We would like to thank Referee #1 for his/her thoughtful comments and detailed suggestions to our manuscript. In the following, we answer to the reviewer's comments and indicate the changes in the manuscript that were implemented as a consequence of the recommendations. The comments are in black and italic. Our answers are in blue and plain text.

Anonymous Referee #1

Received and published: 5 August 2019

Review of Lian et al. (2019) Analysis of temporal and spatial variability of atmospheric CO_2 concentration within Paris from the GreenLITE™ laser imaging experiment.

Lian et al. describe the application of a long open-path spectroscopy technique for the measurement of CO_2 mixing ratios above a complex urban canopy, which could influence existing emissions estimates at the city scale. They compare the data measured using the GreenLITE™ system with fixed-site CO_2 measurements within the same urban environment, and contrast the results against two urban canopy schemes within the WRF-Chem model.

It is a well written paper which I would recommend for publication in ACP. The content of the paper, which covers greenhouse gas measurements with a possible climate change impact, is relevant to the journal and its readers.

We thank the reviewer for these very supportive comments.

General comments:

The authors acknowledge that calibration of long open-path spectroscopy techniques is difficult. A separate paper (Zaccheo et al., 2019), detailing a new calibration procedure applied to the GreenLITE™ data, is referenced by the authors here. This calibration procedure appears to use the fixed-site installations within the city to calibrate the open-path data. Whilst the authors state this "has no significant impact on chord-to-chord variations", they do not discuss the potential implications of using point-source measurements to adjust area/path averaged measurements. Zaccheo et al. (2019) does go into more detail but considering this is a key element of the calibration procedure, I believe it needs some more attention here.

We thank the reviewer for raising this question of data calibration. As he/she points out, the calibration procedure is described in some detail in (Zaccheo et al., 2019) so that we felt there is no need to go into the same level of detail. As it has been addressed in Zaccheo et al. (2019), while not desirable, it is often necessary to apply post-calibration corrections to such data to rectify residual differences between observation types. We acknowledge nevertheless that the reader may want to see more, and we therefore have provided more information about this calibration method as well as its limitation in the revised version of the manuscript:

"These slowly time-varying differences were most likely due to a slight systematic long-term drift in both the on- and off-line wavelengths as a function of continuous operations. Such drift may induce some non-linear impacts on the measured concentrations. It is therefore more appropriate to adjust the wavelengths rather than to apply a linear calibration to the retrieved concentrations. Unlike in-situ point measurement systems, there is no established method for calibration of long open-path systems to the WMO mole fraction scale used as an international standard for atmospheric CO_2 monitoring (Tans et al., 2011). Therefore, a bias correction method was developed by AER (Zaccheo et al., 2019) for addressing observed slowly drifting biases between the GreenLITE™ prototype system and the two in-situ sensors (CDS and JUS) that are near the GreenLITE™ chords. This method computed a time-varying adjustment to the offline

wavelength based on a non-linear optimization mechanism. This non-linear approach adjusts the GreenLITE™ offline wavelength considering not only the average values of hourly CO_2 concentrations at two in-situ stations, but also the corresponding average temperature, relative humidity, atmospheric pressure along the chord and an optimized online wavelength value during the measurement period. Finally, the median on- and off-line values over a 4-day window was used to recompute the GreenLITE™ data from all chords using a radiative transfer based iterative retrieval scheme based on the LBLRTM model (Clough et al., 2005). Even though this approach is not ideal as the two in-situ stations and the GreenLITE™ system do not sample the exact same area, it does provide a well-defined mechanism that reduces the systematic long-term biases with no significant impact on the chord-to-chord variations."

As most of our analyses focus on the spatial gradient on the concentrations, we feel that the important point of the calibration procedure is that it has no significant impact (and thus no significant uncertainty) on the spatial gradient between chords. (See below, our answers to the comment: Page 4, Line 23: What is meant by "no significant impact" – significant in what way?)

Font sizes in some figures could be larger. Some of the text is hard to read on a computer screen without zooming in.

This suggestion is well taken. We have increased the font size in all figures (see Figure 1 to 7 in the revised manuscript).

The authors should address the following points in a revised manuscript:

Page 3, Line 30: What is meant by 15/100 m above ground level? I assume there are two sampling inlets? This is not made clear.

Yes, the atmospheric CO_2 concentrations at SAC station are measured with two sampling inlets at 15 m and 100 m above ground level, on a tall tower at that location. This has been clarified in the text:

"OVS site is located about 26 km southwest of Paris center with the sampling height of 20 m above the ground level (AGL) on the top of a building. The SAC tall tower is located on the Plateau de Saclay (9.5 km southeast of OVS) with two air inlets placed at 15 m and 100 m AGL respectively."

In addition, we also changed "15/100" to "15 and 100" in Table 1.

Page 4, Line 23: What is meant by "no significant impact" – significant in what way?

We agree that the term "no significant impact" is an overly vague statement and should be clarified more precisely. We have added a figure in the revised supplement (Figure S2). It shows the distributions of the original and re-processed GreenLITE™ absolute CO_2 concentration differences between all pairs of chords for each transceiver. The differences between the medians of the inter-chord range of the re-processed and original data are within in the range of ±0.5 ppm for T1 and ±2 ppm for T2 with the respective yearly mean plus/minus one standard deviation of 0.04 ± 0.16 ppm for T1 and 0.48 ± 0.43 ppm for T2, which are relatively small. The modified text in the revised manuscript is as follows:

"Top panels in Figure S2 (a) and (b) show the distribution of the absolute values of the daily averaged CO_2 concentration difference between all pairs of chords for each transceiver before and after the calibration. The differences between the medians of the re-processed and original inter-chord range, shown in bottom panels, are within in the range of ±0.5 ppm for T1 and ±2 ppm for T2 with the respective yearly mean plus/minus one standard deviation of 0.04 ± 0.16 ppm for T1 and 0.48 ± 0.43 ppm for T2."

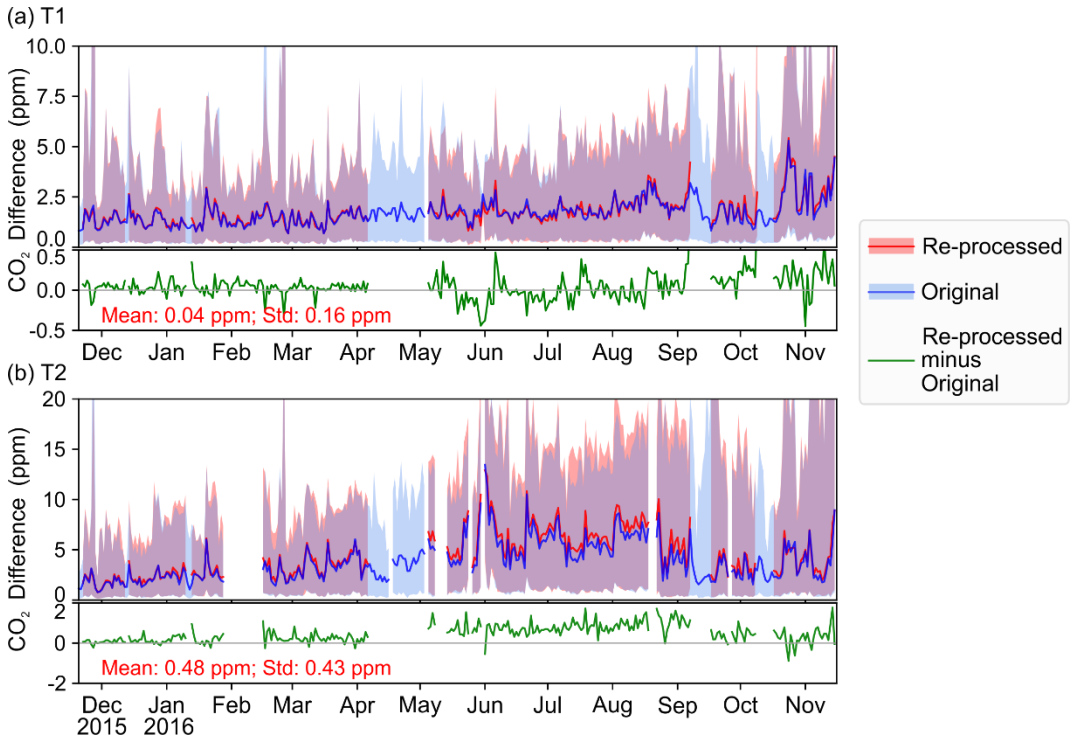

Figure S2: Distribution of the original and re-processed GreenLITE™ absolute CO_2 concentration differences between all pairs of chords for (a) T1 and (b) T2 from December 2015 to November 2016. The solid lines in top panels of (a) and (b) indicate the 0.5 quantile, and the shaded areas represent the 0.1 and 0.9 quantile intervals for original data in blue and re-processed data in red. The green line in bottom panels of (a) and (b) indicates the differences between the median values of the re-processed and original inter-chord range.

Page 4, Line 27: Why is a threshold of standard deviation < 10 ppm CO_2 applied to the hourly data? Is this because CO_2 is not expected to change by more than 10 ppm over the course of one hour? Is this justifiable?

The outlier detection for the 4-minute GreenLITE™ data is mainly based on the 3-sigma rule, which is used to remove the data outside three standard deviations from a mean in the positive direction. We have added a sentence in the main body of the revised manuscript and Figure S3 in the supplement to answer the reviewer's comment:

"The 10 ppm threshold was selected to be rough 3 times the typical standard deviation of the 4-minute measurements for any given chord within a one-hour period (Figure S3)."

For better clarity, we have also added the following statement in the supplement together with Figure S3.

"The outlier detection for the 4-minute GreenLITE™ data is mainly based on the 3-sigma rule, which is used to remove the data outside three standard deviations from a mean in the positive direction. Figure S3 (a) shows the frequency distribution of the standard deviations of the 4-minute CO_2 concentrations measured within one hour for one given chord (e.g. T2R08). Figure S3 (b) shows the three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord. In general, the threshold varies between 6.5 ppm and 11.9 ppm from chord to chord. We therefore choose to use a uniform threshold value of 10 ppm to remove the outliers for all chords."

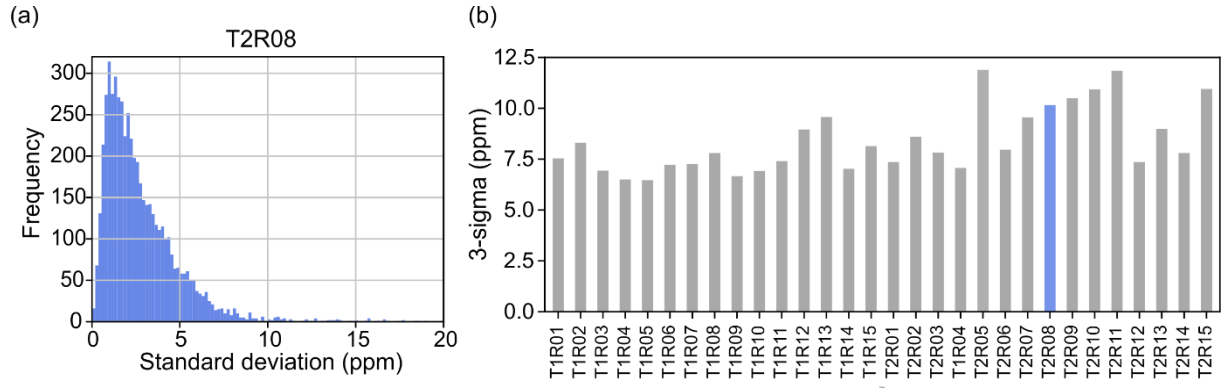

Figure S3: (a) Frequency distribution of the standard deviations of the 4-minute CO_2 concentrations measured within one hour for one chord (e.g. T2R08); (b) Three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord.

Page 5, Line 19: Can you quantify "much larger differences"?

For better clarity, we have added the following sentence in the revised manuscript:

"In order to select an adequate model physical configuration for Paris, we carried out some preliminary sensitivity experiments to test the impact of different physical schemes on the simulated CO_2 concentrations. These tests use up to five different PBL schemes and two urban canopy schemes. The simulations were carried out for two months, including one winter month (January 2016) and one summer month (July 2016). These preliminary sensitivity results indicate that different PBL schemes in the WRF-Chem model lead to monthly average differences of 2-3 ppm on the simulated CO_2 concentrations over Paris, whereas the two different urban canopy schemes lead to much larger differences of 8-10 ppm. Thus in this study, we carried out the 1-year simulation with two different urban canopy schemes as they are sufficient to address the paper main question regarding the ability of a configuration of the WRF-Chem model to simulate the CO_2 atmospheric transport in an urban environment, but also to provide an estimate of the modeling uncertainty. All of the other physics options remained the same for the two experiments (Table 2)."

Page 5, Line 29: Typo - "details" should be "detail".

Correction made.

Page 6, Line 5: Consider "accounting for" rather than "taking up"?

Text changed as suggested.

Page 11, Line 9: Rephrase "city surrounding" to "areas surrounding the city", or similar.

Text changed as suggested. The modified text is as follows:

"On the other hand, both models show similar performances in the areas surrounding the city."

Table 3: What are the colour scales showing? Better or worse values? This needs to be made more clear particularly because high correlation coefficient (red) is good but high RMSE (also red) is bad?

We agree with the reviewer that the color scales in Table 3 can be misleading. The color only represents the values from minimum (blue) to maximum (red) in the cells instead of indicating the goodness of fit between model and observation. We have added the following text in the caption of Table 3 in order to clarify this issue:

"The color highlights the value in the cell with the minimum in blue, the median in white and the maximum in red. All other cells are colored proportionally."

Figure 1: Some text is very small – a possible solution would be to refer the reader to the panel in Fig 2 in the caption and remove the chord labels. Also the caption refers to Figure S1 but this doesn't appear relevant to the text – the authors might mean Figure S2?

For better clarity, we have added a second panel in Figure 1 and noted the previous Figure 1 as Figure 1a. Now, Figure 1a shows the distribution of in-situ CO_2 stations and the GreenLITE™ laser system without the chord labels. Figure 1b is a high-resolution zoom of the inner Paris area and shows the GreenLITE™ laser system layout in detail.

Corrected, thanks. It should refer to Figure S5 in the revised manuscript.

Figure 2: Does the caption need to state that these emissions are taken from an emissions inventory i.e. not measured or modelled.

Yes. We have added the following sentence into the caption to stress this point:

"Figure 2: Total CO_2 emissions, according to the AirParif inventory (within IdF) and the IER inventory (outside IdF), for a weekday in March 2016."

Figure 3: See Figure 2.

We have modified the caption:

"Figure 3: Averaged anthropogenic CO_2 fluxes along each GreenLITE™ chord according to the AirParif inventory."

Figure 4b: Is there some way of better highlighting that this is not a continuous time series of data? Perhaps either thicker/bolder lines or a gap between each monthly diurnal cycle.

As suggested we have used thicker lines between each monthly diurnal cycle in the revised Figure 4b.

Figure 5: The blue (observation) line is quite difficult to see on these plots.

We have changed the line colors to a sharper contrast between the observation and the model results. (PS: following the recommendation from Referee #3, we have moved this figure to the supplement as Figure S7 in the revised manuscript)

Figure 7: y-axis titles should probably read "CO_2 difference (ppm)" as in Figure 6.

Changed as suggested. This figure is now Figure 6 in the revised manuscript.

We would like to thank Referee #2 for his/her thoughtful comments and detailed suggestions to our manuscript. In the following, we answer to the reviewer's comments and indicate the changes in the manuscript that were implemented as a consequence of the recommendations. The comments are in black and italic. Our answers are in blue and plain text.

Anonymous Referee #2

Received and published: 6 August 2019

This paper describes urban measurements of CO_2 by in-situ and by the novel open-path laser system "GreenLITE" with multiple reflectors and transceivers deployed in the Paris area. Observations are compared to high-resolution WRF-Chem simulations with a representation of CO_2 fluxes from anthropogenic emissions and biosphere atmosphere exchange. The paper is well written, and I recommend publishing after the following minor comments are addressed.

We thank the reviewer for his/her work and suggestions.

