this case,

the uncertainty in the retrieved emissions depends on how accurately the biogenic fluxes can be modeled, as well as the associated transport model uncertainty in simulating the biogenic XCO₂ pattern. Note that we assume negligible transport uncertainty for the anthropogenic XCO₂ pattern in order to distinguish the retrieved emission errors due only to the biogenic XCO₂ pattern. In order to account for this modeling-related error, we consider scenario S03. In S03, we assume an extreme case where biogenic XCO₂ cannot be modeled at all; hence biogenic XCO₂ is treated as the “perturbation” seen in the measurement vector (y) of the inversion system (see Fig(s). 8c and 9c). However, it should be noted that in reality biospheric modeling uncertainty is not expected to be as high as this assumption. A simple biosphere model such as VPRM used in this study could capture 50 to 65% of the biospheric flux variability in most of the cases (squared correlation coefficient (VPRM vs. observations), $R^2 \sim 0.50-0.65$).

The systematic errors of the retrieved emissions for S03 are found to be significantly higher compared to the errors for the above-mentioned scenarios than those for S01 and S02 (see Fig. 11). Noteworthy is that this uncertainty is not related to CarbonSat measurement errors, but arises due to the inability of the model to simulate the biospheric contribution. Hence this uncertainty should be treated as model-related error. Due to the extreme assumption of modeling error in S03, the uncertainty values reported in this section have to be considered as the extreme upper limits of the possible total uncertainties in the retrieved fluxes due to biogenic modeling error. Despite this, the SE of the retrieved emission for S03 is within the range of 20 to 25% (10 to 15 MtCO₂ yr⁻¹) for most of the scenes although we assumed the largest uncertainty in modeling biogenic XCO₂. The reason for this is that the spatial biospheric XCO₂ pattern in the target region that “disturbs” the inverse system typically differs from the anthropogenic XCO₂ pattern in many of the good CarbonSat overpasses, enabling these two sources/sinks (anthropogenic and biogenic) to be disentangled reasonably well for cities like Berlin.

Additionally, we define other scenarios, S05 and S06, to investigate the impact of the biogenic modeling errors in combination with other error sources, such as CarbonSat measurement errors and high-resolution aerosol related errors. Systematic error estimations for these scenarios are summarized in Table 1 and these results suggest that a dominant part of the retrieved emission error is caused by the unknown biogenic variability.

4.3. Inversion Experiment using different prior emission fluxes (S07-S11)

The inversion results presented so far have not taken into account the impact of imperfect knowledge of the spatial pattern of emission fluxes and the different time dependences of the emissions; hence the inverse optimization adjusts only the amplitude of the emission plume corresponding to the anthropogenic CO₂ emission in the target region. Although the error arising from these unknown spatial emission structures is not directly related to CarbonSat measurement errors, we attempt to perform an experiment using two different flux inventories, with one of the flux inventories representing the prior fluxes and the other as the “true” fluxes. The experiment is designed with an inversion set-up, which is essentially the same as that described in Sect. 3.3, but with the following exception: Here the prior emission fluxes are prescribed from the IER emission inventory (Fig. 2b); hence the modeled anthropogenic XCO₂ is based on IER emission fluxes (see Fig. 4b and Sec. 2.1.1). Similar to the sections above, the EDGAR emission inventory is considered to be the “true” fluxes and the measurement vector (y) which corresponds to CarbonSat simulated observations is based on the EDGAR emission inventory as described in sec. 3.1. The retrieved posterior fluxes of this inversion optimization are compared with “true” fluxes to estimate the retrieved posterior flux errors and to assess how well inversion studies can benefit from CarbonSat measurements in the case of discrepancy between “true” and prior fluxes in terms of spatial patterns of distribution.

Similar to the above section, systematic errors of the retrieved fluxes are estimated specifically for each source of errors or combination of errors by defining scenarios S07 through S11 (see Table 1). It should be noted that the IER and EDGAR fluxes are not entirely different in terms of temporal variations, though the magnitude of the emissions in the target region is notably different (see Fig. 5). However, there exists a dissimilarity of approximately 70% of the spatial patterns between these two inventories (based on the correlation of spatial variability between two inventories, $R^2 \sim 0.30$) in the target region.

For most of the overpasses, the random errors of the retrieved emission fluxes over the target region (single overpass) are found to be less than 20% (approximately $10 \text{ MtCO}_2 \text{ yr}^{-1}$) of the emission fluxes for both swath widths (not shown). These values are comparable to those shown in Fig. 6, indicating the potential of CarbonSat simulated observations to retrieve surface fluxes, even when uncertainties in the spatial pattern of the prior emission fluxes are present. Figure 12 shows the SE of the retrieved emissions estimated for the scenario S09, where CarbonSat measurement errors and high resolution errors are considered in addition to the uncertainty in the spatial pattern of the prior fluxes. For both swath widths, the estimated SE for S09 is found to be less than $10 \text{ MtCO}_2 \text{ yr}^{-1}$ in many instances (for about 85% of useful overpasses). Systematic errors for other scenarios are summarized in Table 1. Depending on the error sources, the inversion experiment shows that the mean and standard deviation of SE, considering all “N” useful overpasses and the 500 km swath width, ranges from -0.2 to $10.8 \text{ MtCO}_2 \text{ yr}^{-1}$ and 6.9 to $11.1 \text{ MtCO}_2 \text{ yr}^{-1}$ respectively. Furthermore, the systematic errors of the retrieved emission fluxes for both swath widths are found to be lower than the difference between the prior fluxes and the “true” fluxes except for a very few cases, providing confidence in the inverse results although only a simple inverse optimization methodology is used.

5. Discussion

In this section, we discuss the merits of instrument like CarbonSat for retrieving emission fluxes and its potential in disentangling anthropogenic and biogenic CO_2 fluxes over cities like Berlin. Caveats related to the simple inversion approach used here are discussed.

For the study of CO_2 emissions, it is necessary to assess whether local anthropogenic XCO_2 enhancements are large enough to be detected by using the retrieved XCO_2 data products from the satellite-borne instrument, taking into account the measurement noise. Our analysis shows that anthropogenic XCO_2 enhancements around Berlin are well above the retrieval biases for most of the overpasses and the number of potential observations, after filtering out the contaminated pixels, is large enough to minimize the random error component (not shown). Given the availability of such a dense sampling coverage with similar retrieval biases, one can be confident in utilizing CarbonSat’s observations for retrieving city emission trends or absolute emission fluxes via appropriate inverse modeling.

In a real scenario the question arises whether it is possible to clearly separate local anthropogenic XCO_2 enhancements from CarbonSat’s total column measurements, which are in addition influenced by biospheric sources or sinks. Moreover, in order to isolate the XCO_2 enhancement caused by local sources (such as city emissions), it is necessary to specify the “background” signal, representing the CO_2 column without any influence of local fluxes. These additional biospheric and background influences can be ignored if the target city is well isolated from other strong urban sources and/or active biospheric regions as well as has negligible local biospheric activity. However, only a few cities or urban areas meet the above criteria, and a typical European city, in general, has considerable local or nearby biogenic influences. Under these conditions it is necessary to disentangle biogenic, anthropogenic and background contributions from CarbonSat’s observations. To assess the relative contribution of biogenic and anthropogenic sources, one can utilize additional co-emitted tracers such as CO and NO_x (Newman et al. 2013; Silva et al., 2013; Berezin et al., 2013 and Reuter et al. 2014). In the time frame of a potential CarbonSat mission, Sentinel-5 will be providing data on CO and tropospheric NO_2 (Ingmann et al., 2012), which when combined with CarbonSat data is expected to provide information for the attribution of air masses originating from fossil fuel combustion. Depending on the extent of the variability and the possible uncertainties, we can also rely on the biospheric and global model simulations to differentiate different source-sink contributions.

By assuming that the biospheric patterns are accurately modeled and that these biogenic signals can be subtracted from the measurement vector to isolate the anthropogenic contribution of XCO_2 , our simple inversion system is constructed such that it takes into account the impact of CarbonSat sampling errors on the retrieved city emissions over Berlin. The applicability of our results to a scenario where these assumptions are not valid needs to be examined, but the current set-up is not well suited for this purpose since we have not taken into account additional

state vectors for biospheric contributions. On the other hand, the current setup allows us to investigate the extremely pessimistic scenario where we assume that we cannot model the biospheric contribution at all (see Sec. 4.2.3).

When using observations at CarbonSat's 2 km spatial resolution, as mentioned in Sec. 4.1, it is likely that the magnitude and variability of local anthropogenic XCO₂ enhancement would be higher than our estimation that is based on simulations at 10 km spatial resolution. One of the main advantages of CarbonSat's resolution is its ability to provide a large number of cloud-free observations and this study identified the potential observations over Berlin by utilizing CarbonSat's 2 km spatial resolution.

Although we utilize high-resolution forward simulations, at present our inversion system uses only one scaling factor for the entire target region for each useful overpass. This means that the current set-up cannot provide posterior estimates for each pixel or emission sector within the target region. In other words, the flexibility to capture the true spatial variation of fluxes is more limited in our simple inversion system than in pixel- or parameter-wise inversions. Using this simple inversion system may thus overestimate the retrieved flux uncertainty. While interpreting our results, one should keep in mind that we do not specify other important sources of errors in the inversion system such as transport error. As previously noted, the main focus of this study is to estimate the retrieved flux uncertainties that are caused only by CarbonSat's measurement errors. However, these transport related errors, which provide proper weight to the observations depending on the capability of the transport model, need to be taken into account when estimating the total flux uncertainty via inverse modeling.