General comments:

For the WRF-Chem modeling CO_2 emissions at annual and national scale for scaling the high spatial resolution emissions to the year of interest have been taken from the Global Carbon Atlas (GCA), however it is unclear what these data are based on (e.g. UNFCCC reporting, BP statistical reports, or other sources). The Global Carbon Atlas has some missing links in the "Data contributors" section making traceability of the emissions impossible. This needs to be clarified.

We recognize that the sentence was insufficiently detailed. We have now modified the following statements with a supplement table (Table S1) to address the data sources and the corresponding references.

"This is accomplished by rescaling the maps with the ratio of the annual budgets of national CO_2 emissions for the countries within the domain between the base year 2005 for IER and 2010 for AirParif and the year of simulation (2015/2016), taken from Le Quéré et al. (2018) (https://www.icos-cp.eu/GCP/2018. See also Table S1 in the supplement for details about original data sources)."

Table S1. National CO_2 emissions from fossil-fuel combustion and cement production for the countries within the WRF-Chem domain used in this study (unit: $MtCO_2/yr$). The data in the following table are taken from Le Quéré et al. (2018), available at https://www.icos-cp.eu/GCP/2018, last access: August 2019. (The use of data is conditional on citing the original data sources: data in black are from the CDIAC inventory (Boden et al., 2017), data in red are from the UNFCCC national inventory reports (UNFCCC, 2018), data in purple are from the BP Statistical Review of World Energy (BP, 2018). Cement emissions are updated from Andrews (2018))

	Austria	Belgium	France (including Monaco)	Germany	Italy (including San Marino)	Luxembourg	Netherlands	Spain	Switzerland	United Kingdom
2005	79.37	125.64	432.64	867.22	495.23	12.05	177.53	368.96	45.78	570.00
2010	72.38	113.58	397.90	833.68	424.87	11.15	182.18	283.88	45.05	512.21
2015	66.70	100.23	348.16	797.08	355.48	9.26	165.03	271.73	38.74	422.66
2016	67.40	100.24	350.10	801.75	350.32	9.00	165.52	260.99	39.20	398.55

References:

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth System Science Data, 10, 2141-2194, 2018.

Andrew, R. M.: Global CO_2 emissions from cement production, Earth System Science Data, 10, 195-217, https://doi.org/10.5194/essd-10-195-2018, 2018.

Boden, T. A., Marland, G., and Andres, R. J.: Global, Regional, and National Fossil-Fuel CO_2 Emissions, available at: http://cdiac.ornl.gov/trends/emis/overview_2014.html (last access: July 2017), Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA, 2017.

BP: BP Statistical Review of World Energy June 2018, available at: https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf, last access: June 2018.

UNFCCC, 2018. National Inventory Submissions 2018. United Nations Framework Convention on Climate Change. Available at: http://unfccc.int/process/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/national-inventory-submissions-2018; accessed June 2018.

It is somewhat unclear how the statistics shown in Table 3 and Fig. S4 have been calculated for the GreenLITE vs. WRF-Chem measurements in section 4.1. Have the data from all chords related to e.g. T1 been combined and then the statistics is derived, or has each chord been treated independent and the resulting statistics shown in Table 3 and Fig. S4 reflect the average across all chords?

We agree with the reviewer that this point needs clarification. (PS: following the recommendation from Referee #3, we have split the previous Table 3 into Table 3 and Table S2. The previous Figure S4 is now Figure S6 in the revised manuscript)

We have added the following sentence to make it clearer:

"The statistics shown in Table 3, Table S2 and Figure S6 also indicate the ability of the models to reproduce the CO_2 at two urban in-situ stations (JUS & CDS) and the GreenLITE™ measurements. As for the GreenLITE™ data, we first compute the hourly averages of the observed and modeled CO_2 concentrations over all 15 chords for each transceiver (T1 and T2), and then calculate the respective statistics."

The discussion of the results in section 4.2.2 regarding the spatial gradients between different chords of the GreenLITE observations and the simulated counterparts, as well as the corresponding discrepancy between observations and model results should at least mention the potential impact of turbulent eddies and thermals. Those are likely to form in a convectively unstable atmosphere, i.e. during summer, and are unlikely to be represented properly in the MYJ PBL scheme (a local closure model) deployed in the WRF-

Chem simulations (c.f. Xiao-Ming et al., 2010). Ref.: Hu, Xiao-Ming, John W Nielsen-Gammon, and Fuqing Zhang. 2010. "Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model." Journal of Applied Meteorology and Climatology 49 (9): 1831–44. doi:10.1175/2010JAMC2432.1

We greatly appreciate the reviewer's suggestion and fully agree that the vertical mixing associated with turbulent eddies and thermals plays an important role in the CO_2 transport and dispersion. The impact of insufficient vertical mixing, local eddy diffusion and entrainment flux under convective conditions reproduced by the local closure MYJ PBL scheme is a plausible explanation for the model-observation misfits. The revised manuscript has included the following discussion as suggested by the reviewer:

"Another potential source of measurement-model discrepancy is the atmospheric transport modeling as proposed in H2. According to previous studies (e.g. Hu et al., 2010), the turbulent eddies and thermals are unlikely to be reproduced properly by the local closure MYJ PBL scheme, which results in insufficient vertical mixing under convective (unstable) conditions, e.g. during summer. It may also indicate that the WRF-Chem model at a 1-km horizontal resolution cannot reproduce the fine-scale (sub-kilometer) CO_2 concentration features over a complex urban environment in Paris, as the analysis of JUS and CDS in-situ measurements has shown in Section 4.2.1."

Reference:

Hu, X. M., Nielsen-Gammon, J. W., and Zhang, F.: Evaluation of three planetary boundary layer schemes in the WRF model. Journal of Applied Meteorology and Climatology, 49(9), 1831-1844, 2010.

Specific comments:

P1 L36: I suggest replacing "have been used" with e.g. "have been or will be used" as you are refering also to future satellites.

Text changed as suggested.

P7 L21: please rephrase "low atmosphere", e.g. "lower part of the atmosphere"

Text changed as suggested. The modified text is as follows:

"It is well known that the lower part of the atmosphere is, on average, more stable in winter than in summer."

P7 L38: "in some respect superior" this should be formulated clearer. What I see from Table 3 is that RMSE with the BEP model is always better for T1 than for JUS, and better for T1 than for CDS with one exception.

To further clarify this point, the sentence has been refined as follows:

"The RMSE with the BEP scheme is within the range of 4.5 to 9.6 ppm for T1 which is substantially superior to those of JUS and CDS, with only one exception at CDS during summer when the value is slightly better for CDS than for T1."

P8 L21, P8 L37, and P10 L19: please rephrase "The std values", e.g. "Standard deviations"

Text changed as suggested.

P9 L38: What is the difference between the first two of the three hypotheses? Is H2 meant to refer to only transport model deficiencies, excluding inaccuracies in emissions? This should be made clearer. Also it should be made clear at the end of section 4.2.2 which hypothesis remains the most probable one.

Yes, Hypothesis 1 is about potential inaccuracies or uncertainties of the emission inventory for the Paris urban area, whereas Hypothesis 2 refers to the imperfect modeling of the atmospheric transport and dispersion of CO_2 over the complex urban area. We have modified the statement to make it clearer:

"• H2 The models fail in the description of CO_2 concentrations within the Paris city because of imperfect representations of atmospheric transport processes, excluding inaccuracies in emissions;"

Our analyses indicate that the model-GreenLITE™ discrepancy during the summer is more likely the consequence of the measurement noise and bias in some of the chords, whereas it is hard to fully rule out the possibility of impacts of the emission spatial structure and the atmospheric transport that have been discussed in section 4.2.2. Therefore, we tend to be more cautious to make such an assessment based on our current knowledge. We have added the following sentence in the conclusion and discussion section to address this point:

"Although it is not yet fully understood, several evidences suggest an increase of measurement noise and bias in some of the GreenLITE™ chords during the summer season, that must be resolved or reduced before assimilating the whole dataset into the CO_2 atmospheric inversion system that aims at retrieving urban fluxes."

We would like to thank Referee #3 for his/her thoughtful comments and detailed suggestions to our manuscript. In the following, we answer to the reviewer's comments and indicate the changes in the manuscript that were implemented as a consequence of the recommendations. The comments are in black and italic. Our answers are in blue and plain text.

Anonymous Referee #3

Lian et al., present long open-path spectroscopy measurements for the City of Paris from December 2015 to November 2016 in conjunction with in-situ observations from towers in and around Paris as well as WRF-Chem simulated observations from two different urban canopy schemes. It is assumed that the authors are using the GreenLITE measurements along with the in-situ tower observations to discern which WRF-Chem urban canopy scheme can best represent vertical mixing and transport in urban areas.

My main concern with this article is that the specific objective/conclusions of the paper are unclear. The authors have conducted a lot of work analyzing data from many components, but it is uncertain as to whether they have drawn any solid conclusions.

The objective, which I assume is using GreenLITE and the in-situ observations to evaluate WRF-Chem urban canopy configurations, should be more clearly stated in the Introduction and the title of the paper should be changed.

Without a clear narrative, the paper mainly comes across as a presentation of data which is difficult to evaluate as a reader. I do not recommend that this work be accepted for publication in ACP without substantial revisions to clarify scientific objectives/conclusions.

We thank the reviewer for his/her work and comments on our paper. To answer this criticism, we have added a few sentences in the introduction and conclusion sections that make clearer the objectives of the paper and the conclusions that were derived.

Urban areas are significant sources of fossil fuel CO_2 emissions. CO_2 measurements in urban areas are used in conjunction with atmospheric transport models and statistical inversion techniques to estimate city CO_2 emissions. The novel GreenLITE™ laser imaging system deployed in Paris provides a much wider spatial coverage of atmospheric CO_2 concentrations over the complex urban environment, which makes it possible to provide new insights into the CO_2 characteristics compared to the highly accurate in-situ measurements that can only be made at point locations, and can be influenced by local sources to a poorly known extent. In this paper, we analyze the measurements provided by this novel system, together with the more classical in-situ sampling and high-resolution modeling and we focus on the temporal and spatial variability of atmospheric CO_2 concentrations. The main purpose of the paper is therefore an evaluation of the new system capabilities to provide information on both the emission and the atmospheric transport, typically whether the new system can falsify the emission inventory used or point out to transport modeling deficiencies (the two hypotheses formulated in the paper).

We have added the following paragraph in the introduction to better clarify the objectives of this work:

"The detailed objectives of this paper are:

- To analyze in detail the information content of the GreenLITE™ data in addition to conventional in-situ CO_2 measurements in order to better understand the temporal and spatial variations of near-surface CO_2 concentrations over Paris and its vicinity.

- To evaluate the performance of the high-resolution WRF-Chem model coupled with two urban canopy schemes (UCM, BEP) for the transport of CO_2 over the Paris megacity area based on the two types of CO_2 measurements.
- To discuss the potential implications of assimilating the GreenLITE™ data into the CO_2 atmospheric inversion systems with the ultimate goal of increasing the robustness of the quantification of city emissions and constraining the spatial distribution of the emissions within the urban area."

The main conclusions of this study are:

- Two urban canopy schemes (UCM, BEP) as part of the WRF-Chem model are capable of reproducing the seasonal cycle and most of the synoptic variations in the atmospheric CO_2 in-situ measurements over the suburban areas, as well as the general corresponding spatial differences in CO_2 concentration between pairs of in-situ stations that span the urban area.
- The GreenLITE™ measurements are less sensitive to local unresolved sources than the in-situ point measurements, and are then better suited for the comparison to km-scale modeling. In our analysis, the GreenLITE™ data have been used to show a deficiency of the UCM scheme during the winter, linked to underestimated vertical mixing. Conversely, the model-GreenLITE™ discrepancy that is observed during the summer is not yet fully understood. Several evidences suggest an increase of measurement noise and bias in some of the GreenLITE™ chords during the summer season, that must be resolved or reduced before assimilating the whole dataset into the CO_2 atmospheric inversion system that aims at retrieving urban fluxes.
- Within the city, the misfit between the observed and simulated CO_2 concentrations is found to be highly sensitive to the WRF-Chem configuration for the urban canopy scheme, which affects the atmospheric vertical mixing. We also show that the CO_2 concentrations are impacted by the spatial distribution of the emission and the presence of local sources that are poorly resolved in the inventory. This study stresses the difficulty in reproducing precisely the atmospheric CO_2 concentration within the city because of our inability to represent the detailed spatial structure of the emission and because of the sensitivity of the concentration to the strength of vertical mixing. From the model results analysis, we infer that the uncertainty on the vertical mixing is much larger than the uncertainty on the emissions so that atmospheric concentration measurements within the city can hardly be used to constrain the emission inventories.

We have modified the conclusion and discussion section based on the main conclusions listed above.

I have the following suggestions regarding the technical components of the analysis.

Main Comments:

(1) The observational network: can you provide some indication as to areas in which the observations are most sensitive especially CDS, JUS, and any others that are situated in or adjacent to major sources/sinks? I understand that the authors use WRFCHEM and not a Lagrangian approach so that footprints cannot be generated but having some understanding would help the comparison of GreenLITE and in-situ observations presented later in the paper since these two different types of measurements represent different spatial extents.

The footprint of the measurement station very much depends on the wind speed and direction, as well as the atmospheric stability. It is then difficult to interpret a mean footprint that aggregates a wide range of atmospheric situations. In response to the reviewer's concern about the major sources of emissions to the station, we carried out some sensitivity experiments for the one-month period of March 2016 with

anthropogenic and biogenic emissions limited to a given area within the simulation domain in order to quantify their respective contributions to the simulated CO_2 concentrations at a certain measurement site. This set of experiments includes the assignment of emissions to: 1) ONE: one grid cell that contains the in-situ station, 2) GRP: all grid cells within the GReater Paris except the one where the station is located, 3) IDF: all grid cells within the IdF region except those of the Greater Paris, 4) OUT: all grid cells outside the IdF region, as shown in Figure S10 (a) (b) (c) (d) respectively. The contribution from sources outside the model domain is small enough so that its influence is negligible. Figure S11 shows the relative contributions (in percentages) of each component to the modeled total anthropogenic and biogenic CO_2 concentrations for one urban site JUS and one suburban site COU respectively. The simulated monthly mean concentrations of anthropogenic CO_2 are 11.0 ppm at JUS and 5.4 ppm at COU, which are much larger than those of biogenic CO_2 (0.6 ppm at JUS and 0.7 ppm at COU). In general, an urban station like JUS is under a strong influence of the anthropogenic emissions within the IdF region. The contributions of anthropogenic emissions in the vicinity of the station (ONE) and from the Greater Paris (GRP) areas to the simulated anthropogenic CO_2 concentrations are around 16% and 60% respectively, whereas the remote anthropogenic emissions account for less than 20%. For a suburban station like COU, the Parisian emissions (GRP) and the remote ones (OUT) have a comparable influence (~40%) on the simulated anthropogenic concentrations, with very large variations depending on the wind direction (downwind or upwind of the city). The biogenic CO_2 concentrations mainly come from outside of the IdF region (~86%).

We have added the following statement in the main body of the revised manuscript and two figures (Figure S10 and S11) in the supplement based on the discussion above:

"Atmospheric transport simulations make it possible to assess the respective contributions of various areas/sectors to the measurements. Our preliminary sensitivity experiments (see Figure S10 and S11 for details) have shown that the anthropogenic emission from the Greater Paris area is the dominant contribution (~80%) to the anthropogenic CO_2 signal at the urban measurement stations. In order to get further insights into the characteristics of CO_2 spatial variations within the Paris city, it is therefore necessary to analyze the CO_2 differences with the consideration of the anthropogenic CO_2 emissions shown in Figure 2 and Figure 3."

For better clarity, we have also added the following analyses in the supplement together with Figure S10 and S11.

"In order to determine respective contributions of various areas/sectors to the simulated CO_2 concentrations at a certain measurement site, we carried out a set of sensitivity experiments for the one-month period of March 2016 with anthropogenic and biogenic emissions limited to a given area. This set of experiments includes the assignment of emissions to: 1) ONE: one grid cell that contains an in-situ station, 2) GRP: all grid cells within the GReater Paris except the one where the station is located, 3) IDF: all grid cells within the IdF region except those of the Greater Paris, 4) OUT: all grid cells outside the IdF region, as shown in Figure S10 (a) (b) (c) (d) respectively.

Figure S11 shows the relative contributions (in percentages) of each component to the modeled total anthropogenic and biogenic CO_2 concentrations for one urban site JUS and one suburban site COU respectively. The simulated monthly mean concentrations of anthropogenic CO_2 are 11.0 ppm at JUS and 5.4 ppm at COU, which are much larger than those of biogenic CO_2 (0.6 ppm at JUS and 0.7 ppm at COU). In general, an urban station like JUS is under a strong influence of the anthropogenic emissions within the IdF region. The contributions of anthropogenic emissions in the vicinity of the station (ONE) and from the Greater Paris (GRP) areas to the simulated anthropogenic CO_2 concentrations are around 16% and 60% respectively, whereas the remote anthropogenic emissions account for less than 20%. For a suburban station

like COU, the Parisian emissions (GRP) and the remote ones (OUT) have a comparable influence (~40%) on the simulated anthropogenic concentrations, with very large variations depending on the wind direction (downwind or upwind of the city). Note that in these experiments, the emission inventory and the WRF-Chem modeling cannot describe the CO_2 patterns (both emission and concentration) at a scale finer than 1 km, and the simulation shows that the "local" contribution is significant. The unresolved spatial distribution of the emission can therefore be a significant contribution to the uncertainty. The biogenic CO_2 concentrations mainly come from outside of the IdF region (~86%)."

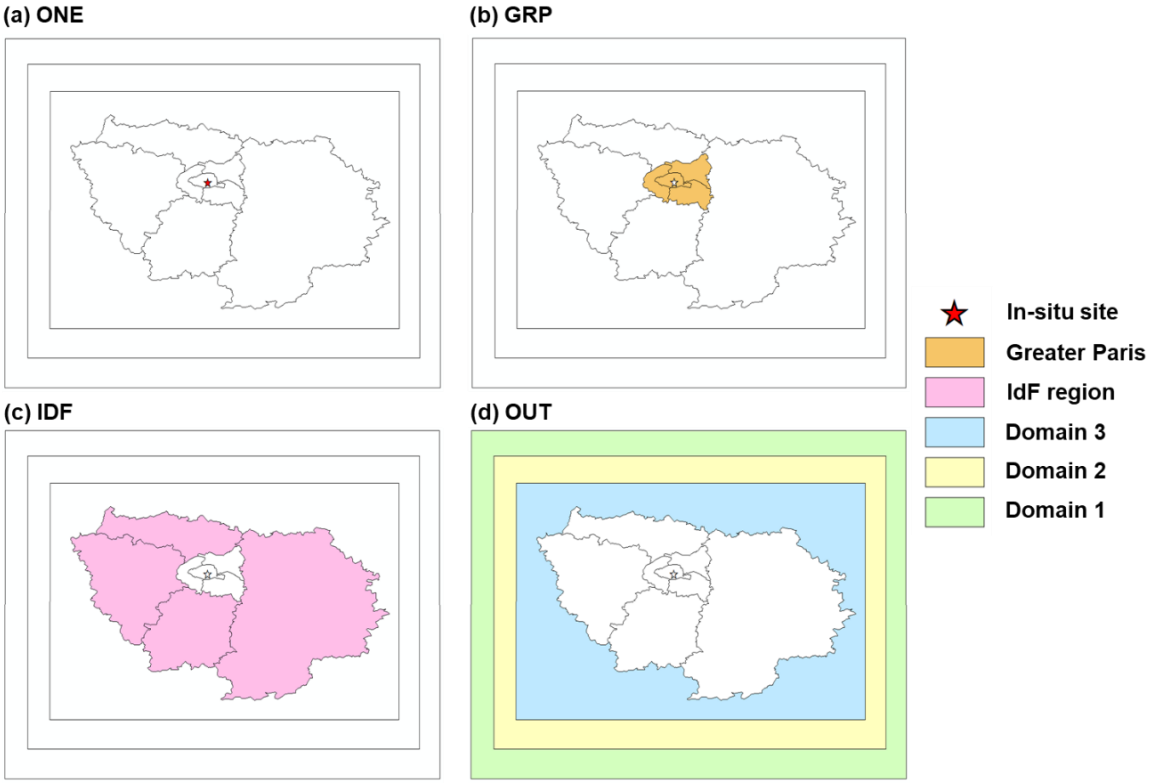

Figure S10. Four experiments are carried out for the JUS station with the assignment of emissions to: (a) ONE: one grid cell that contains an in-situ station; (b) GRP: all grid cells within the Greater Paris except the one where the station is located; (c) IDF: all grid cells within the IdF region except those of the Greater Paris; (d) OUT: all grid cells outside the IdF region. Another four experiments are carried out for the COU station.

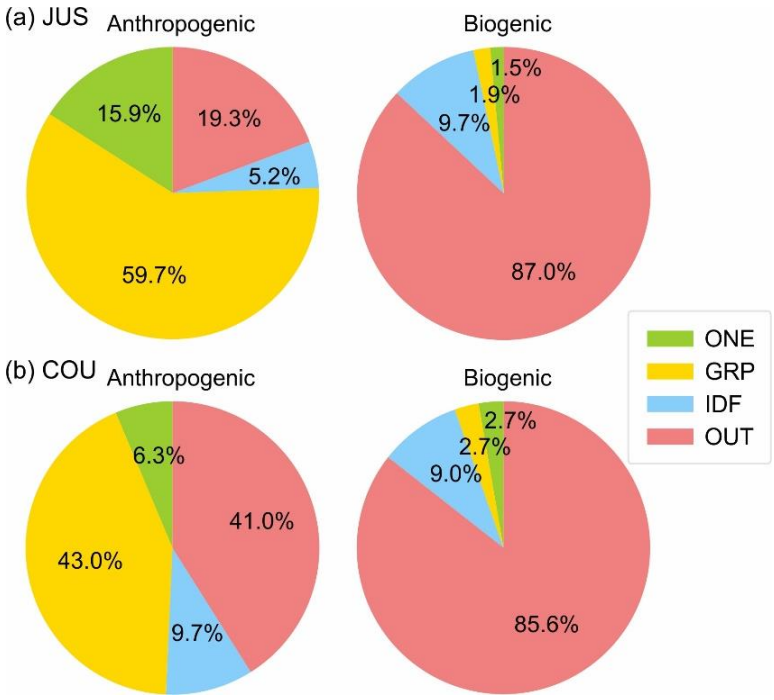

Figure S11. Relative contributions (in percentages) of each component flux to the modeled total anthropogenic and biogenic CO_2 concentrations for (a) urban site JUS and (b) suburban site COU. Note that only the afternoon data (11-16 UTC) are used in the analysis.

(2) GreenLITE campaign – can you provide more description than citing the Zaccheo paper as to how the GreenLITE observations are calibrated?

As mentioned by the reviewer, the calibration method is extensively described in Zaccheo et al. (2019). Nevertheless, we have followed the suggestion and now provide a better description (although very much summarized) of the procedure:

"These slowly time-varying differences were most likely due to a slight systematic long-term drift in both the on- and off-line wavelengths as a function of continuous operations. Such drift may induce some non-linear impacts on the measured concentrations. It is therefore more appropriate to adjust the wavelengths rather than to apply a linear calibration to the retrieved concentrations. Unlike in-situ point measurement systems, there is no established method for calibration of long open-path systems to the WMO mole fraction scale used as an international standard for atmospheric CO_2 monitoring (Tans et al., 2011). Therefore, a bias correction method was developed by AER (Zaccheo et al., 2019) for addressing observed slowly drifting biases between the GreenLITE™ prototype system and the two in-situ sensors (CDS and JUS) that are near the GreenLITE™ chords. This method computed a time-varying adjustment to the offline wavelength based on a non-linear optimization mechanism. This non-linear approach adjusts the GreenLITE™ offline wavelength considering not only the average values of hourly CO_2 concentrations at two in-situ stations, but also the corresponding average temperature, relative humidity, atmospheric pressure along the chord and an optimized online wavelength value during the measurement period. Finally, the median on- and off-line values over a 4-day window was used to recompute the GreenLITE™ data from all chords using a radiative transfer based iterative retrieval scheme based on the LBLRTM model (Clough et al., 2005). Even though this approach is not ideal as the two in-situ stations and the GreenLITE™ system do not sample the exact same area, it does provide a well-defined mechanism that reduces the systematic long-term biases with no significant impact on the chord-to-chord variations."

A general response to comments (3) (4) and (5):

We fully agree with the reviewer that the uncertainties associated with atmospheric transport, anthropogenic emissions, biogenic fluxes and background conditions could all have a more or less impact on the model performance. Over the years, many studies have been carried out at different scales and regions on the analysis, derivation and quantification of critical sources of uncertainties that lead to the model-observation misfits. Nevertheless, even with a state-of-art atmospheric transport model and an inventory at a high spatio-temporal resolution, the uncertainties associated with the modeling are inevitable and cannot be completely eliminated. The ensemble-based sensitivity study and a full analysis on the uncertainties derived from different anthropogenic inventories, biogenic fluxes, atmospheric transports or background conditions are out of the scope of this study that focuses on the potential contribution of GreenLITE™ observing system, but will be specifically addressed within another dedicated study.

(3) WRF-Chem – Is this paper an analysis of WRF-Chem urban canopy models for cities like Paris (evaluated using GreenLITE and in-situ observations)? If so, please substantiate/provide reference for the claim that "previous sensitivity tests indicate that different physical schemes in the WRF-Chem model lead to mean differences of 2-3ppm on the simulated CO_2 concentrations over Paris, whereas the various urban canopy schemes lead to much larger differences." This seems like the motivation for much of the work presented within the paper, but I am not sure that this claim, if substantiated, holds true across most urban areas.

As our answers to the general comment, our main objective is not fully to test the modeled CO_2 sensitivity to the use of different physical schemes and to discern which urban canopy scheme could reach best results when comparing the model to observations. Several studies have demonstrated that the city-scale physical and dynamical processes in the atmospheric modeling system remain a challenge. In order to select an adequate WRF-Chem model configuration for Paris, we did perform some preliminary sensitivity experiments to test the impact of different physical schemes (5 PBL schemes + UCM, 2 PBL schemes + BEP) on the simulated CO_2 concentrations. The simulations were carried out for two months, including one winter month (January 2016) and one summer month (July 2016). These preliminary sensitivity results indicate that different PBL schemes in the WRF-Chem model lead to monthly average differences of 2-3ppm on the simulated CO_2 concentrations over Paris, whereas the urban canopy schemes lead to much larger differences of 8-10 ppm. We thus carried out the 1-year simulation with two different urban canopy schemes as they are sufficient to address the paper main question regarding the ability of a configuration of the WRF-Chem model to simulate the CO_2 atmospheric transport in an urban environment, but also to provide an estimate of the modeling uncertainty.

We have added the following sentence in the manuscript to account for the reviewer's comment:

"In order to select an adequate model physical configuration for Paris, we carried out some preliminary sensitivity experiments to test the impact of different physical schemes on the simulated CO_2 concentrations. These tests use up to five different PBL schemes and two urban canopy schemes. The simulations were carried out for two months, including one winter month (January 2016) and one summer month (July 2016). These preliminary sensitivity results indicate that different PBL schemes in the WRF-Chem model lead to monthly average differences of 2-3 ppm on the simulated CO_2 concentrations over Paris, whereas the two different urban canopy schemes lead to much larger differences of 8-10 ppm. Thus in this study, we carried out the 1-year simulation with two different urban canopy schemes as they are sufficient to address the paper main question regarding the ability of a configuration of the WRF-Chem model to simulate the CO_2 atmospheric transport in an urban environment, but also to provide an estimate of the modeling uncertainty. All of the other physics options remained the same for the two experiments (Table 2)."

a. If this is not the focus, and the purpose is to use the meteorology to understand the variability of the measurements, then I believe the authors should pick a model and use it throughout the rest of the analysis. It seems (from Figure 5) that the BEP model is largely better. The rest of the analysis using UCM could be put in the supplementary information. As an aside, I do think that the authors could use the ensembles in a way that would help them draw some robust conclusions. Their ensembles provide some measure of the atmospheric transport and dispersion uncertainty which can be used to contextualize their comparison between GreenLITE and the in-situ observations (S1 and S4).

The main purpose of the study is to assess the potential contribution of the GreenLITE™ system, in addition to the more classical in-situ sampling, for a better understanding of the temporal and spatial variations of near-surface CO_2 concentrations over Paris and its vicinity (due to emissions / atmospheric transport). Even though the two urban canopy schemes do not represent the full range of uncertainty in the atmospheric transport, in some extent they do provide an insight into the critical impact of the atmospheric transport on simulated atmospheric CO_2 concentrations. It thus appears to us that it is necessary to keep the UCM analysis within the main body of the paper. We also tested different PBL schemes for the impact of the atmospheric transport, but decided to show results for the two urban canopy schemes because they caused the largest differences in simulated CO_2 concentrations.

(4) Anthropogenic Fluxes – The use of the anthropogenic fluxes in the analysis should be reconsidered or better explained. For example, does IER have any temporal variability? If so, please explain. If not, the authors could consider scaling using published methods. The authors could use other emission products that have temporal variability if needed. The loss of spatial scale (e.g. going from 5-10km) seems less important than preserving some temporal structure in emissions. Other products are also more recent and thus more represented of ex-urban fluxes which constitute a large portion of CO_2 inflow to Paris.

Yes, the IER inventory used in this study has a detailed country-specific temporal profiles (monthly, daily and hourly) at spatiotemporal resolutions of 5 km and 1 h. Given the fact that it has a higher spatial resolution than some other emission products and it has been rescaled to account for annual changes in emission between the base year and simulation timeframe, this inventory is sufficient to be used in this study.

We have added the information about the IER temporal variability in the revised manuscript:

"CO_2 emissions from fossil fuel CO_2 sources outside the IdF region are taken from the inventory of the European greenhouse gas emissions, together with country-specific temporal profiles (monthly, daily and hourly) at a spatial resolution of 5 km (updated in October 2005). This inventory was developed by the Institute of Economics and the Rational Use of Energy (IER), University of Stuttgart, under the CarboEurope-IP project (http://www.carboeurope.org/)."

Could the authors also further explain "we interpolate the emissions to the WRF-Chem grids following the principle of mass conservation?" This is unclear in both its meaning and why it is important.

The total magnitude of anthropogenic emissions should be consistent before and after the interpolation to model grid cells. We have modified the statement to make it clearer:

"Finally, we interpolate the emissions onto the WRF-Chem grids, making sure to conserve the total budget of emission in the process, as done in previous studies (e.g. Ahmadov et al., 2007)."

As with the WRF-Chem comments, the authors could use an ensemble of anthropogenic emission products (those outside of Paris) to help contextualize the GreenLITE and in-situ observations in terms of emission uncertainty (refer to Martin et al., 2018).

Please see the analysis above. The contribution of anthropogenic emissions outside the IdF region to the simulated anthropogenic CO_2 concentrations over urban areas is relatively small (~20%), and our present simulations are capable of reproducing the seasonal cycle and most of the synoptic variations in the atmospheric CO_2 point measurements over the suburban areas. In addition, this distant contribution is much smoother, both in temporal and spatial scales, than the impact of more local emissions. We then do not expect a significant impact of these distant emissions on the CO_2 signatures that are analyzed in the paper. We then do not feel that an ensemble of anthropogenic emission products outside the IdF region will bring critical insights on the main conclusions at two urban in-situ stations (JUS & CDS) and the GreenLITE™ measurements. A deeper analysis of the impact of uncertainties in the anthropogenic emissions outside the IdF region is out of the scope of the paper.

(5) Biogenic Fluxes – The use of VPRM to represent the urban biosphere is an active area of research and there are lots of questions as to how well a biospheric model captures the urban biogenic emissions. When VPRM was optimized using flux data, were urban towers used to help parameterize the "urban" areas of Paris? The paper mentions that the western portion of Paris has much green space and thus biogenic sources might be important in this area of the city and impact the analysis. How was Paris-VPRM (or VPRM) validated, e.g. comparison to in-situ data from towers outside of the city that are surrounded by vegetation (maybe OVS)? Has it been used in other studies? How does it vary as a function of time in comparison to the anthropogenic fluxes like what is shown in Figure 3?