6. Summary and Conclusion

In the present study, we examine the potential of a satellite mission like CarbonSat for improving the current knowledge on the surface-atmosphere exchange of atmospheric CO₂. A significant contribution by the CarbonSat greenhouse gas (GHG) observations will be the ability to retrieve the emissions of localized (moderate to strong CO₂ and CH₄) emission sources such as cities, power plants, methane seeps, etc., as a result of its unique sampling capability at high spatial resolution (approximately 2 km × 2 km) with a good spatial coverage using a much wide swath. To demonstrate this, we have investigated the error on the retrieved fluxes using synthetic data which are similar to that expected from CarbonSat. We have simulated emissions from a medium-size city (in terms economic contribution and trade) and assessed the capability to retrieve anthropogenic emission fluxes for the city and its surrounding region (Berlin-centered target region investigated here: ~100 km × 100 km) from CarbonSat simulated observations. The results show that these potential space-based top-down flux estimates have high accuracy; hence this study contributes to the definition of achievable targets for emission fluxes at the city scale.

The study utilizes a Bayesian inversion approach based on the WRF-GHG modeling system at a high spatial resolution to optimize anthropogenic CO₂ emissions for the target region using CarbonSat simulated observations for a time period of one year. The inverse system is designed in such a way that one can quantify the random and systematic errors of the retrieved anthropogenic emission fluxes for a given set of XCO₂ measurement and modeling errors. The CarbonSat measurement errors are estimated using the error parameterization scheme of Buchwitz et al. (2013a), which takes into account different sources of uncertainties including scattering related errors. Based on the EDGAR emission inventory, the local anthropogenic XCO₂ enhancement over Berlin is found to be approximately 0.80 to 1.35 ppm. The latter is similar to the detectable limit of single CarbonSat ground pixels. However typically there will be several hundred observations available per overpass, sampling the emission plume and its surrounding. The impact of CarbonSat measurement errors on the retrieved emissions is assessed for two swath widths (240 km and 500 km). By performing a Bayesian inversion based on one year of CarbonSat simulated observations, we show that the random error of the retrieved Berlin CO₂ emissions is typically less than 15 to 20% of the total city emissions. In other words, the CarbonSat measurements can be utilized in atmospheric top-down approaches to quantify emissions of medium sized cities such as Berlin with a precision better than 8 to 10 MtCO₂ yr⁻¹.

In order to quantify the systematic error (SE) of the retrieved fluxes, we use different scenarios in terms of various sources of systematic error in the inversion system. For scenario S01, we use CarbonSat's "default" XCO₂ systematic errors (retrieval biases) from Buchwitz et al. (2013a), and assume no biogenic XCO₂ modeling error. For S01, we find that SE is in the range of 3 to 6 MtCO₂ yr⁻¹ for most of the cases (40 to 80% of the "good" overpasses as identified by the quality filtering procedure), indicating a high potential of utilizing CarbonSat's measurements to retrieve city emissions. Based on the analysis using a one-year period of CarbonSat simulated observations, we show that narrowing the swath width (from 500 km to 240 km) decreases the total number of useful overpasses, as expected, but we do not find any significant difference between the single overpass SEs estimated for the two swath widths investigated here.

As explained in Buchwitz et al. (2013b), the default XCO₂ systematic errors only reflect aerosol related biases at quite low spatial resolution. On the spatial scale of the city of Berlin, aerosol-related biases may be larger. To consider this, we use the "worst-case" measurement scenario as used by Buchwitz et al. (2013b), in which we assume that the aerosol-related biases may be perfectly correlated with the signal of interest, which is the city CO₂ emission plume in combination with a high amount of aerosols in the plume. For this, we define a scenario S04 and refer to this as "high resolution aerosol error" in this manuscript. The estimated emission uncertainty for this scenario (S04) is found to be higher than that of S01, with mean and standard deviation of approximately -6.1 and 3.8 MtCO₂ yr⁻¹ respectively.

The above-mentioned results, however, are mostly dominated by the assumption that there is a negligible influence of biospheric fluxes that perturb the emission plume over the target region, or that these biospheric contributions can be modeled very well. By further investigating the extreme case in which the biospheric contribution is assumed to be totally unknown and treated as perturbation in the inversion system (scenario S03), we find that the single overpass SE of the retrieved emission is significantly increased to 7.5 ± 9.5 MtCO₂ yr⁻¹ (mean \pm standard deviation). Nevertheless, the magnitude of the uncertainty is not overwhelmingly large over the target region, despite the worst-case assumption used here. It should be kept in mind that the above-mentioned uncertainty is not directly related to the performance of CarbonSat measurements, but more towards the models' inability in simulating the biospheric contribution well. Hence, for the effective utilization of these measurements, the "noises" induced from other sources have to be taken into account, which requires careful design of the inverse optimization methodology using transport models at high resolution, enabling them to handle the information contained in those measurements. On comparing the results from different scenarios, we show that the systematic error of the retrieved fluxes depends largely on the accuracy of the CarbonSat simulated observations and more importantly on the modeling related errors.

Further investigation by designing a synthetic inversion experiment is motivated by the possible impact of spatial structural variability of the emission fluxes, which is not considered in the above-mentioned inversions. We acknowledge that our current inversion set-up is too simple to examine how suitable CarbonSat measurements are for this purpose, as we use only one scaling factor for the entire target region. Nevertheless, we find promising results from this experiment in which the modeled and true XCO₂ concentrations are based on two distinct emission inventories (IER and EDGAR) differing in spatiotemporal patterns. By showing that the systemic error of the retrieved fluxes is lower than the difference between the prior fluxes and the true fluxes in most of the cases, the results from the inversion experiment build confidence in our uncertainty estimations and ensure that the optimization is done correctly. The random error of the retrieved emissions for a single overpass is estimated to be less than 10 MtCO₂ yr⁻¹ for both swath widths. Hence it is expected that given the availability of the high-resolution CarbonSat simulated observations, it is likely to deduce the structural patterns of the emission fluxes. Based on the above analysis, however, no firm conclusion can be made on the magnitude of the retrieved flux uncertainty when prior fluxes significantly deviate from true fluxes in representing the structural variations of emissions. For this purpose, a more sophisticated inverse methodology involving additional extended state vectors and calculation of the response function of the elements of the state vector (adjoint calculation) is required. Since we use the same

transport model to generate the (pseudo) observations and the influence functions, the inversion results shown here may be slightly optimistic. Although it is not within the scope of this study, the transport-related errors are expected to be non-negligible and should be properly addressed in the inverse modeling applications of satellite data.

5 Using the dense CarbonSat measurements in an inverse modeling framework at high resolution is expected to improve the inference of CO₂ fluxes by disentangling different sources of variations. But to what extent one can differentiate regional contributions from different sources should be investigated in further detail.

10 Overall, the present study demonstrates that an instrument like CarbonSat has high potential to provide important information on city CO₂ emissions when exploiting the atmospheric XCO₂ observations using a high-resolution inverse modeling system. Utilizing these measurements together with in-situ, airborne and other satellite measurements is expected to provide more detailed and reliable information on natural and anthropogenic fluxes, facilitating the monitoring of future climate mitigation strategies.

15 **Acknowledgements.** We thank all principle investigators involved in the eddy covariance measurements, and all scientists involved in the L4 eddy covariance dataset that has been accessed from <http://www.europe-fluxdata.eu/>. We would also like to thank François-Marie Bréon for his helpful suggestions and careful review. This study has received funding from ESA (projects LOGOFLUX-I and LOGOFLUX-II) and the State and the University of Bremen.

References:

- Amstel, A. Van, Olivier, J. and Janssen, L.: Analysis of differences between national inventories and an Emissions Database for Global Atmospheric Research (EDGAR), *Environ. Sci. Policy*, 2(3), 275–293, doi:10.1016/S1462-9011(99)00019-2, 1999.
- 5 Beck, V., Koch, T., Kretschmer, R., Marshall, J., Ahmadov, R., Gerbig, C., Pillai, D., and Heimann, M.: The WRF Greenhouse Gas Model (WRF-GHG), Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany, [online] Available from: <http://www.bgc-jena.mpg.de/bgc-systems/index.shtml>, 2011.
- Berezin, E. V., Konovalov, I. B., Ciais, P., Richter, A., Tao, S., Janssens-Maenhout, G., Beekmann, M., and Schulze, E.-D.: Multiannual changes of CO₂ emissions in China: indirect estimates derived from satellite
10 measurements of tropospheric NO₂ columns, *Atmos. Chem. Phys.*, 13, 9415–9438, doi:10.5194/acp-13-9415-2013, 2013.
- Bergeron, O. and Strachan, I. B.: CO₂ sources and sinks in urban and suburban areas of a northern mid-latitude city, *Atmos. Environ.*, 45(8), 1564–1573, doi:10.1016/j.atmosenv.2010.12.043, 2011.
- 15 Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V. and Goede, a. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, *J. Atmos. Sci.*, 56(2), 127–150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2, 1999.
- 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 CO₂ emissions from space and related applications, *Atmos. Meas. Tech.*, 3(4), 781–811, doi:10.5194/amt-3-781-2010, 2010.
- 20 Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O. and Ciais, P.: An attempt at estimating Paris area CO₂ emissions from atmospheric concentration measurements, *Atmos. Chem. Phys.*, 15(4), 1707–1724, doi:10.5194/acp-15-1707-2015, 2015.
- 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):
25 Assessment of atmospheric CO₂ and CH₄ retrieval errors by error parameterization, *Atmos. Meas. Tech.*, 6(12), 3477–3500, doi:10.5194/amt-6-3477-2013, 2013a.
- 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):
30 assessment of scattering related atmospheric CO₂ and CH₄ retrieval errors and first results on implications for inferring city CO₂ emissions, *Atmos. Meas. Tech. Discuss.*, 6, 4769–4850, doi:10.5194/amtd-6-4769-2013, 2013b.
- Burrows, J. P., Hölzle, E., Goede, a. P. H., Visser, H. and Fricke, W.: SCIAMACHY-scanning imaging absorption spectrometer for atmospheric cartography, *Acta Astronaut.*, 35(7), 445–451, doi:10.1016/0094-5765(94)00278-T, 1995.
- 35 Canadell, J. G., Ciais, P., Dhakal, S., Dolman, H., Friedlingstein, P., Gurney, K. R., Held, A., Jackson, R. B., Le Quéré, C., Malone, E. L., Ojima, D. S., Patwardhan, A., Peters, G. P. and Raupach, M. R.: Interactions of the carbon cycle, human activity, and the climate system: a research portfolio, *Curr. Opin. Environ. Sustain.*, 2(4), 301–311, doi:<http://dx.doi.org/10.1016/j.cosust.2010.08.003>, 2010.
- Corbin, K. D., Denning, a. S. and Gurney, K. R.: The space and time impacts on U.S. regional atmospheric CO₂ concentrations from a high resolution fossil fuel CO₂ emissions inventory, *Tellus B*, 62(5), 506–511,
40 doi:10.1111/j.1600-0889.2010.00480.x, 2010.