As for the urban biogenic emissions mentioned by the reviewer, we certainly agree that it might be important for the simulated CO_2 concentrations and it is an active area of research. To which extent the biogenic fluxes affect the simulated CO_2 concentrations in the Paris urban areas remains an open question. Whereas there is no eddy covariance measurement in the Paris urban area that is available for the biospheric flux optimization and we are not able to make an evaluation of the Paris-VPRM model in this study. Nevertheless, we have performed some further analyses and validations of the VPRM model at a suburban station at SAC in a dedicated study mentioned above. Since these analyses at SAC do not reflect the model performance of the biosphere mode in the urban area, it is out of the scope of this study. Mean diurnal cycles of CO_2 biogenic flux (NEE) for 12 calendar months and for 8 vegetation classes used in VPRM over Domain 03 are shown in Figure 4 and the related texts are in Section 3.2.2.

(6) Results – (4.1) There are a lot of moving pieces in this analysis and it is hard to ascertain the main conclusions from the statistical analysis. Do you think that the uncertainties associated with the other components (e.g. anthropogenic emissions and vprm sources and sinks) would have changed some of these results especially during the growing seasons or per your analysis of the seasonality of the sectors? From the Table, it is unclear that BEP outperforms UCM for much of the year. As with 4.1, I am not sure what the main takeaway is from this analysis.

In fact, each paragraph in Section 4.1 relates to a certain aspect regarding the statistics for observed and modeled CO_2 concentrations for periods of the day (all hourly data, hourly afternoon data) and two urban canopy schemes (UCM, BEP). In general, the model performance is better during the afternoon than it is for the full day. UCM and BEP have different performances for four seasons and for urban/suburban areas (see answers below, this is also the main takeaway). The statistics further confirm the fact that the GreenLITE™ measurements represent an average over a wide area, and are then less sensitive to local unresolved sources than the in-situ measurements.

We do not know to what extent the uncertainties associated with the other components (e.g. anthropogenic emissions and VPRM sources and sinks) would have changed some of these results since we have not made the relevant sensitivity experiments. A full analysis of these uncertainties would be a paper by itself.

In the third paragraph of Section 4.1, we have already discussed the different performances of UCM and BEP for four seasons and for urban/suburban areas with the following statements: 1) The statistics for BEP compared to the observations within the urban areas are significantly better than UCM during autumn and winter; 2) CO_2 concentrations are better reproduced by both UCM and BEP in the spring; 3) Both models show lower correlations during summer; 4) the UCM and BEP also have comparable performances at peri-urban areas while the BEP is slightly better at some suburban sites as shown by the statistics.

(4.2.1) Why did you use the wind per ECMWF versus wind measurements at the upwind tower(s)? I am sure, on average, the ECMWF winds are similar to what is measured at the towers but since you are comparing hourly measurements, this may make a difference.

The ECMWF wind product is used here for 2 reasons: Firstly, our previous study has shown that the wind speeds provided by the ECMWF high-resolution operational forecasts (HRES) are, in general, closer to the observations than those provided by WRF (Lian et al., 2018). Secondly, the WRF model was run with two configurations (UCM and BEP urban canopy schemes) in this study. If we make use of the modeled winds, the UCM and BEP modeled CO_2 spatial differences should be analyzed using their corresponding modeled wind fields, and the observed winds are then needed for the analysis of the observed CO_2 spatial differences. However, given the small-scale wind variations reproduced by the model, it is hard to determine that the wind data at which station should be used in the analysis. For the purpose of a fair and uniform comparison, we thus use an independent wind product. The HRES with a horizontal resolution of about 16 km could provide a synoptic wind pattern as a proxy for all stations located within the IdF region. We have added the following sentence in the manuscript to account for the reviewer's comment:

"The HRES wind product is used here for two reasons: Firstly, our previous study has shown that the wind speeds provided by HRES are, in general, closer to the observations than those provided by WRF (Lian et al., 2018). Secondly, the WRF-Chem model was run with two configurations (UCM and BEP urban canopy schemes) in this study. If we make use of the modeled winds, the UCM and BEP modeled CO_2 spatial differences should be analyzed using their corresponding modeled wind fields, and the observed winds are then needed for the analysis of the observed CO_2 spatial differences. However, given the small-scale wind variations reproduced by the model, it is hard to determine that the wind data at which station should be used in the analysis. For the purpose of a fair and uniform comparison, we thus use an independent wind product."

Also, how much time does it take to traverse some of the towers that are farther apart (e.g. COU and SAC)? Did you compare observations from similar times or did you account for a lag in the measurements via travel time?

We ignore the time lag needed to transport information from upwind to downwind sites spanning the city by computing spatial gradients between concentrations at a given time. This is mainly due to the fact that the consideration such a time lag might be somewhat meaningless given the wind shear in the PBL during the afternoon when the mixing layer is usually well developed. Typical wind speed over Paris at 700 m above ground level is 7 m/s (25 km/h) and the distance between COU and SAC is approximately 38 km so that air masses take, on average, less than 2 hours to travel between the two sites at this height. Conversely, the wind speed at ground level is much smaller so that there is not a single time-lag that can be used. We thus assume that the analysis that is based on CO_2 concentration differences measured during the same 1-hour window is a minor issue. Note that, when comparing observation and models, the time lags are consistent.

Minor Comments:

Be specific as to what model you are using. I think in most cases you are referring to WRF-Chem models but there are others too such as VPRM, etc.

Following the reviewer's recommendation, we have attempted to make it clear and specific.

Grammar should be checked in many places throughout the article to improve clarity. Examples include lines 34 through 36 (page 2), ~10 (page 8).

As suggested, we have carefully done thorough English editing and corrected the grammatical mistakes in the revised manuscript.

Figures should be modified to improve clarity:

For example, Figure 1 should include a depiction of adjacent urban areas to show how remote AND, COU, OVS, and SAC are from ex-urban sources. This will help the reader know whether or not they sample "clean" air.

We feel that there is no need to add a depiction in Figure 1 for 3 reasons: 1) the distributions of the 1-km anthropogenic CO_2 emissions together with all in-situ measurement stations are shown in Figure 2a. 2) the dominant land use categories together with all in-situ measurement stations are shown in Figure S5. These two figures could be sufficient to provide such a depiction. 3) As we have mentioned in Section 2.1, the in-situ stations are installed on the rooftops or on towers to minimize the impact of local surface emissions. Moreover, the distance to the localized emissions was also taken into account as a necessary aspect in the design of this CO_2 monitoring network to ensure they sample air that is not in the immediate proximity of large anthropogenic emissions.

Figure 2 should include roads and other infrastructure in the second panel especially since the authors have made spent time discussing sectorial emissions. Also note (a) and (b) on Figure 2.

This suggestion is well taken. We have added these infrastructures into a second panel in Figure 1 where it might be more appropriate.

As suggested, we have noted (a) and (b) on Figure 2.

For Figure 4, zoom into similar area as in Figure 2 to show if VPRM is capturing urban biospheric flux which can significantly impact the urban fluxes especially their variability.

Figure R1 a high-resolution zoom of Paris and shows the daytime (06-18 UTC) average of CO_2 biogenic flux (NEE) in June 2016. Due to the 1-km SYNMAP land use data used for the VPRM model, the biogenic fluxes in Paris are almost zero except for a few grid cells containing two big parks that are located in the eastern and western Paris.

We thus feel that there is no need to make this high-resolution zoom of the GreenLITE™ covering areas since we have already mentioned it in Section 3.2.2 "The model simulates negative values of NEE (uptake of more than 5 $gCO_2/m^2/day$) over most of the region with the exception in urban areas where the values are assigned to zero."

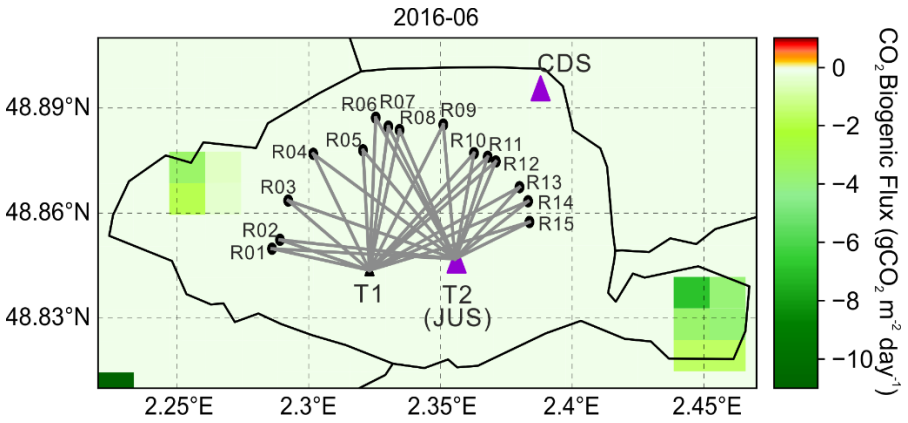

Figure R1. Daytime (06-18 UTC) average of CO_2 biogenic flux (NEE) over Paris in June 2016

The authors should consider moving Figure 5 to the supplemental information. It doesn't provide much information, especially given the timeframe that makes it hard to see, expect to show that the UCM transport model yields extreme outliers in the winter.

Following the reviewer's recommendation, we have moved Figure 5 to the supplement as Figure S7.

For Table 3, explain by what criteria did you color code the Tables. It seems like the better models for the correlation coefficients have "red" shading where int the RMSE and MBE the colors are switched (aka blue is better while red is worse). I would remove "all hours" to make the table clearer - all hours not really needed.

We agree with the reviewer that the color scales in Table 3 can be misleading. The color only represents the values from minimum (blue) to maximum (red) in the cells instead of indicating the goodness of fit between model and observation. We have added the following text in the caption of Table 3 in order to clarify this issue:

"The color highlights the value in the cell with the minimum in blue, the median in white and the maximum in red. All other cells are colored proportionally."

Following the reviewer's recommendation, we have removed "all hours" to make Table 3 clearer and put it in the supplement as Table S2.

Analysis of temporal and spatial variability of atmospheric CO_2 concentration within Paris from the GreenLITE™ laser imaging experiment

Jinghui Lian[1], François-Marie Bréon[1], Grégoire Broquet[1], T. Scott Zaccheo[2], Jeremy Dobler[3,a], Michel Ramonet[1], Johannes Staufer[4,b], Diego Santaren[1], Irène Xueref-Remy[5,b], and Philippe Ciais[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[2]Atmospheric and Environmental Research, Lexington, Massachusetts, United States
[3]Spectral Sensor Solutions LLC, Fort Wayne, Indiana, United States
[4]Thales, Labège, France
[5]Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE), Aix Marseille Université, CNRS, IRD, Avignon Université, Aix-en-Provence, France
[a]formerly at: Harris Corporation, Fort Wayne, Indiana, United States
[b]formerly at: Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

Correspondence to: Jinghui Lian (Jinghui.Lian@lsce.ipsl.fr)

Abstract. In 2015, the Greenhouse gas Laser Imaging Tomography Experiment (GreenLITE™) measurement system was deployed for a long-duration experiment in the center of Paris, France. The system measures near-surface atmospheric CO_2 concentrations integrated along 30 horizontal chords ranging in length from 2.3 km to 5.2 km and covering an area of 25 km^2 over the complex urban environment. In this study, we use this observing system together with six conventional in-situ point measurements and the WRF-Chem model coupled with two urban canopy schemes (UCM, BEP) at a horizontal resolution of 1 km to analyze the temporal and spatial variations of CO_2 concentrations within the Paris city and its vicinity for the 1-year period spanning December 2015 to November 2016. Such an analysis aims at supporting the development of CO_2 atmospheric inversion systems at the city scale. Results show that both urban canopy schemes in the WRF-Chem model are capable of reproducing the seasonal cycle and most of the synoptic variations in the atmospheric CO_2 point measurements over the suburban areas, as well as the general corresponding spatial differences in CO_2 concentration that span the urban area. However, within the city, there are larger discrepancies between the observations and the model results with very distinct features during winter and summer. During winter, the GreenLITE™ measurements clearly demonstrate that one urban canopy scheme (BEP) provides a much better description of temporal variations and horizontal differences in CO_2 concentrations than the other (UCM) does. During summer, much larger CO_2 horizontal differences are indicated by the GreenLITE™ system than both the in-situ measurements and the model results, with systematic east-west variations.

1 Introduction

Urban areas account for almost two-thirds of global energy consumption and more than 70% of carbon emissions (IEA, 2008). Human activities, such as fossil fuel burning (Duren and Miller, 2012) and cement production (Wang et al., 2012) produce a net increase of atmospheric CO_2 concentration within and downwind of the emission sources. Over the years, many instruments have been or will be used to measure the urban atmospheric CO_2 concentrations, including (i) ground-based monitoring networks in e.g., Paris (Xueref-Remy et al., 2018), Indianapolis (Davis et al., 2017), Los Angeles (Feng et al., 2016), Washington, DC (Mueller et al., 2017), Boston (Sargent et al., 2018); (ii) airborne campaigns conducted in e.g., Colorado (Graven et al., 2009), London (Font

et al., 2015); (iii) existing space-based measurements, e.g., GOSAT (Hamazaki et al., 2004), OCO-2 (Crisp et al., 2008, 2015) and (iv) future satellites with imaging capabilities, e.g., OCO-3 (Elderling et al., 2019), GeoCarb (Moore et al., 2018) and CO2M (Buchwitz, 2018). These observations are used or could be used for estimating emissions of CO_2 over large cities using atmospheric inverse modeling, or to detect emission trends if these data are collected over a sufficiently long period of time. High-accuracy continuous in-situ ground-based measurements of CO_2 concentrations, using the Cavity Ring-Down Spectroscopy (CRDS) technology, have been used in previous urban atmospheric inversion studies for the quantification of CO_2 emissions of large cities (Bréon et al., 2015; Staufer et al., 2016; Lauvaux et al., 2016; Feng et al., 2016; Boon et al., 2016; Sargent et al., 2018). However, many in-situ stations may be needed to accurately capture the CO_2 emission budget of a large city (Wu et al., 2016). Deploying such a network is expensive to install and maintain. The sparseness of CO_2 concentration sampling sites limits the ability of inversions to estimate the large spatial and temporal variations of the CO_2 emissions within the city, even though high-resolution emission inventories are available (e.g. AIRPARIF, 2013).

New concepts and technologies are desirable for a full sampling of atmospheric CO_2 concentrations within a city. These concepts may rely on moderate precision but low-cost sensors that could be deployed at many sites for a high spatial density sampling (Wu et al., 2016; Arzoumanian et al., 2019). An alternative to in-situ point measurements is a remote sensing system based on the spectroscopic techniques which could provide long-path measurements of atmospheric trace gases over extended areas of interest. An example of this is the differential optical absorption spectroscopy (DOAS). It has been applied to monitor atmospheric air pollutions such as nitrogen dioxide (NO_2) and aerosol in a complex urban environment (Edner et al., 1993). A novel laser absorption spectroscopy based system for monitoring greenhouse gases was developed by Spectral Sensor Solutions and Atmospheric and Environmental Research (AER). This system, known as the GreenLITE™, consists of a set of continuously operating laser-based transceivers and a set of retroreflectors separated by a few kilometers. Both data collection and data processing components are based on the Intensity Modulated Continuous Wave (IM-CW) measurement technique, which is described in detail in Dobler et al. (2017). This instrument provides estimates of the average CO_2 concentrations along the line of sight defined by the path between a laser-based transceiver and any given retroreflector. The path between a transceiver and a retroreflector is referred to as a "chord". The GreenLITE™ system was developed and deployed as part of several field campaigns over the past several years (Dobler et al., 2013; Dobler et al., 2017). These field tests have included extended operations at industrial facilities, and have shown that the system is capable of identifying and spatially locating point sources of greenhouse gases (CO_2 and CH_4) within a test area (~1 km^2). In conjunction with the 21st Conference of Parties to the United Nations Framework Convention on Climate Change (COP 21), the GreenLITE™ system was deployed for a long-duration field test over central Paris, France. The objective was to demonstrate the potential of CO_2 concentration measurements along 30 horizontal chords ranging in length from 2.3 km to 5.2 km and covering an area of 25 km^2. The aim of this field campaign was to demonstrate the ability of GreenLITE™ to monitor the temporal and spatial variations of near-surface atmospheric CO_2 concentrations over the complex urban environment. In addition, these measurements may be used for post-deployment analysis of the CO_2 distribution with the ultimate goal of revealing the CO_2 emission distribution. As a first step, the objectives of this work are to assess the information content of the GreenLITE™ data, to analyze the atmospheric CO_2 distribution and to characterize precisely the processes that lead to dilution and mixing of the anthropogenic emissions, which can provide new insights compared to the present in-situ point measurement approaches due to a much wider spatial coverage.