- 5 Crisp, D., Atlas, R. M., Breon, F. M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. J., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D. and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, *Adv. Sp. Res.*, 34(4), 700–709, doi:10.1016/j.asr.2003.08.062, 2004.
- EPA, Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2014, Office of Atmospheric Programs, Washington, D.C., [online] Available from: <https://www3.epa.gov/climatechange/ghgemissions/usinventoryreport.html>, 2016
- 10 ESA: Report for Mission Selection: CarbonSat, ESA SP-1330/1 (2 volume series), Noordwijk, The Netherlands. [online] Available from: http://esamultimedia.esa.int/docs/EarthObservation/SP1330-1_CarbonSat.pdf, 2015.
- European Commission: Commission Regulation (EU) No 601/2012 of 21 June 2012 on the monitoring and reporting of greenhouse gas emissions pursuant to Directive 2003/87/EC of the European Parliament and of the Council Text with EEA relevance, *Official Journal of the European Union L 181*, 12.7.2012, p. 30–104, [online] Available from <http://data.europa.eu/eli/reg/2012/601/oj>, 2012.
- 15 Göckede, M., Michalak, A. M., Vickers, D., Turner, D. P. and Law, B. E.: Atmospheric inverse modeling to constrain regional-scale CO₂ budgets at high spatial and temporal resolution, *J. Geophys. Res. Atmos.*, 115(15), doi:10.1029/2009JD012257, 2010.
- Gregg, J. S., Andres, R. J. and Marland, G.: China: Emissions pattern of the world leader in CO₂ emissions from fossil fuel consumption and cement production, *Geophys. Res. Lett.*, 35(8), doi:10.1029/2007GL032887, 2008.
- 20 Gurney, K. R., Law, R. M., Denning, a S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T. and Yuen, C.-W.: Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models., *Nature*, 415(6872), 626–630, doi:10.1038/415626a, 2002.
- 25 Gurney, K. R., Chen, Y. H., Maki, T., Kawa, S. R., Andrews, A. and Zhu, Z.: Sensitivity of atmospheric CO₂ inversions to seasonal and interannual variations in fossil fuel emissions, *J. Geophys. Res. D Atmos.*, 110(10), 1–13, doi:10.1029/2004JD005373, 2005.
- 30 Hase, F., Frey, M., Blumenstock, T., Groß, J., Kiel, M., Kohlhepp, R., Mengistu Tsidu, G., Schäfer, K., Sha, M. K. and Orphal, J.: Application of portable FTIR spectrometers for detecting greenhouse gas emissions of the major city Berlin, *Atmos. Meas. Tech.*, 8(7), 3059–3068, doi:10.5194/amt-8-3059-2015, 2015.
- Heimann, M. and Körner, S.: The Global Atmospheric Tracer Model TM3, Technical Report No. 5, Max Planck Institute for Biogeochemistry, Jena, Germany, [online] Available from: <http://www.bgc-jena.mpg.de/bgc-systems/index.shtml>, 2003
- 35 Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H. and Courrèges-Lacoste, G. B.: Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and -5p, *Remote Sens. Environ.*, 120, 58–69, doi:http://dx.doi.org/10.1016/j.rse.2012.01.023, 2012.
- IPCC, IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds).Published: IGES, Japan, 2006.
- 40 Keppel-Aleks, G., Wennberg, P. O., O'Dell, C. W. and Wunch, D.: Towards constraints on fossil fuel emissions from total column carbon dioxide, *Atmos. Chem. Phys.*, 13(8), 4349–4357, doi:10.5194/acp-13-4349-2013, 2013.

- Kort, E. a., Frankenberg, C., Miller, C. E. and Oda, T.: Space-based observations of megacity carbon dioxide, *Geophys. Res. Lett.*, 39(17), doi:10.1029/2012GL052738, 2013.
- 5 Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, a., Vogel, F., Hammer, S. and Totsche, K. U.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO₂, *Atmos. Chem. Phys.*, 14(14), 7149–7172, doi:10.5194/acp-14-7149-2014, 2014.
- Krings, T., Gerilowski, K., Buchwitz, M., Reuter, M., Tretner, A., Erzinger, J., Heinze, D., Pfluger, 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, 5 doi:10.5194/amt-4-1735-2011, 2011. 4772, 4778, 4792, 4799
- 10 Kuze, A., Suto, H., Nakajima, M. and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring., *Appl. Opt.*, 48(35), 6716–6733, doi:10.1364/AO.48.006716, 2009.
- 15 Levin, I., Hammer, S., Eichelmann, E. and Vogel, F. R.: Verification of greenhouse gas emission reductions: the prospect of atmospheric monitoring in polluted areas., *Philos. Trans. A. Math. Phys. Eng. Sci.*, 369(1943), 1906–1924, doi:10.1098/rsta.2010.0249, 2011.
- Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y. and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO₂ exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochem. Cycles*, 22(2), doi:10.1029/2006GB002735, 2008.
- 20 Marland, G.: Uncertainties in accounting for CO₂ from fossil fuels, *J. Ind. Ecol.*, 12(2), 136–139, doi:10.1111/j.1530-9290.2008.00014.x, 2008.
- 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(20), 7816–7823, doi:10.1021/es901326b, 2009.
- 25 McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R. and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, *Proc. Natl. Acad. Sci.*, 109(22), 8423–8428, doi:10.1073/pnas.1116645109, 2012.
- 30 Newman, S., Jeong, S., Fischer, M. L., Xu, X., Haman, C. L., Lefer, B., Alvarez, S., Rappenglueck, B., Kort, E. a., Andrews, a. E., Peischl, J., Gurney, K. R., Miller, C. E. and Yung, Y. L.: Diurnal tracking of anthropogenic CO₂ emissions in the Los Angeles basin megacity during spring 2010, *Atmos. Chem. Phys.*, 13(8), 4359–4372, doi:10.5194/acp-13-4359-2013, 2013.
- Oda, T. and Maksyutov, S.: A very high-resolution (1km × 1 km) global fossil fuel CO₂ emission inventory derived using a point source database and satellite observations of nighttime lights, *Atmos. Chem. Phys.*, 11(2), 543–556, doi:10.5194/acp-11-543-2011, 2011.
- 35 Pacala, S. W., Breidenich, C., Brewer, P. G., Fung, I., Gunson, M. R., Heddle, G., Law, B., Marland, G., Paustian, K., Prather, M., Randerson, J. T., Tans, P., Wofsy, S. C., Linn, A. M., Sturdivant, J. and Al., E.: *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements*, The National Academies Press. [online] Available from: <http://www.nap.edu/catalog/12883.html>, 2010.
- 40 Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., and Karstens, U.: High resolution modeling of CO₂ over Europe: implications for representation errors of satellite retrievals, *Atmos. Chem. Phys.*, 10, 83–94, doi:10.5194/acp-10-83-2010, 2010.

- Pillai, D., Gerbig, C., Ahmadov, R., Rödenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B. and Lavrié, J. V.: High-resolution simulations of atmospheric CO₂ over complex terrain-representing the Ochsenkopf mountain tall tower, *Atmos. Chem. Phys.*, 11(15), 7445–7464, doi:10.5194/acp-11-7445-2011, 2011.
- 5 Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B. and Heimann, M.: Comparing Lagrangian and Eulerian models for CO₂ transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, *Atmos. Chem. Phys. Discuss.*, 12(1), 1267–1298, doi:10.5194/acpd-12-1267-2012, 2012.
- 10 Pregger, T., Scholz, Y. and Friedrich, R.: Documentation of the anthropogenic GHG emission data for Europe provided in the frame of CarboEurope GHG and CarboEurope IP, Final Report, University of Stuttgart, IER – Institute of Energy Economics and the Rational Use of Energy, [online] Available from: http://carboeurope.org/ceip/products/files/Pregger_IER_Final_Report_Feb2007.pdf, 2007
- Reuter, M., Buchwitz, M., Hilboll, A., Richter, A., Schneising, O., Hilker, M., Heymann, J., Bovensmann, H. and Burrows, J. P.: Decreasing emissions of NO_x relative to CO₂ in East Asia inferred from satellite observations, *Nat. Geosci.*, 7(11), 792–795 [online] Available from: <http://dx.doi.org/10.1038/ngeo2257>, 2014.
- 15 Rödenbeck, C.: Estimating CO₂ sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport. [online] Available from: http://www.bgc-jena.mpg.de/uploads/Publications/TechnicalReports/tech_report6.pdf, 2005.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics--Vol. 2, World Scientific. [online] Available from: <http://books.google.de/books?id=9IoscAAACAAJ>, 2000.
- 20 Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Reuter, M., Notholt, J., Macatangay, R., and Warneke, T.: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite – Part 1: Carbon dioxide, *Atmos. Chem. Phys.*, 8, 3827–3853, doi:10.5194/acp-8-3827-2008, 2008.
- 25 Schneising, O., Heymann, J., Buchwitz, M., Reuter, M., Bovensmann, H. and Burrows, J. P.: Anthropogenic carbon dioxide source areas observed from space: Assessment of regional enhancements and trends, *Atmos. Chem. Phys.*, 13(5), 2445–2454, doi:10.5194/acp-13-2445-2013, 2013.
- Silva, S. J., Arellano, A. F. & Worden, H. M: Toward anthropogenic combustion emission constraints from space-based analysis of urban CO₂/CO sensitivity, *Geophys. Res. Lett.*, 40, 4971–4976, 2013.
- 30 Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C. and Mukai, H.: The CO₂ release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: Effects from varying oxidative ratios, *Atmos. Chem. Phys.*, 11(14), 6855–6870, doi:10.5194/acp-11-6855-2011, 2011.
- Tolk, L. F., Meesters, A. G. C. A., Dolman, A. J. and Peters, W.: Modelling representation errors of atmospheric CO₂ mixing ratios at a regional scale, *Atmos. Chem. Phys.*, 8(22), 6587–6596, doi:10.5194/acp-8-6587-2008, 2008.
- 35 Wang, Y., Munger, J. W., Xu, S., McElroy, M. B., Hao, J., Nielsen, C. P. and Ma, H.: CO₂ and its correlation with CO at a rural site near Beijing: Implications for combustion efficiency in China, *Atmos. Chem. Phys.*, 10(18), 8881–8897, doi:10.5194/acp-10-8881-2010, 2010.
- White, T., Jonas, M., Nahorski, Z. and Nilsson, S.: Greenhouse gas inventories: dealing with uncertainty, edited by J. M. N. Z. N. S. White Th., Springer, Dordrecht, Netherlands. [online] Available from: <http://books.google.com/books?hl=en&lr=&id=jfNFYeWVxhQC&oi=fnd&pg=PA2&dq=Greenhouse+Gas+Inventory+-+Dealing+with+uncertainties&ots=O2K-rYp4Xy&sig=V0eEblUdPkcVIfamIlgQKh6SIGGk>, 2011.