The collection of the GreenLITE™ atmospheric CO_2 measurements in Paris makes it possible to evaluate and potentially improve meteorological and atmospheric transport models coupled to CO_2 emission inventories. On the other hand, the modeling system is expected to provide interpretations of the temporal and spatial variations of the GreenLITE™ data, with the aim of supporting the development of CO_2 atmospheric inversion systems at the city scale. Here we compare GreenLITE™ CO_2 data with simulations

performed with the Weather Research and Forecasting Model coupled with a chemistry transport model (WRF-Chem). The WRF-Chem model allows various choices of physics parameterizations and data assimilation methods for constraining the meteorological fields (Deng et al., 2017; Lian et al., 2018). Previous studies have shown that it is necessary to account for specific urban effects when modeling the transport and dispersion of CO_2 over complex urban areas such as Salt Lake City, UT and Los Angeles, CA (Nehrkorn et al., 2013; Feng et al., 2016). Nevertheless, even when the urban environment is accounted for, the modeling of atmospheric transport is a challenge. Significant mismatches remain between modeled and measured concentrations that could be explained by transport biases, particularly at night, and vertical mixing during the day.

In this study, we present the results from a set of 1-year simulations (from December 2015 to November 2016) of CO_2 concentrations over the Paris megacity based on the WRF-Chem model coupled with two urban canopy schemes at a horizontal resolution of 1 km. The simulated CO_2 concentrations are compared with observations from the GreenLITE™ laser system as well as in-situ CO_2 measurements taken continuously at six stations located within the Paris city limits and surrounding area. The detailed objectives of this paper are: (i) to analyze in detail the information content of the GreenLITE™ data in addition to conventional in-situ CO_2 measurements in order to better understand the temporal and spatial variations of near-surface CO_2 concentrations over Paris and its vicinity; (ii) to evaluate the performance of the high-resolution WRF-Chem model coupled with two urban canopy schemes (UCM, BEP) for the transport of CO_2 over the Paris megacity area based on the two types of CO_2 measurements; (iii) to discuss the potential implications of assimilating the GreenLITE™ data into the CO_2 atmospheric inversion system with the ultimate goal of increasing the robustness of the quantification of city emissions and constraining the spatial distribution of the emissions within the urban area.

This paper is organized as follows: Section 2 provides more details about the GreenLITE™ deployment in conjunction with the in-situ CO_2 monitoring network in Paris. The WRF-Chem modeling framework and model configurations are presented in Section 3. In Section 4, we evaluate the performance of the WRF-Chem simulations based on the analyses of the temporal and spatial patterns of observed and modeled CO_2 concentrations. Discussions and conclusions are given in Section 5.

2 The observation network

2.1 In-situ measurements

Since 2010, a growing network of three to six in-situ continuous CO_2 monitoring stations has been established in the Île-de-France (IdF) region in coordination with ongoing research projects (e.g., Bréon et al., 2015; Xueref-Remy et al., 2018). These observations are used to understand the variability of atmospheric CO_2 concentrations, with the aim to improve the existing bottom-up CO_2 emission inventories by providing a top-down constraint through atmospheric inverse modeling. The stations are equipped with high-precision CO_2/CO/CH_4 analyzers installed on rooftops or towers to increase the area of representativity. All instruments have been regularly calibrated against the WMO cylinders (WMO-CO_2-X2007 scale) (Tans et al., 2011).

The locations of the stations are given in Table 1a and are shown in Figure 1a. Four stations are located within the peri-urban area: OVS site is located about 26 km southwest of Paris center with the sampling height of 20 m above the ground level (AGL) on the top of a building. The SAC tall tower is located on the Plateau de Saclay (9.5 km southeast of OVS) with two air inlets placed at 15 m and 100 m AGL respectively. The other two sites are located at the north (AND) and north-east (COU) edges of the Paris urban area in a mixed urban-rural environment with single inlets at 60 m and 30 m AGL respectively. These four peri-urban stations are complemented by in-situ continuous measurements at two urban stations: one at the Cité des Sciences et de l'Industrie (CDS) and one at the former Pierre and Marie Curie University (now Sorbonne University, also called Jussieu; JUS). The inlets for each of the sensors are placed at approximately 34 m and 30 m AGL respectively. The JUS station is on the roof of a building close to

ventilation outlets and may be influenced by this and other localized sources of CO_2. The JUS site was only measuring CO_2 continuously from January to April 2016, and from September 16th 2016 through the end of this study. The spatial distribution of the monitoring sites was chosen a priori to best enable the analysis of gradients due to emissions in Paris when the wind is blowing from either the south-west or north-east directions, which corresponds to the prevailing winds in the region (Bréon et al., 2015; Staufer et al. 2016; Xueref-Remy et al, 2018).

2.2 The GreenLITE™ campaign over Paris

The GreenLITE™ system was deployed in Paris in November 2015 as a proof-of-concept demonstration during the COP 21 conference, and kept operating for one year. This system used two transceivers coupled with 15 retroreflectors to measure the CO_2 concentrations along 30 intertwined lines (chords) of 2.3-5.2 km length covering an area of 25 km^2 over the center of Paris. Each transceiver used two fiber-coupled distributed feedback lasers to generate an absorption line at a wavelength of 1571.112 nm and an offline with significantly lower absorptions (nominally 1571.061 nm). The experimental design and layout examined in this study are given in Table 1b and are illustrated in Figure 1b. The two transceivers were located on two rooftops, one on the lower of the two Montparnasse buildings (T1) (50.3m AGL) and the other on the Jussieu tower (T2) (86.8m AGL) located near the JUS in-situ instrument. These locations were chosen based on a clear line of sight to the retroreflectors which were installed on additional rooftops around the city with heights varying from 16.8-50.4 m AGL. For this implementation, each transceiver scanned the retroreflectors in sequence and made a transmission measurement of each chord with a period of four minutes. The experiment lasted from November 2015 to November 2016 with some sporadic down time of either the transceivers and/or some of the reflectors.

Preliminary analysis shows that the original GreenLITE™ CO_2 concentrations have a slow drift of approximately +/-5 ppm in comparisons to both the nearby in-situ measurements (Figure S1) and simulations with the CHIMERE-ECMWF transport configuration presented in Staufer et al. (2016). These slowly time-varying differences were most likely due to a slight systematic long-term drift in both the on- and off-line wavelengths as a function of continuous operations. Such drift may induce some non-linear impacts on the measured concentrations. It is therefore more appropriate to adjust the wavelengths rather than to apply a linear calibration to the retrieved concentrations. Unlike in-situ point measurement systems, there is no established method for calibration of long open-path systems to the WMO mole fraction scale used as an international standard for atmospheric CO_2 monitoring (Tans et al., 2011). Therefore, a bias correction method was developed by AER (Zaccheo et al., 2019) for addressing observed slowly drifting biases between the GreenLITE™ prototype system and the two in-situ sensors (CDS and JUS) that are near the GreenLITE™ chords. This method computed a time-varying adjustment to the offline wavelength based on a non-linear optimization mechanism. This non-linear approach adjusts the GreenLITE™ offline wavelength considering not only the average values of hourly CO_2 concentrations at two in-situ stations, but also the corresponding average temperature, relative humidity, atmospheric pressure along the chord and an optimized online wavelength value during the measurement period. Finally, the median on- and off-line values over a 4-day window was used to recompute the GreenLITE™ data from all chords using a radiative transfer based iterative retrieval scheme based on the LBLRTM model (Clough et al., 2005). Even though this approach is not ideal as the two in-situ stations and the GreenLITE™ system do not sample the exact same area, it does provide a well-defined mechanism that reduces the systematic long-term biases with no significant impact on the chord-to-chord variations. Top panels in Figure S2 (a) and (b) show the distribution of the absolute values of the daily averaged CO_2 concentration difference between all pairs of chords for each transceiver before and after the calibration. The differences between the medians of the re-processed and original inter-chord range, shown in bottom panels, are within in the range of ±0.5 ppm for T1 and ±2 ppm for T2 with the respective yearly mean plus/minus one standard deviation of 0.04 ± 0.16 ppm for T1 and 0.48 ± 0.43 ppm for T2.

In order to enable the data to be compared to hourly in-situ observations and WRF-Chem outputs, hourly means are computed from the 4-minute GreenLITE™ data after applying the calibration approach described above. Two addition selection criteria were also established for this work: (i) A minimum of 3 valid 4-minute samples were necessary to generate a valid hourly average for a given chord, and (ii) the standard deviation of these samples had to be smaller than 10 ppm. The 10 ppm threshold was selected to be roughly 3 times the typical standard deviation of the 4-minute measurements for any given chord within a one-hour period (Figure S3). Data that do not meet the above criteria, about 1.06 % of the total, were considered invalid and excluded from further analysis.

3 Modeling framework

3.1 WRF-Chem model setup

A set of high-resolution simulations of atmospheric CO_2 concentrations was performed with WRF-Chem V3.9.1 online coupled with the diagnostic biosphere Vegetation Photosynthesis and Respiration Model (VPRM) (Mahadevan et al., 2008; Ahmadov et al., 2007, 2009). The simulations were carried out over the period spanning September 2015 to November 2016, in which the first three months were considered as a spin-up period. Three one-way nested domains were employed with the horizontal grid resolution of 25, 5 and 1 km, covering Europe (Domain 01), Northern France (Domain 02) and the IdF region (Domain 03) respectively (Figure S4). The meteorological initial and lateral boundary conditions were imposed using the ERA-Interim global re-analyses with $0.75\,° \times 0.75\,°$ horizontal resolution and 6 hourly intervals (Berrisford et al., 2011). We nudged the 3D fields of temperature and wind to the ERA-Interim reanalysis in layers above the planetary boundary layer (PBL) of the outer two domains using the grid nudging option in WRF. We also assimilated observation surface weather station data (ds461.0) and upper-air meteorological fields (ds351.0) from the Research Data Archive at the National Center for Atmospheric Research (https://rda.ucar.edu/datasets/ds351.0/; https://rda.ucar.edu/datasets/ds461.0/) using a nudging technique (the surface analysis nudging and observation nudging options of WRF are described in detail in Lian et al., 2018). Details regarding the model configurations used in this study are summarized in Table 2.

The urban canopy parameterization is a critical element in reproducing the lower boundary conditions and thermal structures, which are of vital importance for accurate modeling of the transport and dispersion of CO_2 within the urban areas. We therefore paid special attention, in this study, to examine the impact of the two available urban canopy schemes on WRF-Chem transport results, namely the single-layer Urban Canopy Model (UCM) (Chen et al., 2011) and the multilayer urban canopy model Building Effect Parameterization (BEP) (Martilli et al., 2002). This study does not assess the multilayer urban parameterization BEP+BEM (BEP combined with the Building Energy Model (BEM)) (Salamanca et al., 2010) since this parameterization focuses on the impact of heat emitted by air conditioners, which are not commonly used in Paris. This study used 34 vertical layers in WRF-UCM with the top model pressure set at 100 hPa, and 15 layers arranged below 1.5 km with the first layer top at approximately 19 m AGL. In order to take full advantage of the WRF-BEP configuration, it is necessary to have a fine discretization of the vertical levels close to the surface. This configuration with 44 vertical layers, places 25 of them within the lowest 1.5 km with the lowest level being around 3.8 m AGL. In order to select an adequate model physical configuration for Paris, we carried out some preliminary sensitivity experiments to test the impact of different physical schemes on the simulated CO_2 concentrations. These tests use up to five different PBL schemes and two urban canopy schemes. The simulations were carried out for two months, including one winter month (January 2016) and one summer month (July 2016). These preliminary sensitivity results indicate that different PBL schemes in the WRF-Chem model lead to monthly average differences of 2-3 ppm on the simulated CO_2 concentrations over Paris, whereas the two different urban canopy schemes lead to much larger differences of 8-10 ppm. Thus in

this study, we carried out the 1-year simulation with two different urban canopy schemes as they are sufficient to address the paper main question regarding the ability of a configuration of the WRF-Chem model to simulate the CO_2 atmospheric transport in an urban environment, but also to provide an estimate of the modeling uncertainty. All of the other physics options remained the same for the two experiments (Table 2): WSM6 microphysics scheme (Hong and Lim, 2006), RRTM longwave radiation scheme (Mlawer et al., 1997), Dudhia shortwave radiation scheme (Dudhia, 1989), MYJ PBL scheme (Janjić, 1990, 1994), Eta Similarity surface layer scheme (Janjić, 1996), Unified Noah land-surface scheme (Chen and Dudhia, 2001). The Grell 3D ensemble cumulus convection scheme (Grell and Dévényi, 2002) was applied for Domain 01 only in both experiments.

3.2 CO_2 simulations

3.2.1 Anthropogenic CO_2 fluxes

Anthropogenic CO_2 fluxes within the IdF region are imposed using the AirParif inventory for the year 2010 at spatiotemporal resolutions of 1 km and 1 h (AIRPARIF, 2013). This inventory is based on various anthropogenic activity data, emission factors and spatial distribution proxies, which are described in detail in Bréon et al. (2015). It provides maps and diurnal variations for five typical months (January, April, July, August, and October) and three typical days (a weekday, Saturday and Sunday) to account for the seasonal, weekly and diurnal cycles of the emissions (see Figure 3, Bréon et al., 2015). CO_2 emissions from fossil fuel CO_2 sources outside the IdF region are taken from the inventory of the European greenhouse gas emissions, together with country-specific temporal profiles (monthly, daily and hourly) at a spatial resolution of 5 km (updated in October 2005). This inventory was developed by the Institute of Economics and the Rational Use of Energy (IER), University of Stuttgart, under the CarboEurope-IP project (http://www.carboeurope.org/).

Both inventories are adapted to the WRF-Chem model for the period of simulation (2015.09-2016.11). Moreover, we scale these two data sets to account for annual changes in emission between the base years and simulation timeframe. This is accomplished by rescaling the maps with the ratio of the annual budgets of national CO_2 emissions for the countries within the domain between the base year 2005 for IER and 2010 for AirParif and the year of simulation (2015/2016), taken from Le Quéré et al. (2018) (https://www.icos-cp.eu/GCP/2018. See also Table S1 in the supplement for details about original data sources). Finally, we interpolate the emissions onto the WRF-Chem grids, making sure to conserve the total budget of emission in the process, as done in previous studies (e.g. Ahmadov et al., 2007). Note that for the point sources such as stacks, industries and mines, CO_2 emissions are distributed over a single grid cell corresponding to their locations. Figure 2 shows the spatial distribution of the total CO_2 emissions for a weekday in March over the IdF region at the resolution of 1×1 km^2. It can be seen that there is a large spatial variability of CO_2 emissions ranging from 0 to more than 600 gCO_2/m^2/day in this area and the largest emissions are concentrated over the Greater Paris area, accounting for about 50% of the emitted CO_2.

Based on the analysis of sectoral specific fossil fuel CO_2 emissions over the IdF region by Wu et al. (2016), we group the detailed sectoral AirParif emissions into five main sectors, namely building (43%), energy (14%), surface traffic (29%), aviation-related surface emissions (4%), and all other sectors (10%), where the percentages in parenthesis express the relative contribution of each sector to the yearly total. All emissions are injected in the first model layer. Distinct CO_2 tracers are used for each of the five main sectors in the transport model to record their distinct CO_2 atmospheric signature. Figure 3 shows averages at the monthly scale of emissions below the GreenLITE™ chords for those different sectors. It illustrates that CO_2 emissions have a large seasonal cycle, mostly due to the residential heating (the "building" sector) which is strongly driven by variations of the atmospheric temperature. Figure 3 also reveals lower emissions for those chords (TX and R01-03) in the west of Paris than those in the other quadrants.