Winker, D. M., Vaughan, M. a., Omar, A., Hu, Y., Powell, K. a., Liu, Z., Hunt, W. H. and Young, S. a.: Overview of the CALIPSO mission and CALIOP data processing algorithms, *J. Atmos. Ocean. Technol.*, 26(11), 2310–2323, doi:10.1175/2009JTECHA1281.1, 2009.

5 Wong, K. W., Fu, D., Pongetti, T. J., Newman, S., Kort, E. A., Duren, R., Hsu, Y.-K., Miller, C. E., Yung, Y. L. and Sander, S. P.: Mapping CH₄ : CO₂ ratios in Los Angeles with CLARS-FTS from Mount Wilson, California, *Atmos. Chem. Phys.*, 15(1), 241–252, doi:10.5194/acp-15-241-2015, 2015.

Zimnoch, M., Godlowska, J., Necki, J. M. and Rozanski, K.: Assessing surface fluxes of CO₂ and CH₄ in urban environment: A reconnaissance study in Krakow, Southern Poland, *Tellus, Ser. B Chem. Phys. Meteorol.*, 62(5), 573–580, doi:10.1111/j.1600-0889.2010.00489.x, 2010.

Table 1: Overview of different scenarios, SCE, which are used to investigate the systematic errors of the retrieved emissions. The absolute mean and standard deviations are estimated for two swath widths (SW-500: 500 km and SW-240: 240 km) for all “N” useful overpasses and are expressed in both MtCO₂ yr⁻¹ and in %. Err-L, Err-H, and Err-B indicate errors attributed to CarbonSat measurement, high-resolution aerosol related errors, and biogenic modeling errors respectively. Err-Emi indicates whether inversion experiment uses different prior emission fluxes (see Sec. 4.3)

| SCE | Err-L | Err-H | Err-B | Err-Emi | Prior flux | True flux | SE (SW-500) (mean ±std) % | SE (SW-500) (mean ±std) MtCO ₂ yr ⁻¹ | SE (SW-240) (mean ±std) % | SE (SW-240) (mean ±std) MtCO ₂ yr ⁻¹ |
|-----|-------|-------|-------|---------|------------|-----------|------------------------------------|---|------------------------------------|---|
| S01 | ✓ | | | | EDGAR | EDGAR | -5.3 ± 6.1 | -2.5 ± 2.8 | -6.1 ± 5.7 | -2.8 ± 2.6 |
| S02 | | ✓ | | | EDGAR | EDGAR | -7.5 ± 3.4 | -3.6 ± 2.1 | -7.5 ± 2.9 | -3.1 ± 1.6 |
| S03 | | | ✓ | | EDGAR | EDGAR | 18.5 ± 23.3 | 7.5 ± 9.5 | 20.5 ± 23.7 | 8.2 ± 9.4 |
| S04 | ✓ | ✓ | | | EDGAR | EDGAR | -12.7 ± 7.4 | -6.1 ± 3.8 | -12.8 ± 7.1 | -5.9 ± 3.6 |
| S05 | ✓ | ✓ | ✓ | | EDGAR | EDGAR | 5.8 ± 24.4 | 1.5 ± 10.7 | 7.7 ± 25.3 | 2.3 ± 10.6 |
| S06 | ✓ | | ✓ | | EDGAR | EDGAR | 13.2 ± 23.7 | 5.1 ± 10.0 | 14.4 ± 24.5 | 5.4 ± 10.0 |
| S07 | | | | ✓ | IER | EDGAR | 10.1 ± 15.5 | 4.5 ± 6.9 | 10.1 ± 15.9 | 4.5 ± 7.1 |
| S08 | ✓ | | | ✓ | IER | EDGAR | 6.3 ± 16.2 | 2.8 ± 7.1 | 5.9 ± 18.2 | 2.6 ± 7.9 |
| S09 | ✓ | ✓ | | ✓ | IER | EDGAR | -0.5 ± 16.0 | -0.2 ± 6.9 | -0.9 ± 18.2 | -0.4 ± 7.6 |
| S10 | ✓ | | ✓ | ✓ | IER | EDGAR | 24.3 ± 27.7 | 10.8 ± 11.0 | 24.3 ± 29.1 | 10.8 ± 11.4 |
| S11 | ✓ | ✓ | ✓ | ✓ | IER | EDGAR | 17.6 ± 27.9 | 7.8 ± 11.1 | 17.6 ± 29.6 | 7.8 ± 11.6 |

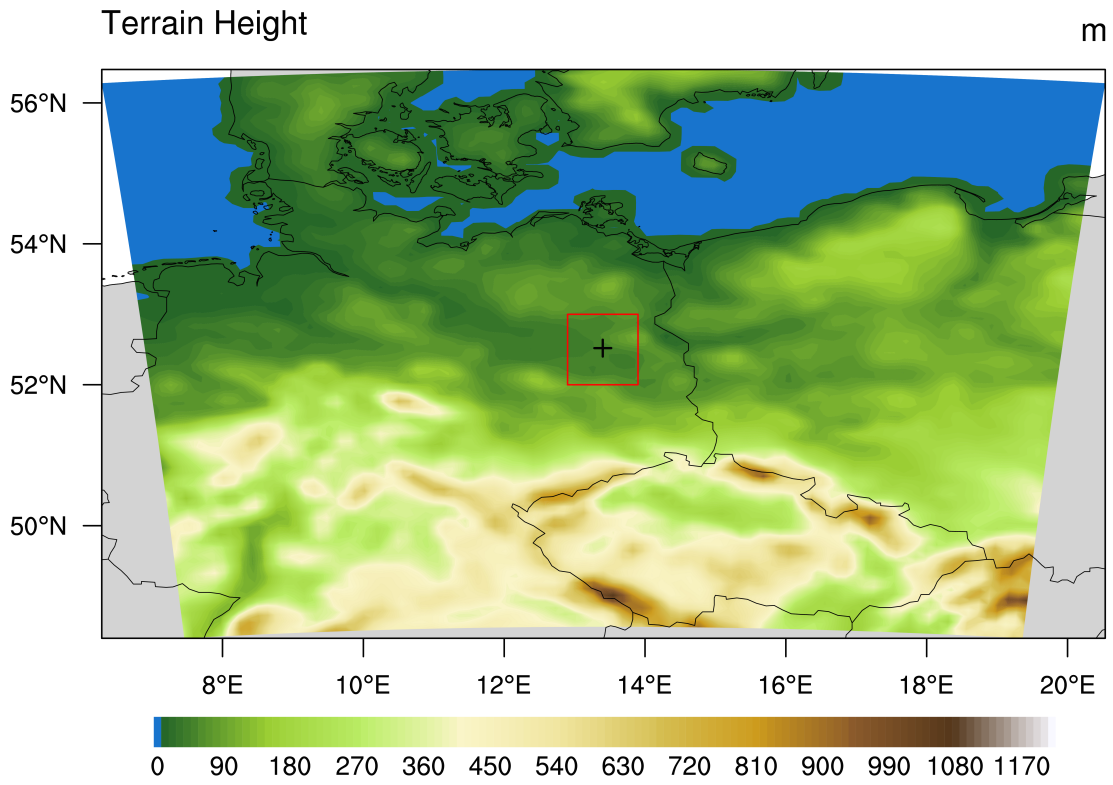


Figure 1: The Berlin-centered WRF-GHG model domain in Lambert Conformal Conic projection used in the study. The red rectangle represents the target region (100 km x 100 km) described in the Sec. 3.2 and the + sign indicates the central location of Berlin city. The color bar indicates the terrain height in meters.