3.2.2 Biogenic CO_2 fluxes

Biogenic CO_2 fluxes are simulated with the VPRM model forced by meteorological fields simulated by WRF, and online-coupled to the atmospheric transport. VPRM uses the simulated downward shortwave radiation and surface temperatures, along with the vegetation indices (EVI, LSWI) derived from the 8-day MODIS Surface Reflectance Product (MOD09A1) and four parameters for each vegetation category (PAR0, λ, α, β) that are optimized against eddy covariance flux measurements over Europe collected during the Integrated EU project "CarboEurope-IP" (http://www.bgc-jena.mpg.de/bgc-processes/ceip/). The land cover data used by VPRM (see Figure S5) are derived from the 1-km global Synergetic Land Cover Product (SYNMAP, Jung et al., 2006) reclassified into 8 different vegetation classes (Ahmadov et al., 2007, 2009).

Figure 4a shows the spatial distribution of daytime-averaged (06-18 UTC) CO_2 biogenic flux (NEE with a negative sign indicating net CO_2 uptake by the vegetation surface) in June 2016. The model simulates negative values of NEE (uptake of more than 5 $gCO_2/m^2/day$) over most of the region with the exception in urban areas where the values are assigned to zero. Figure 4b shows the mean diurnal cycles of NEE for 12 calendar months and for 8 vegetation classes used in VPRM over Domain 03. The magnitude of NEE is highly dependent on the vegetation types, although the diurnal cycles are similar across these vegetation types. From November to January, the VPRM estimates within the IdF region show a small diurnal cycle and a positive NEE explained by ecosystem respiration exceeding gross primary productivity. One exception to positive wintertime NEE is for evergreen trees which, according to the VPRM model, sustain enough gross primary productivity to keep a negative daytime NEE throughout the year. The model shows large CO_2 uptake between late spring and early summer. Note that the seasonal cycle of crops, which dominates over the IdF region, is somewhat different from that of forests, with a NEE that decreases after the harvest in June/July, this crop phenology signal is being driven by the MOD09A1 data. Grasses also have a shorter uptake period than the other vegetation types, with a positive NEE as early as August.

3.2.3 Initial and lateral boundary conditions for CO_2

Initial and lateral boundary conditions for CO_2 concentration fields used in the WRF-Chem model are taken from the 3-hourly fields of the CAMS global CO_2 atmospheric inversion product (Chevallier, 2017a, 2017b) with a horizontal resolution of 3.75 °×1.90 °(longitude × latitude) and 39 vertical levels between the surface and the tropopause.

4 Results

4.1 Time series and general statistics

The continuous CO_2 concentration measurement network in the IdF region provides an invaluable opportunity for model validation and data interpretation. In this work, the correlation coefficient, root-mean-square error (RMSE) and mean bias error (MBE) metrics are first used to compare the performance of the WRF-Chem model with respect to the observed CO_2 concentrations from both the GreenLITE™ laser system and in-situ continuous stations. In order to compare them with the GreenLITE™ measurements, the modeling results are sampled and integrated along the chord lines, accounting for their positions and heights. For the in-situ point measurements, we simply use the CO_2 values from the 1-km WRF-Chem grid cell that contains the observation location.

Table 3, together with Table S2 and Figure S6 in the supplement, shows the statistics of all the hourly differences between the observed and modeled CO_2 concentrations and the hourly afternoon differences (11-16 UTC), from December 2015 to November 2016 using the two model configurations (UCM, BEP). The results presented in the Taylor diagrams (Figure S6) are based on the full year of data and the seasonal statistics are summarized in Table 3. In general, the model performance is better during the afternoon, both in terms of correlation and RMSE, than it is for the full day. These results are consistent with previous findings

that show the model has little skills at reproducing the CO_2 fields during the nighttime due to poor representation of vertical mixing during nighttime conditions, and in the morning due to inadequate depiction of PBL growth (e.g. Bréon et al., 2015; Boon et al. 2016). Given the better performance of the WRF-Chem model in the afternoon, we focus the following analyses on CO_2 concentrations acquired during this period of the day only.

The other significant feature is that the UCM scheme shows a large positive bias (8.7-19.6 ppm) with respect to the observations within the city during autumn and winter. In contrast, the statistics for the BEP scheme compared to the observations are significantly better with clear improvements in the correlation and substantial decreases in both the RMSE and MBE. It is well known that the lower part of the atmosphere is, on average, more stable in winter than in summer (Gates, 1961). As a consequence, a significant fraction of the emitted CO_2 remains close to the surface, so that its atmospheric concentrations is, in winter, highly sensitive to local fluxes and variations in vertical mixing, especially in the complex urban areas. The statistics are highly dependent on the choice of the urban canopy scheme, which strongly suggests that the large UCM model-measurement mismatches in winter are linked to difficulties in modeling the vertical mixing within the urban canopy. It is worth noting that CO_2 concentrations are better reproduced by both UCM and BEP in the spring, with correlations that fluctuate between 0.51 and 0.82 across stations. Both urban canopy schemes show lower correlations during summer (0.45-0.63). These lower values are mostly due to the smaller variability of the concentration rather than a higher measurement-model mismatch. Moreover, the UCM and BEP also have comparable performances at peri-urban areas while the BEP is slightly better at some suburban sites as shown by the statistics. The smallest errors (both in terms of RMSE and bias) are found at Saclay with a measurement inlet that is well above the sources at 100 m AGL (SAC100).

The statistics shown in Table 3, Table S2 and Figure S6 also indicate the ability of the models to reproduce the CO_2 at two urban in-situ stations (JUS & CDS) and the GreenLITE™ measurements. As for the GreenLITE™ data, we first compute the hourly averages of the observed and modeled CO_2 concentrations over all 15 chords for each transceiver (T1 and T2), and then calculate the respective statistics. In general, the model performance is similar for the two types of urban measurements, whereas the performance for urban measurements is slightly inferior to that of the suburban (both in terms of RMSE and correlation). The correlations with observations are better for T1 and T2 than for the two urban in-situ sites, which may be due to the fact that T1 and T2 represent an average over a wide area. Therefore, the GreenLITE™ data are less sensitive to local unresolved sources than the in-situ measurements. The RMSE with the BEP scheme is within the range of 4.5 to 9.6 ppm for T1 which is substantially superior to those of JUS and CDS, with only one exception at CDS during summer when the value is slightly better for CDS than for T1. In terms of the MBE, the values of T1 are similar with those of CDS, while the BEP simulation reveals an underestimation of CO_2 for T2 and JUS, with a negative bias of up to 5.2 ppm.

Figure S7 shows time series of modeled CO_2 against daily afternoon mean GreenLITE™ observations (11-16 UTC). Again, it clearly illustrates that the UCM scheme overestimates the CO_2 concentrations close to the surface within the city during winter. The BEP scheme effectively reproduces the seasonal cycle, as well as most synoptic variations of the atmospheric CO_2 measurements. Note that the UCM model-observation discrepancies for T2 are much smaller than those of T1 as the transceiver T2 is 36.5 m higher in altitude, whereas such a difference in modeled CO_2 between T1 and T2 is not obvious for the BEP scheme.

4.2 Analyze co-variations of CO_2 spatial difference with wind

In this section, we analyze the spatial variations of the CO_2 concentrations that are: (i) measured at the in-situ stations, (ii) provided by the GreenLITE™ system and (iii) simulated by the WRF-Chem model. The analysis of spatial differences rather than individual values should strongly reduce the signature of the large-scale pattern due to boundary conditions, and better highlight that of the

Paris emissions (Bréon et al. 2015). This makes it possible to further evaluate some characteristics of the model and the measurement data.

4.2.1 In-situ measurement

We analyze the horizontal differences between pairs of in-situ stations as a function of wind speed and direction, expecting a larger concentration at the downwind station with respect to the upwind station, in this region of high emission. For wind fields, we use the ECMWF high-resolution operational forecasts (HRES) linearly interpolated at the hourly resolution, and extracted at a height of around 25 m AGL (https://www.ecmwf.int/en/forecasts/datasets/set-i) as a proxy for all stations located within the IdF region. The HRES wind product is used here for two reasons: Firstly, our previous study has shown that the wind speeds provided by HRES are, in general, closer to the observations than those provided by WRF (Lian et al., 2018). Secondly, the WRF-Chem model was run with two configurations (UCM and BEP urban canopy schemes) in this study. If we make use of the modeled winds, the UCM and BEP modeled CO_2 spatial differences should be analyzed using their corresponding modeled wind fields, and the observed winds are then needed for the analysis of the observed CO_2 spatial differences. However, given the small-scale wind variations reproduced by the model, it is hard to determine that the wind data at which station should be used in the analysis. For the purpose of a fair and uniform comparison, we thus use an independent wind product. Furthermore, the hourly afternoon CO_2 data are classified into the wind classes with a bin-width of 1 m/s for wind speed and 11.25 ° for wind direction. Figure 5 shows the patterns of the observed and modeled CO_2 concentration differences between pairs of in-situ stations, averaged accounting for the wind classes. The standard deviations of CO_2 concentration differences for each wind class are shown in Figure S8.

Figure 5a shows the observed and modeled CO_2 horizontal differences between AND and COU, two suburban stations located to the north of the Paris city. One expects that stations downwind of sources of emissions would have a higher CO_2 concentration than those upwind so that the sign of the difference should vary with the wind direction. For this pair of sites (AND and COU), both the model and observations show the expected pattern with a similar amplitude. The values of RMSE and MBE are 4.53 and -0.14 ppm respectively for the BEP scheme, implying a slightly better performance than the UCM scheme (6.34 and -0.47 ppm respectively).

Figure 5b and 5c show similar figures but for the CO_2 differences of (COU-SAC) and (CDS-SAC). The Paris city is located between both pairs of stations when the wind is roughly from the north-east or from the south-west directions. Both COU and SAC are located outside of the city and show a pattern with fairly symmetric positive and negative values. Conversely, CDS is in the Paris city, within an urban environment, and is strongly affected by significant urban emissions from its surroundings. As a consequence, the CDS-SAC differences in concentration are mostly positive for all wind sectors, with the exception of very specific wind conditions (low winds in the 45 ° north-east sector). The wind speed also has a strong influence on the differences. The CO_2 difference signal and its variability are generally larger for smaller wind speeds. The model plots (second and third rows) illustrate that the models reproduce well the expected cross-city upwind-downwind differences in CO_2 concentrations. In term of signal amplitude, the BEP scheme is also in better agreement with the observations than the UCM scheme, which is particularly true for the standard deviations shown in Figure S8.

Conversely, both urban canopy schemes fail to reproduce the wind-related pattern of the observed CDS-JUS difference (Figure 5d). These observed differences do not show any upwind-downwind patterns and are mostly negative, which can be expected since JUS is close to the city center where strong emissions impact the concentration, whereas CDS is in the middle of a park and is therefore less affected by emissions from its surroundings. The model pattern is dominated by the simple upwind-downwind structure and it is very much different from the observed values, especially when the winds are out of west to south-west, where the model values are positive and the observed differences are strongly negative. This model-measurement discrepancy is likely

the result of a poor description of the emissions in the city center that are not well reproduced by the 1-km resolution inventory with periodic temporal profiles. It may also indicate that the complex urban structure and morphology, such as buildings and street canyons affect the energy budget and atmospheric transport, all of which lead to fine-scale (sub-kilometer) CO_2 concentration features that cannot be captured by the WRF-Chem model at a 1-km horizontal resolution. The in-situ point measurement may then not be representative of the average within the larger area (1 km^2) that is simulated by the model.

The analysis of the in-situ point measurement differences within and around Paris, together with the simulations, indicates that the model reproduces both the general structure and the amplitude of the cross-city differences in CO_2 concentrations and the CO_2 difference in the Paris surroundings, but that it fails to simulate CO_2 differences between the two stations located in the inner city.

4.2.2 GreenLITE™ measurement

One expects that the GreenLITE™ principle, that provides averaged CO_2 concentrations along the chord lines, is less affected by the local unresolved sources of CO_2 emissions than the in-situ point measurements. Meanwhile, the wide spatial coverage of the GreenLITE™ system is expected to provide additional information about CO_2 spatial variations within the Paris city. In this section, we focus on the spatial variation of CO_2 concentration measured with the GreenLITE™ system. As a first step, we analyze the distribution of the absolute values of the observed hourly afternoon CO_2 difference between all pairs of chords for each month together with their simulated counterparts shown in Figure 6.

We first focus on the winter period (December to February). During that period, the median value of the measured T1 inter-chord range is mostly on the order of 2 ppm. That of T2 is somewhat larger, on the order of 3-4 ppm with some excursions up to 9 ppm. The two simulations with UCM and BEP respectively show very large differences. Whereas BEP simulates spatial variations that are of the right order of magnitude compared to the GreenLITE™ data, those of UCM are much larger. Thus, the GreenLITE™ measurements provide clear information that favors the BEP over the UCM. During the winter period, there is little vertical mixing which leads to large vertical gradients in CO_2 concentrations close to the surface. The two simulations differ in their representations of this mixing which leads to large differences in the modeled CO_2 concentrations. Figure S9 shows that the UCM scheme reproduces a much larger vertical gradient in CO_2 concentrations close to the surface, a few tens meters above the emissions than the BEP scheme does during afternoon (11-16 UTC). The differences are not as large higher up, neither are they further downwind of the emissions as the vertical gradient is then smoother as a result of mixing.

During the summer period, solar insulation generates more instability and the convection generates vertical mixing that limits the horizontal gradients. Both simulations indicate an inter-chord range of less than a few ppm. Conversely, the GreenLITE™ data indicate much larger values, of 3-4 ppm (the median) for T1 and even larger for T2. Further analysis indicates that this spatial variation is mostly systematic, i.e. that some chords are consistently lower or higher than the in-situ values. At this point, there are three hypotheses:

- H1 The spatial differences of T1 and T2 are true features linked to fine-scale spatial variations of the emissions between the west and east part of Paris, that are under-represented or not included in the emission inventory;
- H2 The models fail in the description of CO_2 concentrations within the Paris city because of imperfect representations of atmospheric transport processes, excluding inaccuracies in emissions;
- H3 There is a chord-dependent bias in some of the GreenLITE™ chords during the summer period.

To resolve this question, we look at the spatial difference between the in-situ sites within the city (JUS-CDS) during summer. Unfortunately, the JUS instrument was not working during the summer of 2016. Therefore, we use the JUS and CDS data over the summers from December 2015 to December 2018 (Figure 6c). In general, the modeled CO_2 concentration differences between pairs of in-situ stations are larger than the modeled inter-chord range of the GreenLITE™ system. During the summer, the observed

absolute differences between JUS and CDS are only of a few ppm (the median is on the order of 2 ppm during July and August). These observations indicate that the spatial differences of CO_2 between these two sites within the Paris city are much smaller during the summer than during the winter, and tend to support the modeling results, which would undermine the hypotheses H1 and H2. However, these two stations do not sample the western part of Paris that is less densely populated with a higher fraction of green areas. The in-situ observations do not fully rule out, therefore, the possibility of an impact of the emission spatial structure. Another potential source of measurement-model discrepancy is the atmospheric transport modeling as proposed in H2. According to previous studies (e.g. Hu et al., 2010), the turbulent eddies and thermals are unlikely to be reproduced properly by the local closure MYJ PBL scheme, which results in insufficient vertical mixing under convective (unstable) conditions, i.e. during summer. It may also indicate that the WRF-Chem model at a 1-km horizontal resolution cannot reproduce the fine-scale (sub-kilometer) CO_2 concentration features over a complex urban environment in Paris, as the analysis of JUS and CDS in-situ measurements has shown in Section 4.2.1.