5

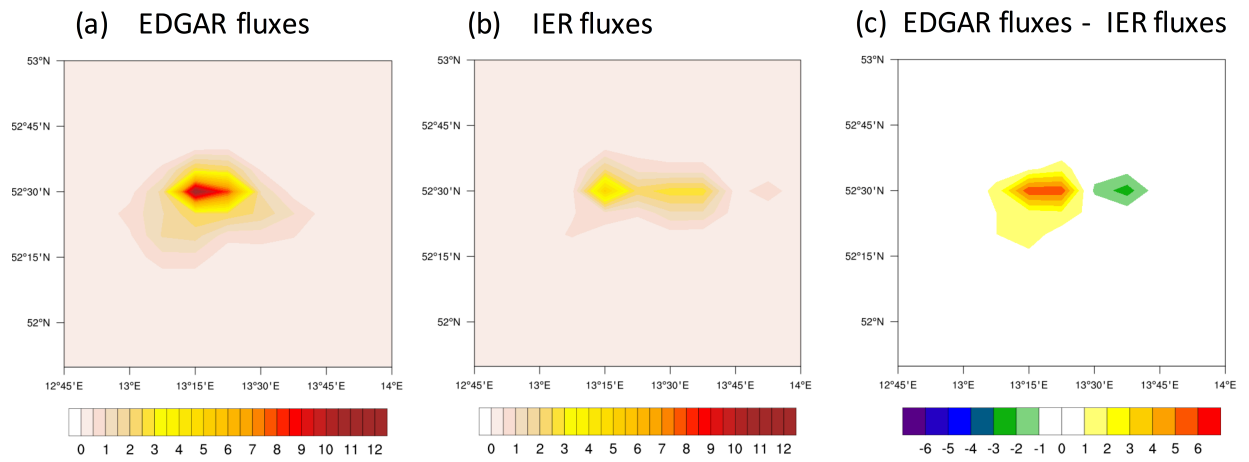


Figure 2: Annual averages of fossil fuel combustion emission fluxes at $10 \text{ km} \times 10 \text{ km}$, zoomed over the Berlin region: (a) EDGAR emissions, (b) IER emissions, and (c) the difference between EDGAR and IER emissions (EDGAR - IER). All units are in MtCO_2 per year per grid cell.

5

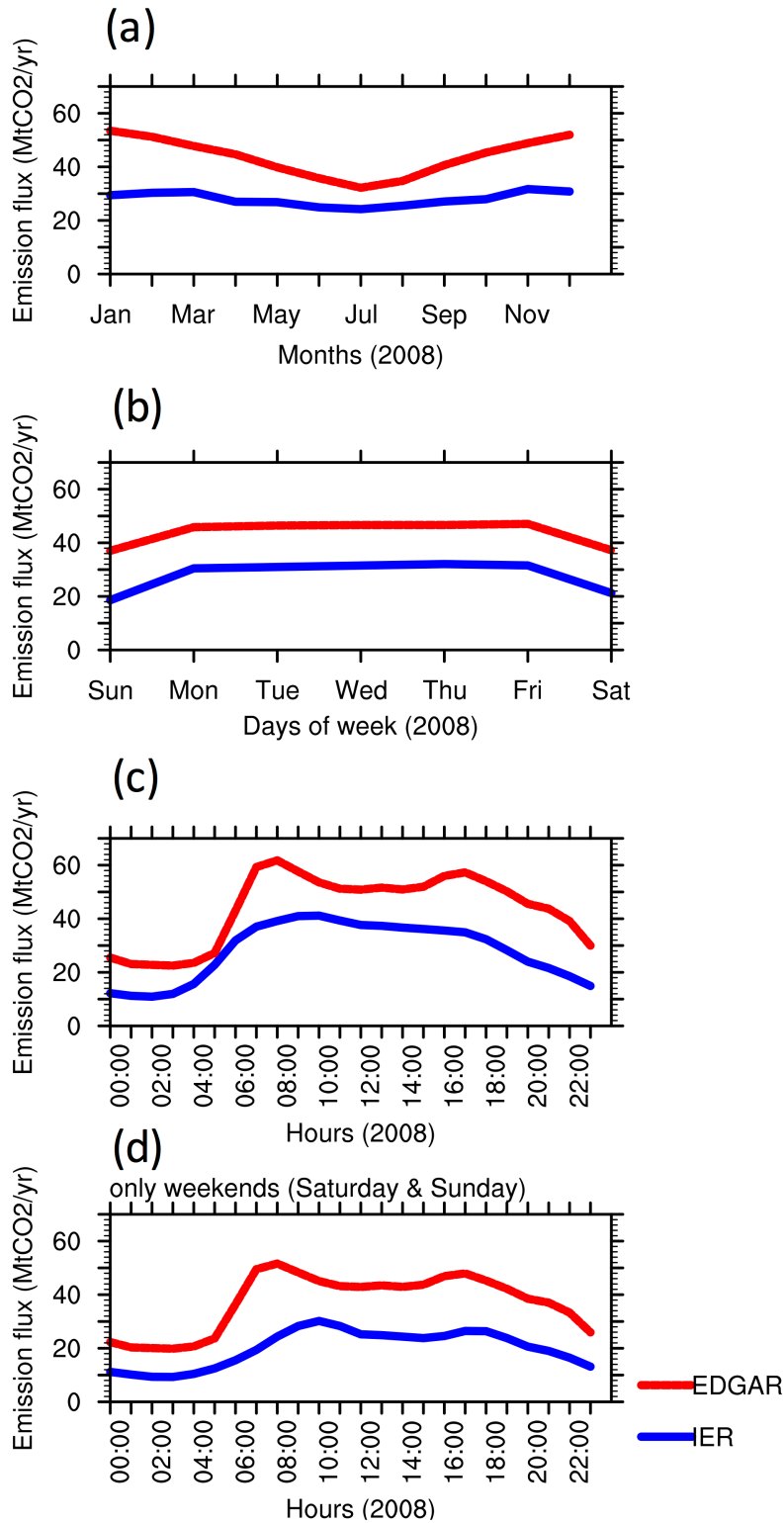


Figure 3: Temporal variability of EDGAR and IER emission fluxes, aggregated over the target region around Berlin (~100 km × 100 km) averaged for different time scales for the year 2008: (a) monthly, (b) weekly and (c & d) hourly. The 4th panel (d) shows the values representing only weekends, while the 3rd panel (c) represents all days of the week. Hours are in UTC (local time CET = UTC+1).

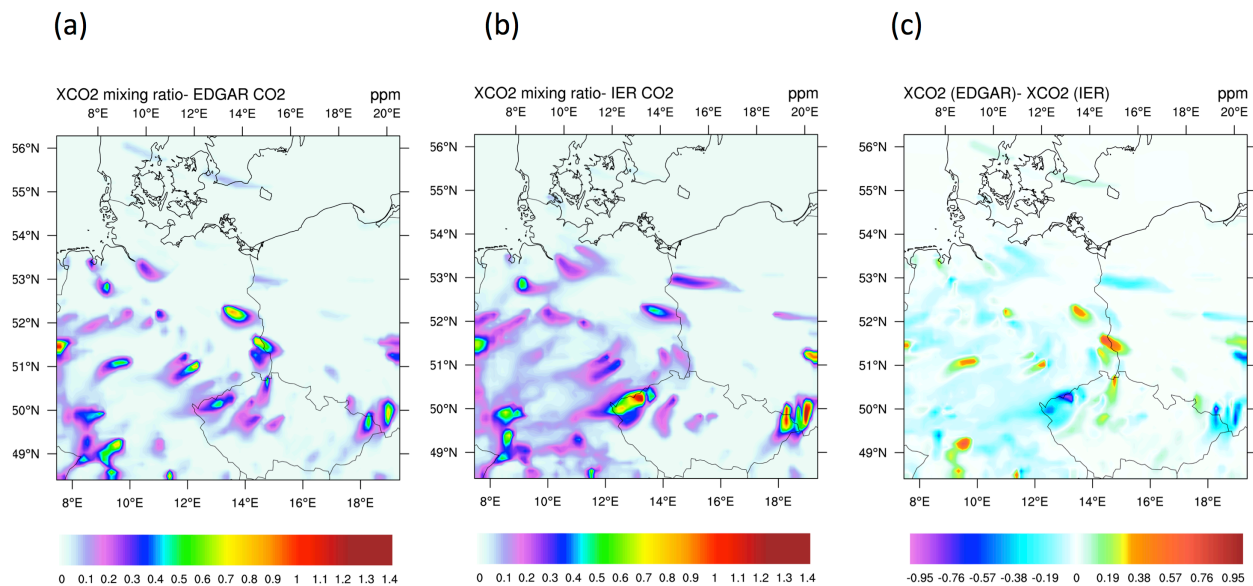


Figure 4: Anthropogenic XCO₂ enhancement on 24th June 2008 at 10:00 UTC (Local time: 12:00 CEST) (a) “true” XCO₂ enhancement (using EDGAR emissions), and (b) XCO₂ enhancement when using IER emissions. Panel (c) shows the discrepancy in XCO₂ enhancement due to the difference between EDGAR and IER emission inventories. All units are in ppm.