Atmospheric transport simulations make it possible to assess the respective contributions of various areas/sectors to the measurements. Our preliminary sensitivity experiments (see Figure S10 and S11 for details) have shown that the anthropogenic emission from the Greater Paris area is the dominant contribution (~80%) to the anthropogenic CO_2 signal at the urban measurement stations. In order to get further insights into the characteristics of CO_2 spatial variations within the Paris city, it is therefore necessary to analyze the CO_2 differences with the consideration of the anthropogenic CO_2 emissions shown in Figure 2 and Figure 3. We thus group the 15 chords from T1 into three bins according to both their geographic locations and the amounts of anthropogenic CO_2 emissions averaged along the chords: the western, middle and eastern parts consist of reflectors R01, R02, R03, reflectors R06, R07, R08, and reflectors R13, R14, R15 respectively overlying three different regions within Paris. Figure 7 shows the co-variations of the GreenLITE™ observed and modeled CO_2 spatial difference with winds. The standard deviations of CO_2 concentration differences for each wind class are shown in Figure S12.

In Figure 7b and 7c, we show the east-west and the middle-west differences, where the CO_2 anthropogenic emissions in the western part are systematically lower than the other two regions, the observed CO_2 concentrations in the middle and east are on average higher than the west. The patterns of observed CO_2 difference are characterized by positive values no matter where the wind blows. The CO_2 differences reproduced by the model are positive in the southwest direction, however, it shows a nearly opposite pattern with those from observations when the wind is from the northeast. A plausible explanation for this is that the influence of km-scale anthropogenic emissions over different parts of Paris on the observed CO_2 concentration has a greater effect than the atmospheric transport and dispersion of the fluxes over the period of study.

Figure 7a shows similar figures but for the east-middle difference. There is a better measurement-model agreement than for Figure 7b and 7c. Indeed, the spatial variations of CO_2 concentrations show, as expected, negative values over upwind directions and positive values over downwind directions both for the observation and the model. According to the inventory, the two Paris areas that are covered by the set of chords used here have similar anthropogenic emissions. As a consequence, the overall CO_2 concentration difference, as shown in Figure 7a, is then better linked to the impact of atmospheric transport.

We therefore conclude that the pattern of CO_2 concentration differences is consistent with winds only over the areas with similar anthropogenic emissions. In other words, if we compare the CO_2 concentrations of the chords overlaying different level of emissions, the model may be insufficient in accurately modulating the dispersion of CO_2 emissions, the ventilation and dilution effects at such a high urban microscale resolution.

5 Summary and Conclusions

In this study, we use conventional in-situ together with novel GreenLITE™ laser measurements for an analysis of the temporal and spatial variations of the CO_2 concentrations within the Paris city and its vicinity. The analysis also uses 1 km-resolution WRF-Chem model coupled with two urban canopy schemes, for the 1-year period from December 2015 to November 2016.

Results show that two urban canopy schemes (UCM, BEP) as part of the WRF-Chem model show similar performances in the areas surrounding the city. They are capable of reproducing the seasonal cycle and most of the synoptic variations in the atmospheric CO_2 in-situ measurements over the suburban areas, as well as the general corresponding spatial differences in CO_2 concentration between pairs of in-situ stations that span the urban area.

Within the city, these results show very distinct features during winter and summer:

During the winter, the emissions within the city are the highest, mainly due to households heating, and the vertical mixing is low. This combination leads to large temporal, vertical and horizontal variations of CO_2 concentrations. The GreenLITE™ measurements are less sensitive to local unresolved sources than the in-situ point measurements, and are then better suited for the comparison to km-scale modeling. In our analysis, the GreenLITE™ data are used to clearly demonstrate that the BEP scheme provides a much better description of the CO_2 fields within the city than the UCM scheme does.

During the summer, the emissions are lower (by a factor of roughly two compared to the cold season) and the sun-induced convection makes the vertical mixing much faster than in winter. For this period, both the in-situ measurements and the modeling indicate that, during the afternoon, the spatial differences are limited to a few ppm. Much larger spatial differences are indicated by the GreenLITE™ system, with systematic east-west variations. Although it is not yet fully understood, several evidences suggest an increase of measurement noise and bias in some of the GreenLITE™ chords during the summer season, that must be resolved or reduced before assimilating the whole dataset into the CO_2 atmospheric inversion system that aims at retrieving urban fluxes.

This study stresses the difficulty in reproducing precisely the atmospheric CO_2 concentration within the city because of our inability to represent the detailed spatial structure of the emission and because of the sensitivity of the CO_2 concentration to the strength of vertical mixing. There are strong indications that the uncertainty on the vertical mixing is much larger than the uncertainty on the emissions so that atmospheric concentration measurements within the city can hardly be used to constrain the emission inventories.

Author contribution

JL, FMB, GB and PC contributed to the design and implementation of the research. JL, FMB, GB, PC, TSZ and JD contributed to the analysis and interpretation of the results. TSZ, JD, MR and IXR performed the measurements. JS and DS contributed to model input preparation. JL and FMB took the lead in writing the manuscript with input from all authors.

Code/Data availability

All data sets and model results corresponding to this study are available upon request from the corresponding author.

Competing interests

The authors have no competing interests to declare.

Acknowledgements

This work is supported by the Ph.D. program funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02 together with Harris Corporation. We would like to thank Harris Corporation and the management of Atmospheric and Environmental Research, Inc. for their support of these ongoing analyses. We acknowledge the support of Francois Ravetta (LATMOS/IPSL) and Sorbonne University, for the installation of the GreenLITE™ system on the Jussieu Campus. Thanks also to Paris Habitat, Elogie and Montparnasse ICADE for providing the locations to install the transceivers and reflectors. Finally, thanks to Marc Jamous at Cité des Sciences et de l'Industrie (CDS), to Cristelle Cailteau-Fischbach (LATMOS/IPSL), to OVSQ, and to LSCE/RAMCES technical staff for the maintenance of the in-situ monitoring network, coordinated by Delphine Combaz.

References

Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neininger, B., Dolman, A. J., and Sarrat, C.: Mesoscale covariance of transport and CO_2 fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model. Journal of Geophysical Research: Atmospheres, 112(D22), 2007.

Ahmadov, R., Gerbig, C., Kretschmer, R., Körner, S., Rödenbeck, C., Bousquet, P., and Ramonet, M.: Comparing high resolution WRF-VPRM simulations and two global CO_2 transport models with coastal tower measurements of CO_2. Biogeosciences, 6(5): 807-817, 2009.

AIRPARIF: Bilan des émissions de polluants atmosphériques et de gaz à effet de serre en Île-de-France pour l'année 2010 et historique 2000/2005, available at: https://www.airparif.asso.fr/_pdf/publications/Emissions_2010_CG75.pdf (last access: 20 April 2019), 2013.

Arzoumanian, E., Vogel, F. R., Bastos, A., Gaynullin, B., Laurent, O., Ramonet, M., and Ciais, P.: Characterization of a commercial lower-cost medium-precision non-dispersive infrared sensor for atmospheric CO_2 monitoring in urban areas. Atmospheric Measurement Techniques, 12(5), 2665-2677, 2019.

Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S., and Simmons, A.: The ERA-Interim archive, version 2.0, 2011.

Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramonet, M., Paris, J. D. and Ciais, P.: Analysis of the potential of near-ground measurements of CO_2 and CH_4 in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model. Atmospheric Chemistry and Physics, 16(11): 6735-6756, 2016.

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 CO_2 emissions from atmospheric concentration measurements. Atmospheric Chemistry and Physics, 15(4): 1707-1724, 2015.

Buchwitz, M.: Copernicus Sentinel Kandidat CO_2 Monitoring (CO2M) Mission, oral (and poster) presentation given at Nationales Forum fuer Fernerkundung und Copernicus 2018 "Copernicus gestaltet", Bundesministerium fuer Verkehr und digitale Infrastruktur, 27-29.11.2018, Berlin, 2018.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. International Journal of Climatology, 31(2): 273-288, 2011.

Chen, F., and Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review, 129(4): 569-585, 2001.

Chevallier, F., Description of the CO_2 inversion production chain. CAMS deliverable CAMS73_2015SC2_D73.1.5.5_201703_CO2 inversion production chain_v1. http://atmosphere.copernicus.eu/, 2017a.

Chevallier, F., Validation report for the inverted CO_2 fluxes, v16r1. CAMS deliverable CAMS73_2015SC2_D73.1.4.2-1979-2016-v1_201707. http://atmosphere.copernicus.eu/, 2017b.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: a summary of the AER codes, Short Communication, Journal of Quantitative Spectroscopy and Radiative Transfer, 91, 233-244, 2005.

Crisp, D., Miller, C. E., and DeCola, P. L.: NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. Journal of Applied Remote Sensing, 2(1), 23508, https://doi.org/10.1117/1.2898457, 2008.

Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2). In Earth Observing Systems XX. International Society for Optics and Photonics, 9607, 960702, 2015.

Davis, K. J., Deng, A., Lauvaux, T., Miles, N. L., Richardson, S. J., Sarmiento, D. P., Gurney, K. R., Hardesty, R. M., Bonin, T. A., Brewer, W. A., Lamb, B. K., Shepson, P. B., Harvey, R. M., Cambaliza, M. O., Sweeney, C., Turnbull, J. C., Whetstone, J., and Karion, A.: The Indianapolis Flux Experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements. Elem Sci Anth, 5:21, 2017.

Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., Wu, K., Sarmiento, D. P., Hardesty, R. M., Bonin, T. A., Brewer, W. A., and Gurney, K. R.: Toward reduced transport errors in a high resolution urban CO_2 inversion system. Elem Sci Anth, 5:20, 2017.

Dobler, J. T., Harrison, F. W., Browell, E. V., Lin, B., McGregor, D., Kooi, S., Choi, Y., and Ismail, S.: Atmospheric CO_2 column measurements with an airborne intensity-modulated continuous wave 1.57 μm fiber laser lidar. Applied Optics, 52(12): 2874-2892, 2013.

Dobler, J. T., Zaccheo, T. S., Pernini, T. G., Blume, N., Broquet, G., Vogel, F., Ramonet, M., Braun, M., Staufer, J., Ciais, P., and Botos, C.: Demonstration of spatial greenhouse gas mapping using laser absorption spectrometers on local scales. Journal of Applied Remote Sensing, 11(1): 014002, 2017.

Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. Journal of the Atmospheric Sciences, 46(20): 3077-3107, 1989.

Duren, R. M., and Miller, C. E.: Measuring the carbon emissions of megacities. Nature Climate Change, 2(8): 560-562, 2012.

Edner, H., Ragnarson, P., Spännare, S., and Svanberg, S.: Differential optical absorption spectroscopy (DOAS) system for urban atmospheric pollution monitoring. Applied optics, 32(3), 327-333, 1993.

Eldering, A., Taylor, T. E., O'Dell, C. W., and Pavlick, R: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data. Atmospheric Measurement Techniques, 12(4), 2341-2370, 2019.

Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A., Díaz-Isaac, L. I., Duren, R. M., Fischer, M. L., Gerbig, C., Gurney, K. R., Huang, J., Jeong, S., Li, Z., Miller, C. E., O'Keeffe, D., Patarasuk, R., Sander, S. P., Song, Y., Wong, K. W., and Yung, Y. L.: Los Angeles megacity: a high-resolution land-atmosphere modelling system for urban CO_2 emissions. Atmospheric Chemistry and Physics, 16(14): 9019-9045, 2016.

Font, A., Grimmond, C. S. B., Kotthaus, S., Morgu í J. A., Stockdale, C., O'Connor, E., Priestman, M., and Barratt, B.: Daytime CO_2 urban surface fluxes from airborne measurements, eddy-covariance observations and emissions inventory in Greater London. Environmental pollution, 196, 98-106, 2015.

Gates, W. L.: Static stability measures in the atmosphere. Journal of Meteorology, 18(4): 526-533, 1961.

Graven, H. D., Stephens, B. B., Guilderson, T. P., Campos, T. L., Schimel, D. S., Campbell, J. E., and Keeling, R. F.: Vertical profiles of biospheric and fossil fuel-derived CO_2 and fossil fuel CO_2: CO ratios from airborne measurements of Δ14C, CO_2 and CO above Colorado, USA. Tellus B: Chemical and Physical Meteorology, 61(3): 536-546, 2009.

Grell, G. A., and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophysical Research Letters, 29(14): 38-1-38-4, 2002.

Hamazaki, T., Kaneko, Y., and Kuze, A.: Carbon dioxide monitoring from the GOSAT satellite. In Proceedings XXth ISPRS conference, Istanbul, Turkey, 1223, 2004.

Hong, S. Y., and Lim, J. O. J.: The WRF single-moment 6-class microphysics scheme (WSM6). Journal of the Korean Meteorological Society, 42(2): 129-151, 2006.

Hu, X. M., Nielsen-Gammon, J. W., and Zhang, F.: Evaluation of three planetary boundary layer schemes in the WRF model. Journal of Applied Meteorology and Climatology, 49(9), 1831-1844, 2010.

IEA: World energy outlook, International Energy Agency (IEA), Paris, 2008.

Janjić, Z. I.: The step-mountain coordinate: Physical package. Monthly Weather Review, 118(7), 1429-1443, 1990.

Janjić, Z. I.: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review, 122(5): 927-945, 1994.

Janjić, Z.I.: The surface layer in the NCEP Eta Model. Eleventh Conference on Numerical Weather Prediction, Norfolk, VA, 19-23 August 1996. American Meteorological Society, Boston, MA, 354-355, 1996.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment, 101, 534-553, 2006.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keeffe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). Journal of Geophysical Research: Atmospheres, 121(10): 5213-5236, 2016.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth System Science Data, 10, 2141-2194, 2018.

Lian, J., Wu, L., Bréon, F. M., Broquet, G., Vautard, R., Zaccheo, T. S., Dobler, J., and Ciais, P.: Evaluation of the WRF-UCM mesoscale model and ECMWF global operational forecasts over the Paris region in the prospect of tracer atmospheric transport modeling. Elem Sci Anth, 6:64, 2018.

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 CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM). Global Biogeochemical Cycles, 22(2), 2008.

Marquis, M., and Tans, P.: Carbon crucible. Science, 320(5875): 460-461, 2008.

Martilli, A., Clappier, A., and Rotach, M. W.: An urban surface exchange parameterisation for mesoscale models. Boundary-layer meteorology, 104(2), 261-304, 2002.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. Journal of Geophysical Research: Atmospheres, 102(D14): 16663-16682, 1997.

Moore, B., Crowell, S., Rayner, P., Kumer, J., O'Dell, C., O'Brien, D., Utembe, S., Polonsky, I., Schimel, D., and Lemen, J.: The potential of the Geostationary Carbon Cycle Observatory (GeoCarb) to provide multi-scale constraints on the carbon cycle in the Americas. Frontiers in Environmental Science, 6, 109, 2018.

Mueller, K., Yadav, V., Lopez‐Coto, I., Karion, A., Gourdji, S., Martin, C., and Whetstone, J.: Siting background towers to characterize incoming air for urban greenhouse gas estimation: a case study in the Washington, DC/Baltimore area. Journal of Geophysical Research: Atmospheres, 123(5): 2910-2926, 2018.

Nehrkorn, T., Henderson, J., Leidner, M., Mountain, M., Eluszkiewicz, J., McKain, K., and Wofsy, S.: WRF simulations of the urban circulation in the Salt Lake City area for CO_2 modeling. Journal of Applied Meteorology and Climatology, 52(2): 323-340, 2013.

Sargent, M., Barrera, Y., Nehrkorn, T., Hutyra, L. R., Gately, C. K., Jones, T., McKain, K., Sweeney, C., Hegarty, J., Hardiman, B., Wang, J. A., and Wofsy, S. C.: Anthropogenic and biogenic CO_2 fluxes in the Boston urban region. Proceedings of the National Academy of Sciences, 115(29), 7491-7496, 2018.

Salamanca, F., Krpo, A., Martilli, A., and Clappier, A.: A new building energy model coupled with an urban canopy parameterization for urban climate simulations-part I. formulation, verification, and sensitivity analysis of the model. Theoretical and Applied Climatology, 99(3-4), 331, 2010.

Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion. Atmospheric Chemistry and Physics, 16(22): 14703-14726, 2016.