5

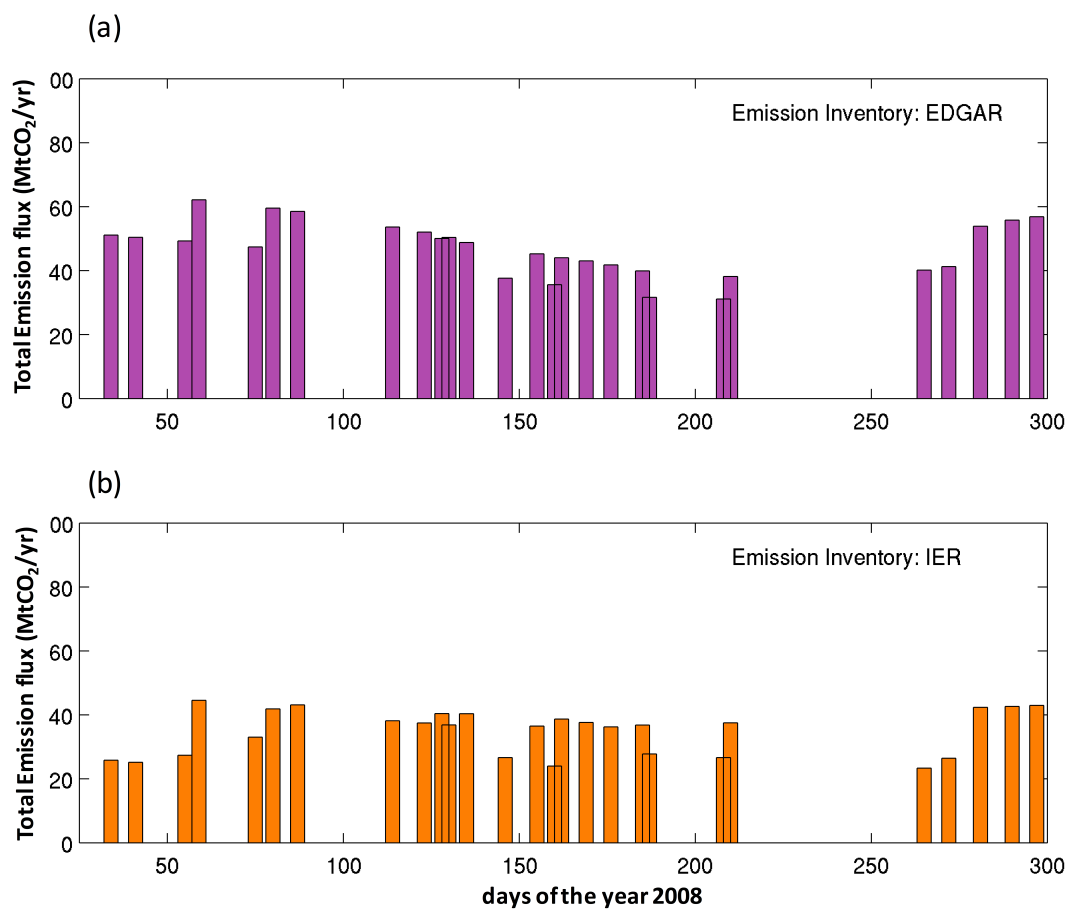


Figure 5: Anthropogenic flux over the target region based on (a) EDGAR inventory, and (b) IER inventory for all of CarbonSat’s “useful overpasses” corresponding to 500km swath width for the year 2008.

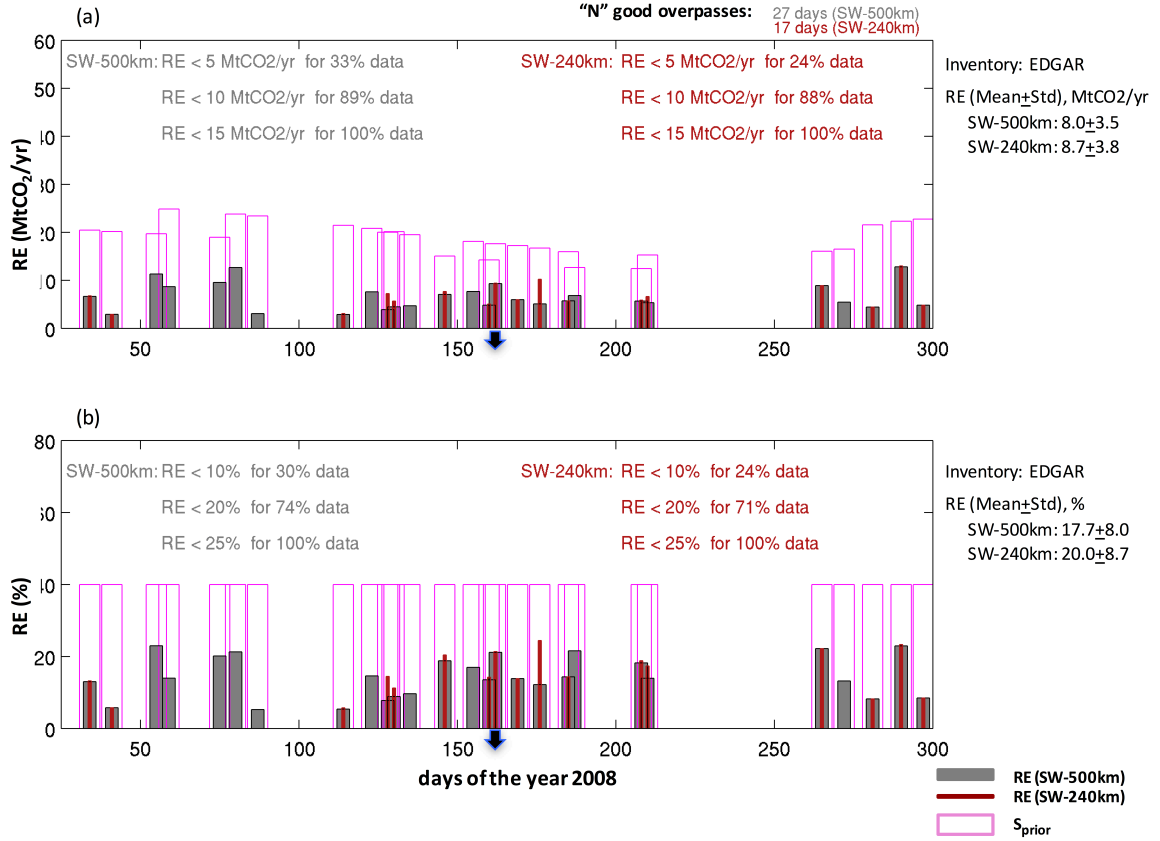


Figure 6: Precision (random errors (RE)) of the retrieved emission fluxes obtained by the inverse optimization using one year of CarbonSat simulated observations. Results of two different swath widths (SW) – 500 km (grey) and 240 km (red) – are shown. S_{prior} values are indicated with magenta bordered bars for visualizing the reduction in uncertainty. The top and bottom panels show RE in MtCO₂ yr⁻¹ and in % respectively. An overview of the statistical distribution of RE, separately for 500 km (grey) and 240 km (red) swath widths, is given inside the panel. The overall mean ± standard deviation is given outside the respective panels. The lower and upper limits of the X-axis (days of the year) is restricted accordingly as there are no good CarbonSat simulated observation during winter months. The arrow marker in the X-axis indicates a particular day (24th June 2008) shown in Figs. 4, 7, and 8.

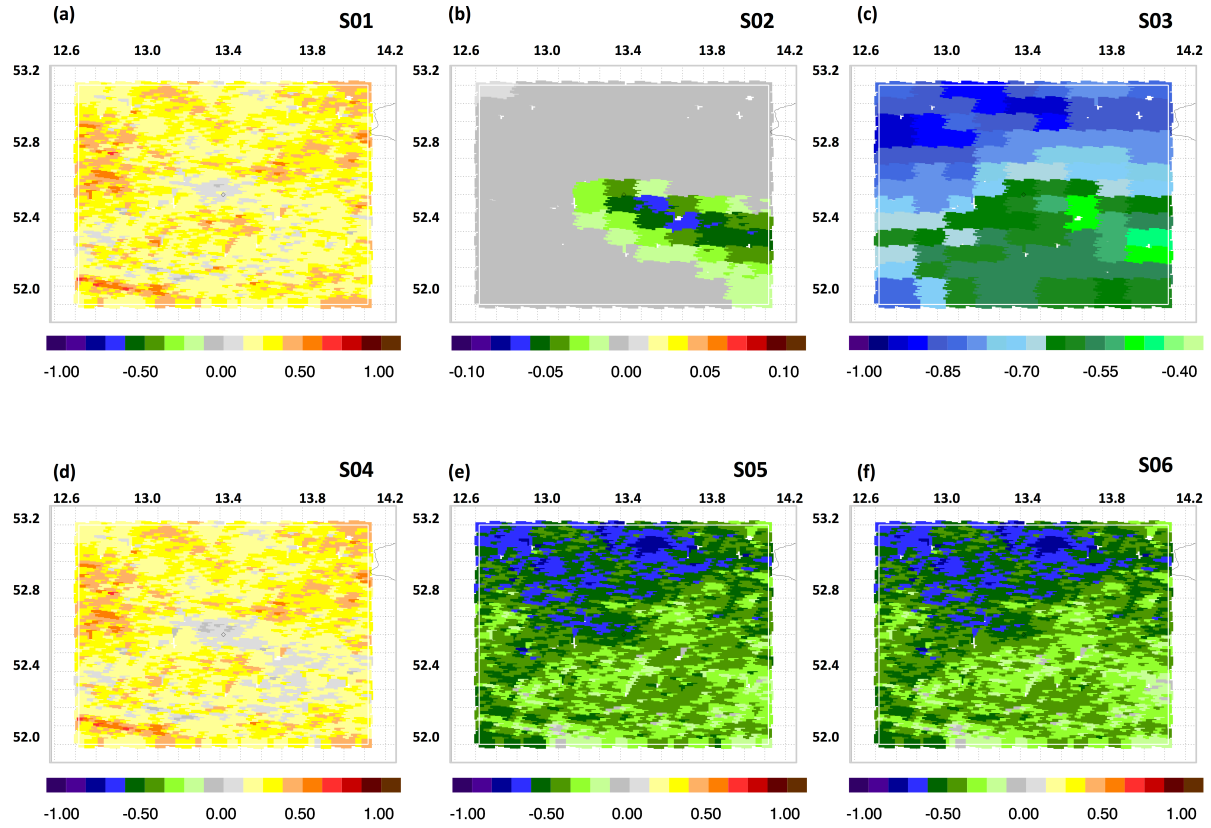


Figure 7: XCO₂ systematic error over the target region on 24th June 2008, assuming a CarbonSat swath width of 500 km. The six scenarios (S01 to S06) are shown with a label inside the respective panel. For S01, S02 and S04, these errors are estimated using the error parameterization scheme of Buchwitz et al. (2013a). The other scenarios additionally utilize biogenic XCO₂ variability in the target region (simulated by WRF-GHG) to derive XCO₂ systematic errors. Note that different color scales are used for S02 and S03. All units are given in ppm.

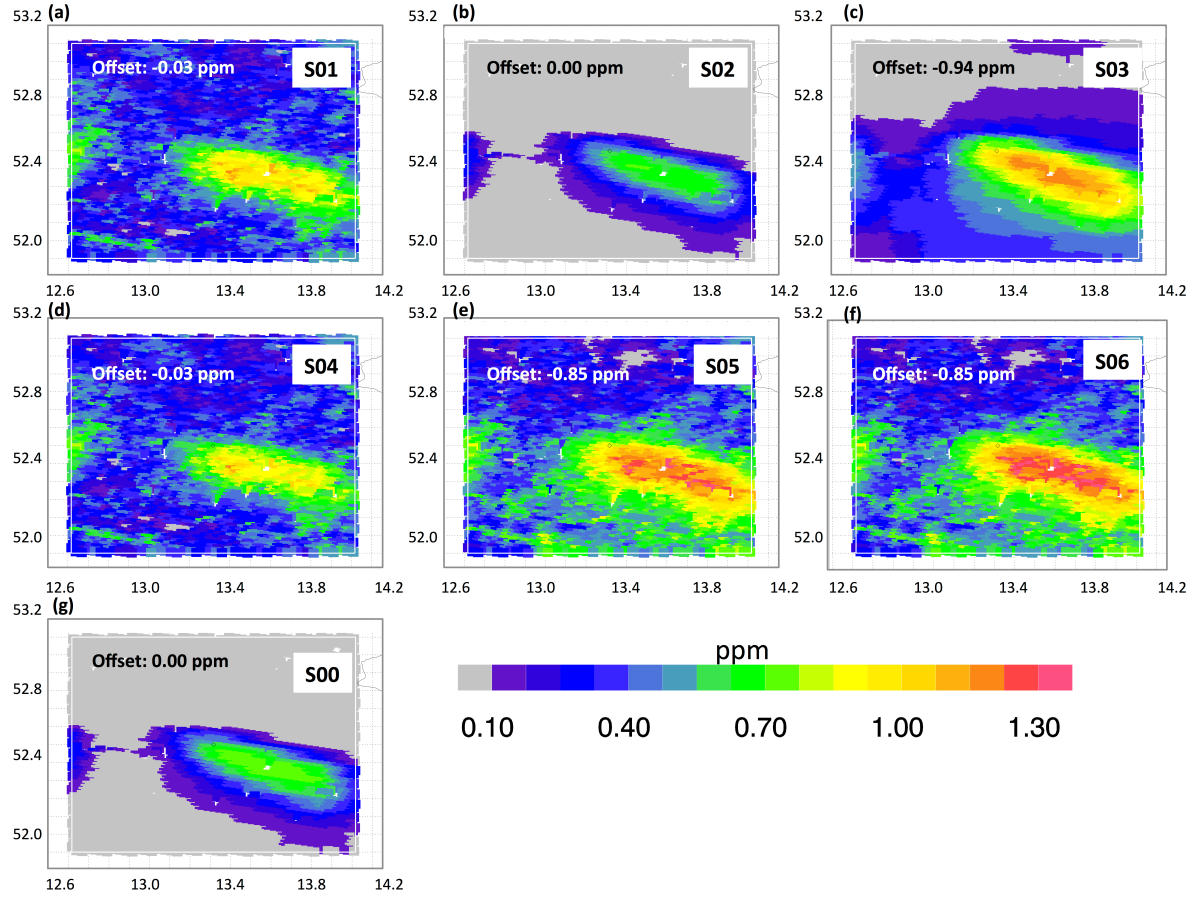


Figure 8: Observed anthropogenic XCO₂ enhancement over the target region during a CarbonSat overpass on 24th June 2008 (swath width: 500 km). Different panels show anthropogenic XCO₂ enhancement, while considering XCO₂ systematic errors for different scenarios as shown in Fig.8. The “True” XCO₂ (fossil fuel (FF)) enhancement (i.e. without any uncertainties) is given in the bottom panel (g) for comparison. Note that an offset, labeled inside each panel, is subtracted from the anthropogenic XCO₂ enhancement to better visualize the details (for the figure only). All units are ppm.

5

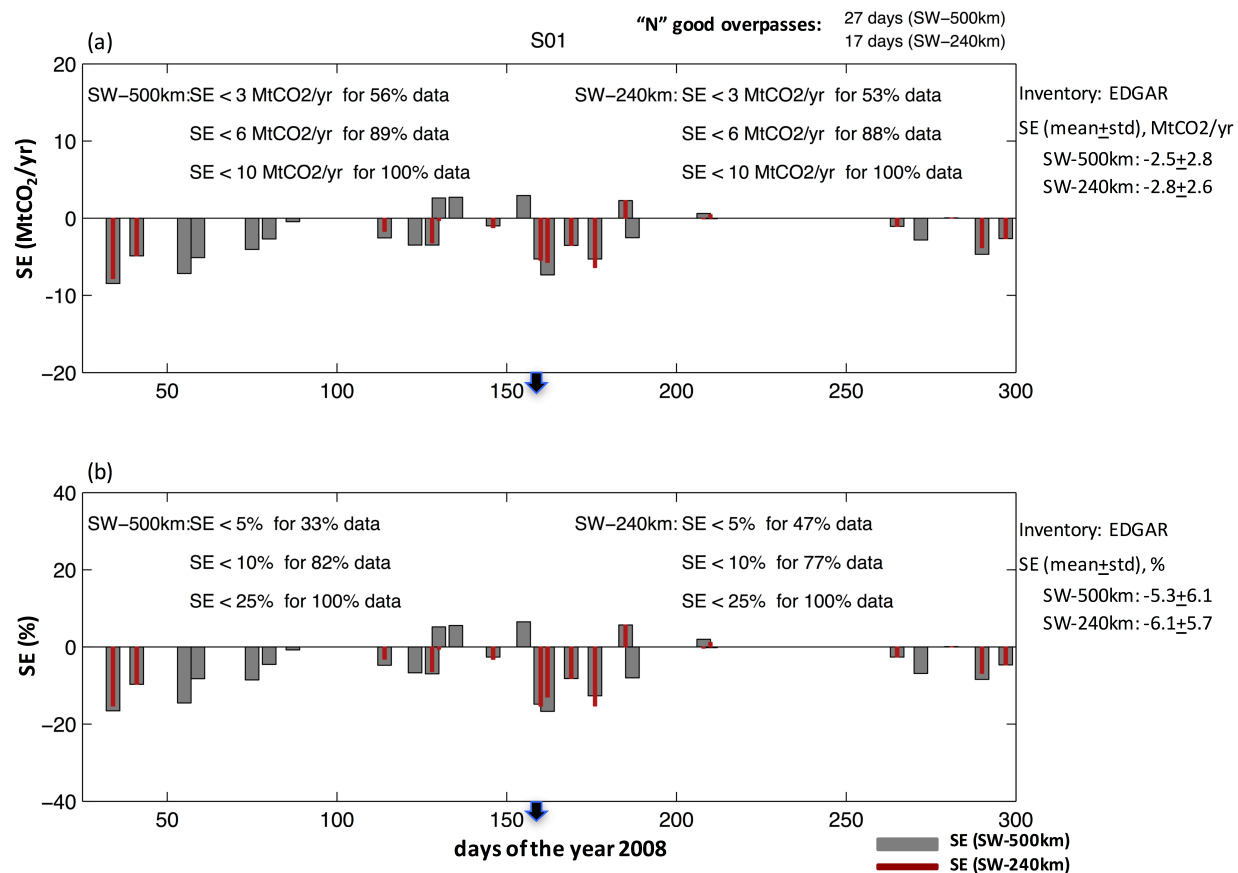


Figure 9: Systematic errors (SE) of the retrieved emission fluxes for S01, obtained by the inverse optimization using one year of CarbonSat simulated observations. Results of two different swath widths (SW) – 500 km (grey) and 240 km (red) – are shown. (a) and (b) show SE in MtCO₂ yr⁻¹ and in % respectively. An overview of the statistical distribution of SE, separately for 500 km (grey) and 240 km (red) swath widths, is given inside the panel. The overall mean +/- standard deviation is given outside the respective panels.

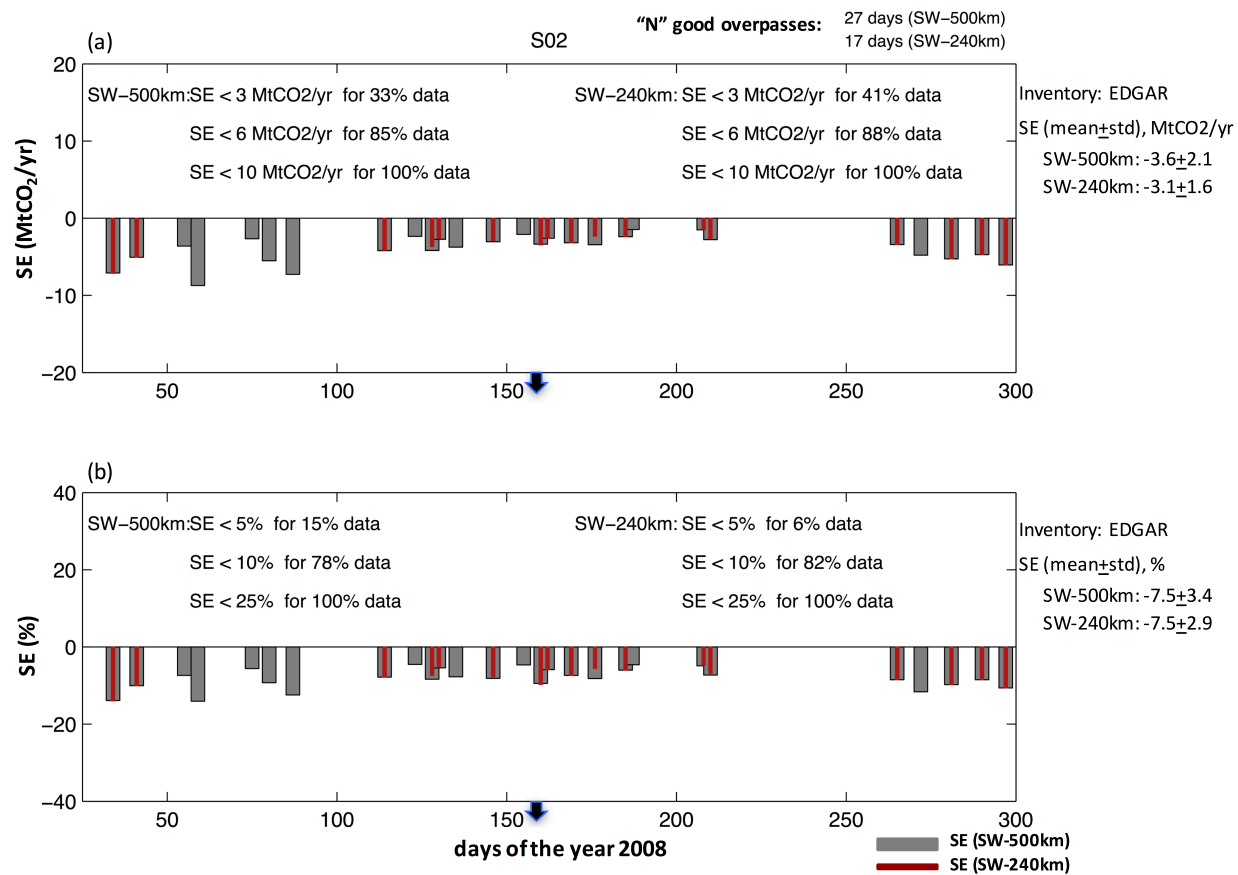


Figure 10: Same as Fig.9, but for S02, quantifying the impact of the worst-case assumption used for aerosol-related biases.

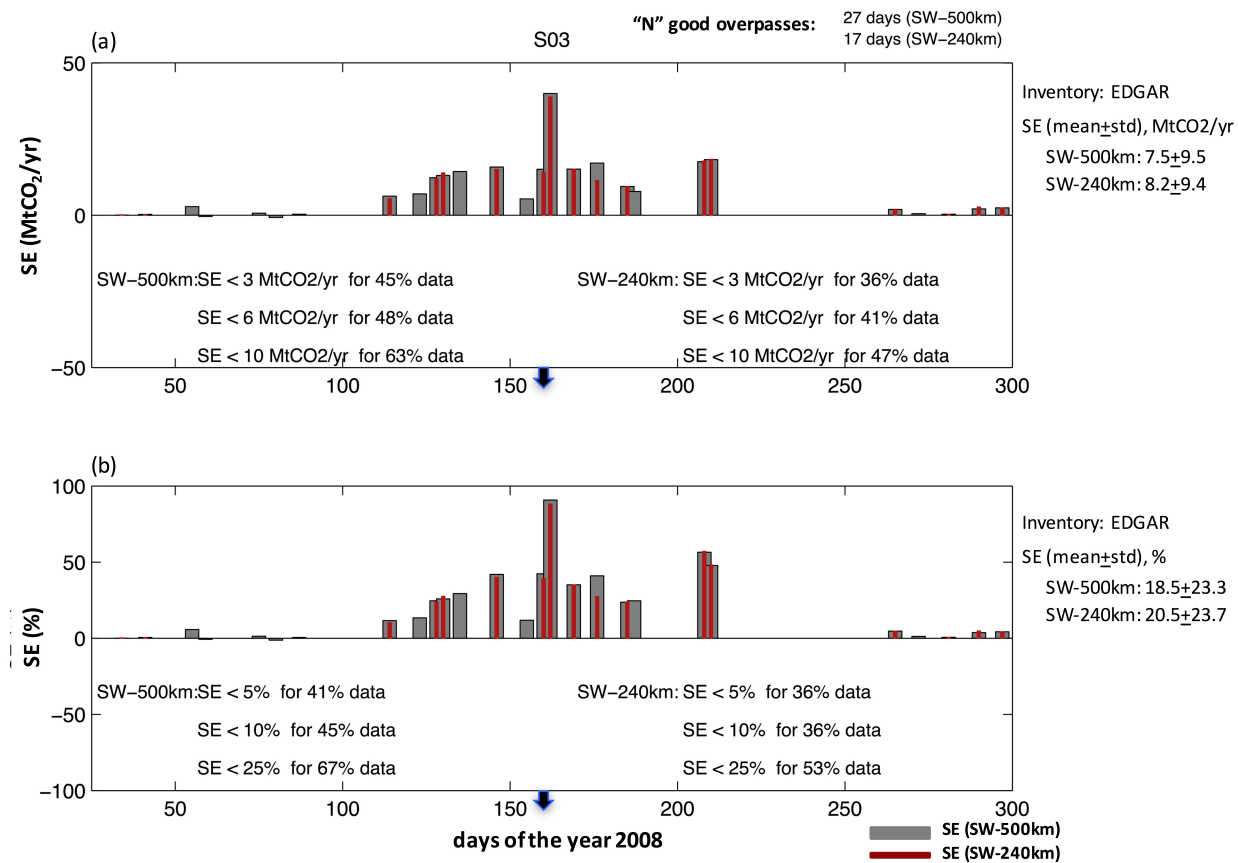


Figure 11: Same as Fig.9, but for S03, quantifying the impact of the worst-case modeling related errors by assuming that biogenic XCO₂ variations cannot be modeled at all.

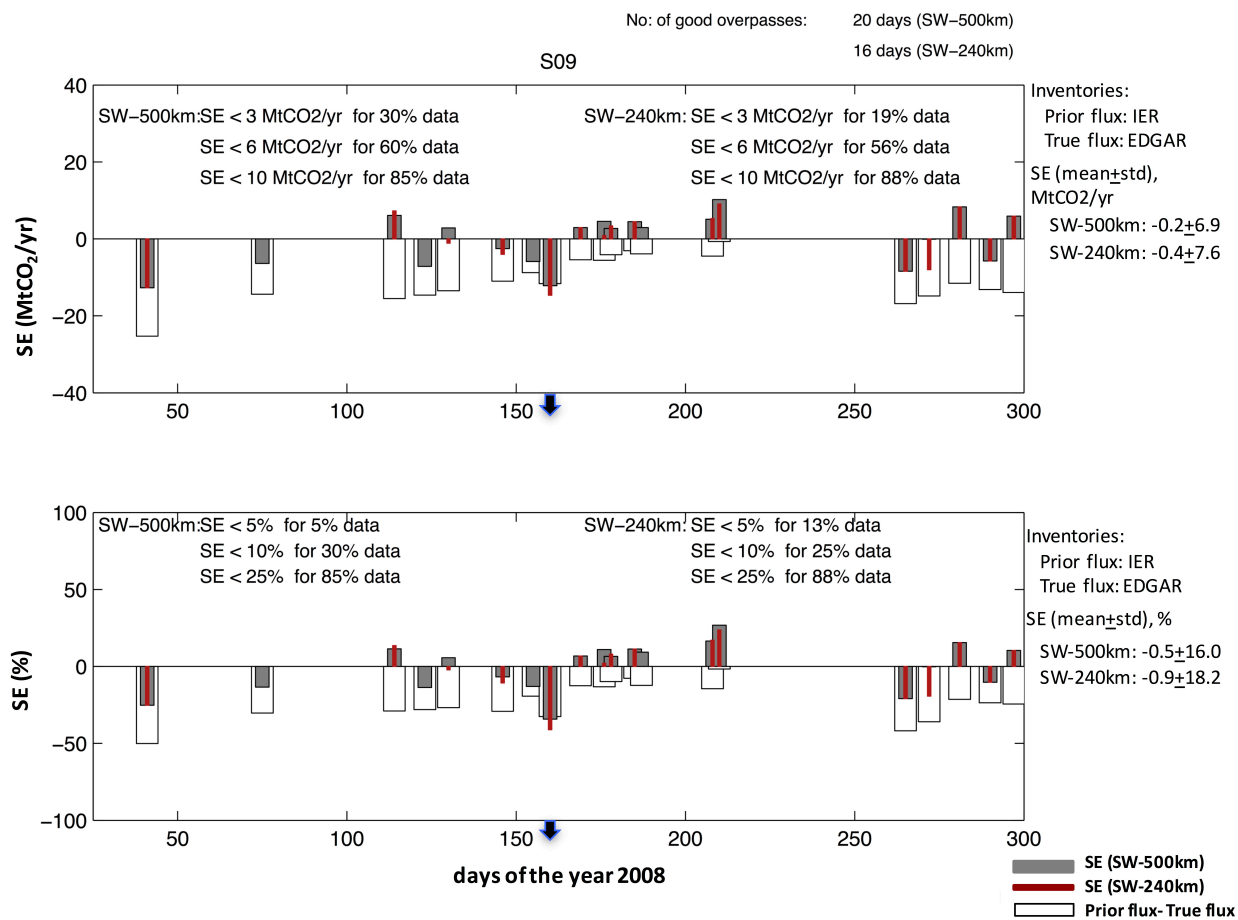


Figure 12: Similar to Fig. 9, but for the inversion experiment S09 using IER (a priori) and EDGAR (true) emission fluxes.