Tans, P., Zhao, C., and Kitzis, D.: The WMO Mole Fraction Scales for CO_2 and other greenhouse gases, and uncertainty of the atmospheric measurements, Report of the 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases, and Related Measurement Techniques, 7-10 September 2009, GAW Report No. 194, WMO TD No. 1553, 152-159, 2011.

Wang, H., Zhang, R., Liu, M., and Bi, J.: The carbon emissions of Chinese cities. Atmospheric Chemistry and Physics, 12(14), 6197-6206, 2012.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO_2 emissions? Atmospheric Chemistry and Physics, 16(12): 7743-7771, 2016.

Xueref-Remy, I., Dieudonné, E., Vuillemin, C., Lopez, M., Lac, C., Schmidt, M., Delmotte, M., Chevallier, F., Ravetta, F., Perrussel, O., Ciais, P., Bréon, F.-M., Broquet, G., Ramonet, M., Spain, T. G., and Ampe, C.: Diurnal, synoptic and seasonal variability of atmospheric CO_2 in the Paris megacity area. Atmospheric Chemistry and Physics, 18, 3335-3362, https://doi.org/10.5194/acp-18-3335-2018, 2018.

Zaccheo, T. S., Blume, N., Pernini, T., Dobler, J.T., and Lian, J.: Bias correction of long-path CO_2 observations in a complex urban environment for carbon cycle model inter-comparison and data assimilation. Atmospheric Measurement Techniques Discussions, https://doi.org/10.5194/amt-2019-199, in review, 2019.

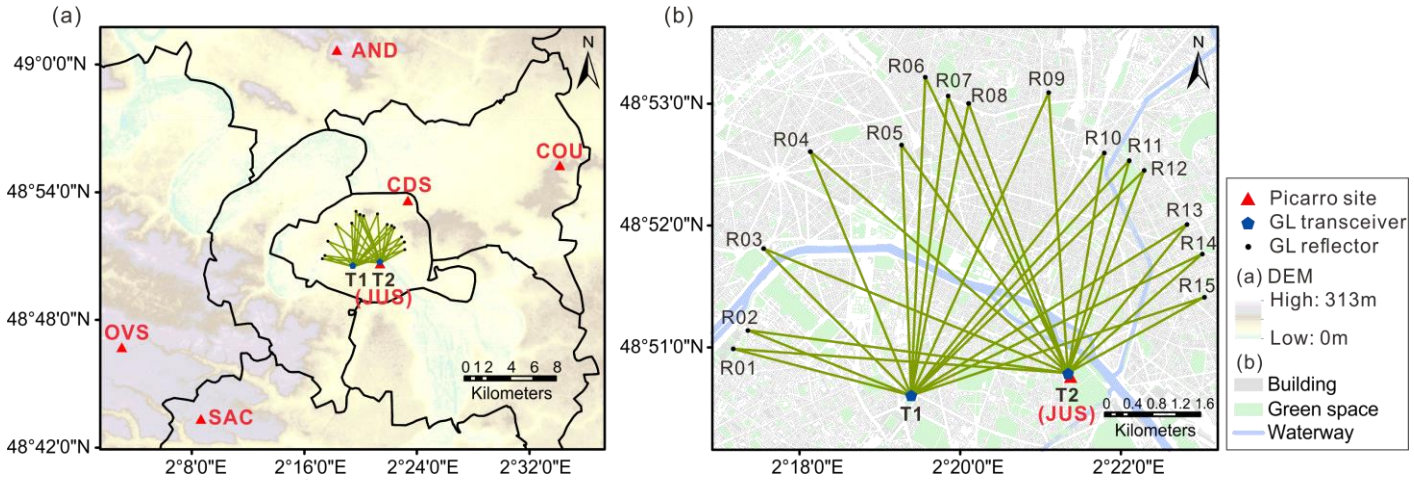

Figure 1: (a) Distribution of in-situ CO_2 measurements and the GreenLITE™ laser system. The Paris city is located within the inner line, but the urban area extends over a larger surface, very roughly within the Greater Paris area (including Paris and the three administrative areas that are around Paris called "Petite Couronne" in French, see Figure S5). The Île-de-France region covers an area that is larger than the domain shown here. (b) The GreenLITE™ laser system layout and its chord labels. (Data sources: the ASTER Global Digital Elevation Model (GDEM) Version 2 data are available at https://lpdaac.usgs.gov/products/astgtmv002/; the administrative division map of the Île-de-France region is available at https://www.data.gouv.fr/en/datasets/geofla-departements-idf/, same for Figure 2, 4, S5; the building, green space and waterway information are from OpenStreetMap available at http://download.geofabrik.de/europe/france/ile-de-france-190907-free.shp.zip)

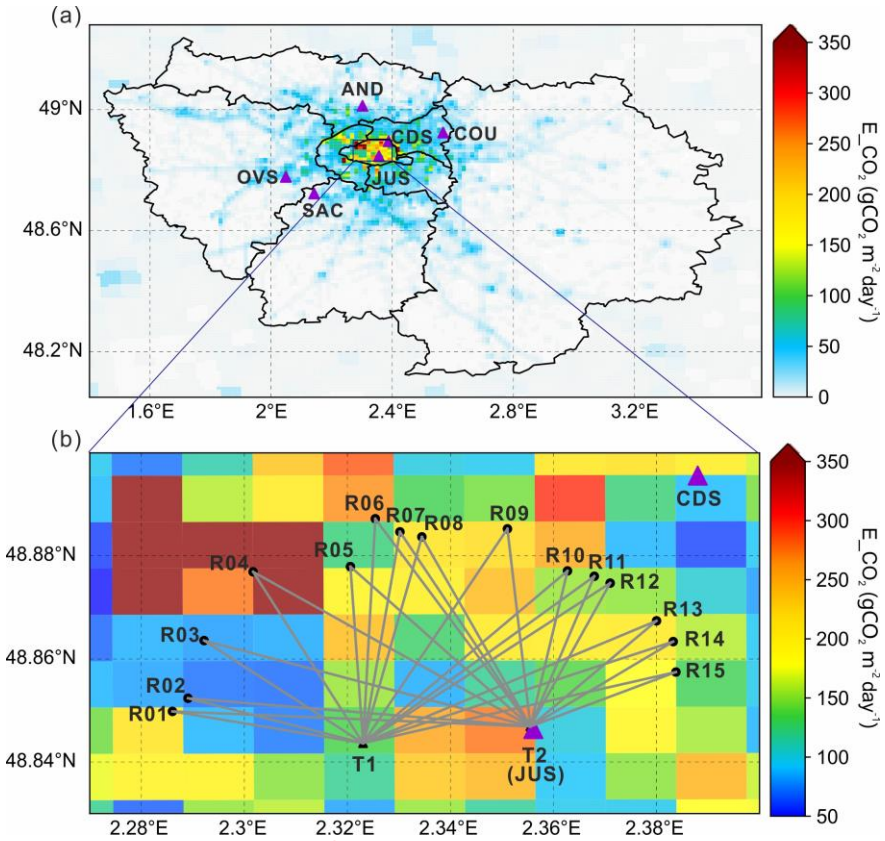

Figure 2: Total CO_2 emissions, according to the AirParif inventory (within IdF) and the IER inventory (outside IdF), for a weekday in March 2016. (a) the top panel shows the CO_2 emissions over the IdF region together with the in-situ measurement stations. (b) the bottom panel is a high-resolution zoom of the inner Paris area and shows the 1-km emissions together with the GreenLITE™ chords and two urban in-situ measurement stations.

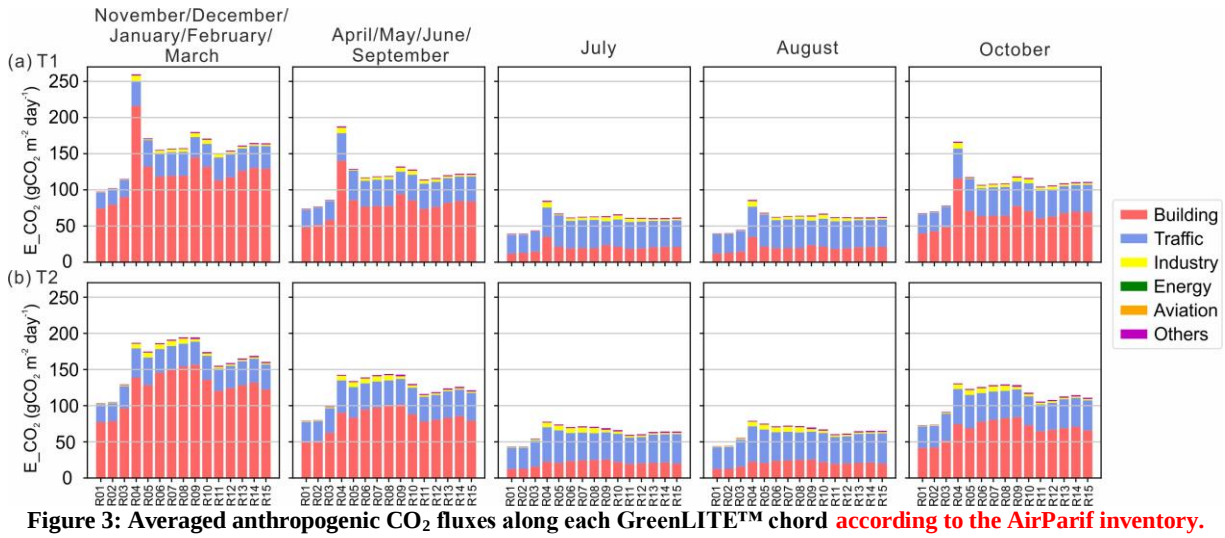

Figure 3: Averaged anthropogenic CO_2 fluxes along each GreenLITE™ chord according to the AirParif inventory.

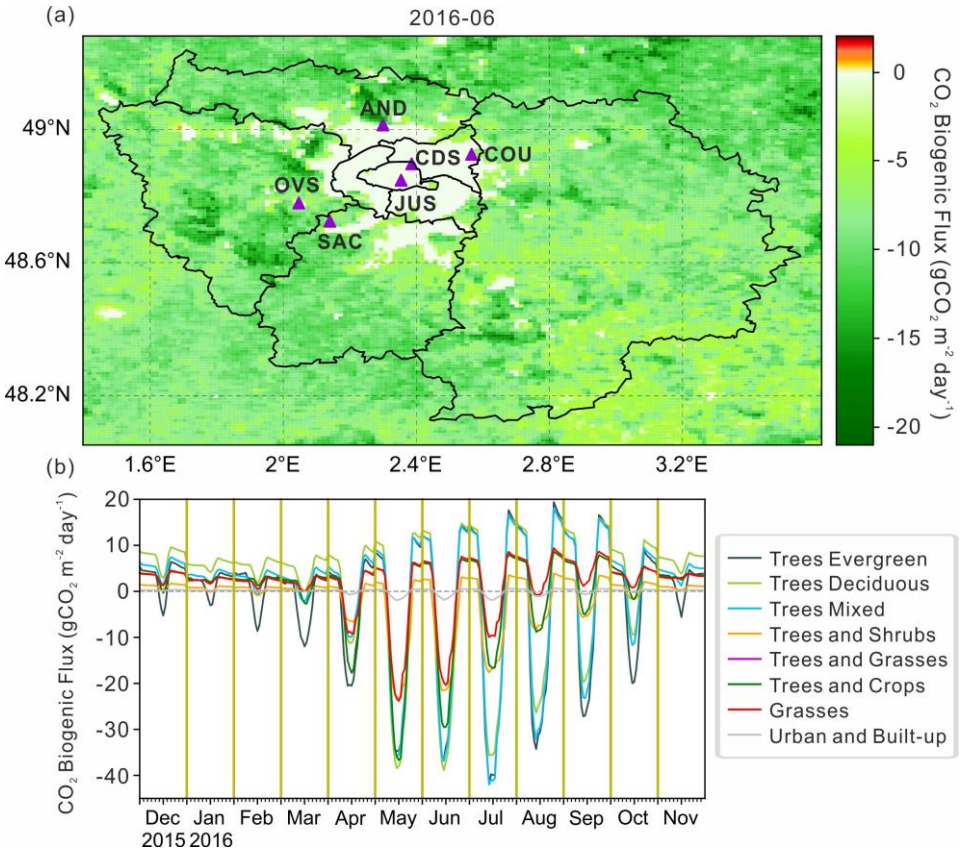

Figure 4: (a) Daytime (06-18 UTC) average of CO_2 biogenic flux (NEE) in June 2016; (b) Mean diurnal cycles of CO_2 biogenic flux (NEE) for 12 calendar months and for 8 vegetation classes used in VPRM over Domain 03.

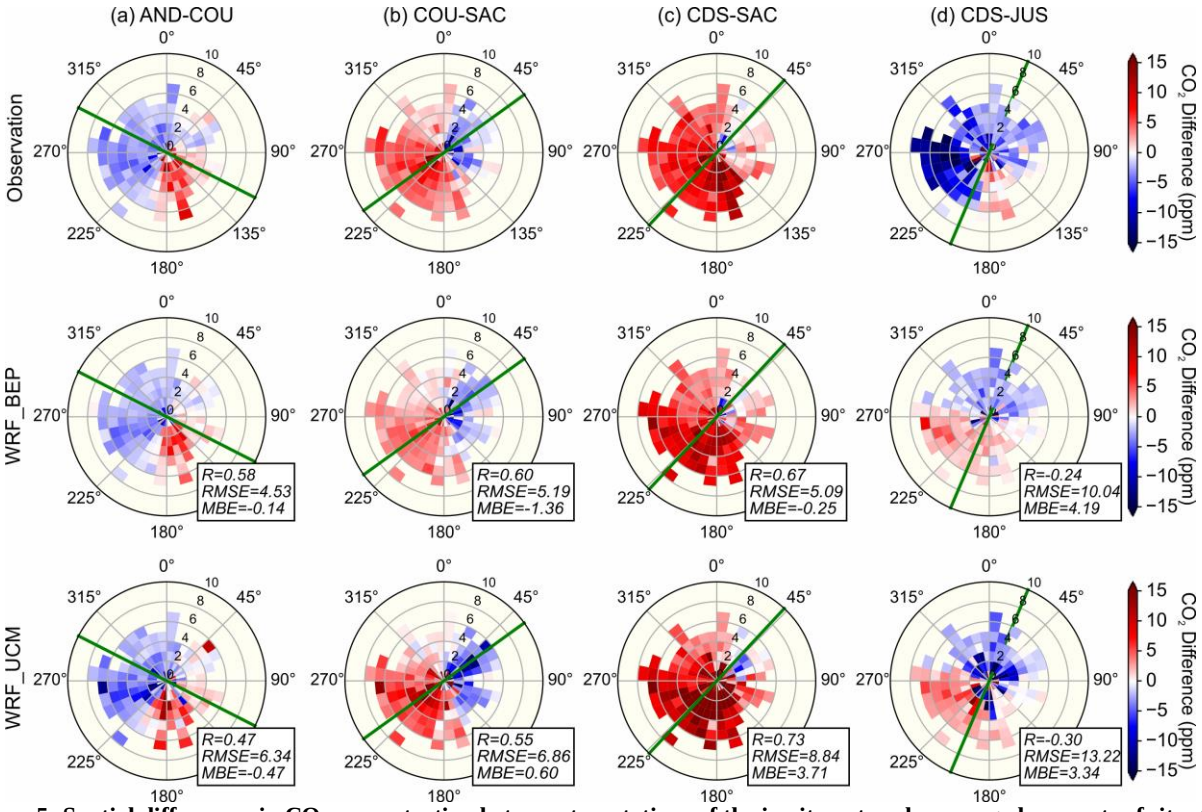

Figure 5: Spatial differences in CO_2 concentration between two stations of the in-situ network, averaged over sets of situation corresponding to bins of wind speed and direction. Only the afternoon (11-16UTC) data are used. The top row shows the observations, whereas the other two rows show the two simulations (UCM, BEP). The green line indicates the direction defined by two in-situ stations. The statistics of hourly values of observed and modeled CO_2 concentration difference are shown in the box.

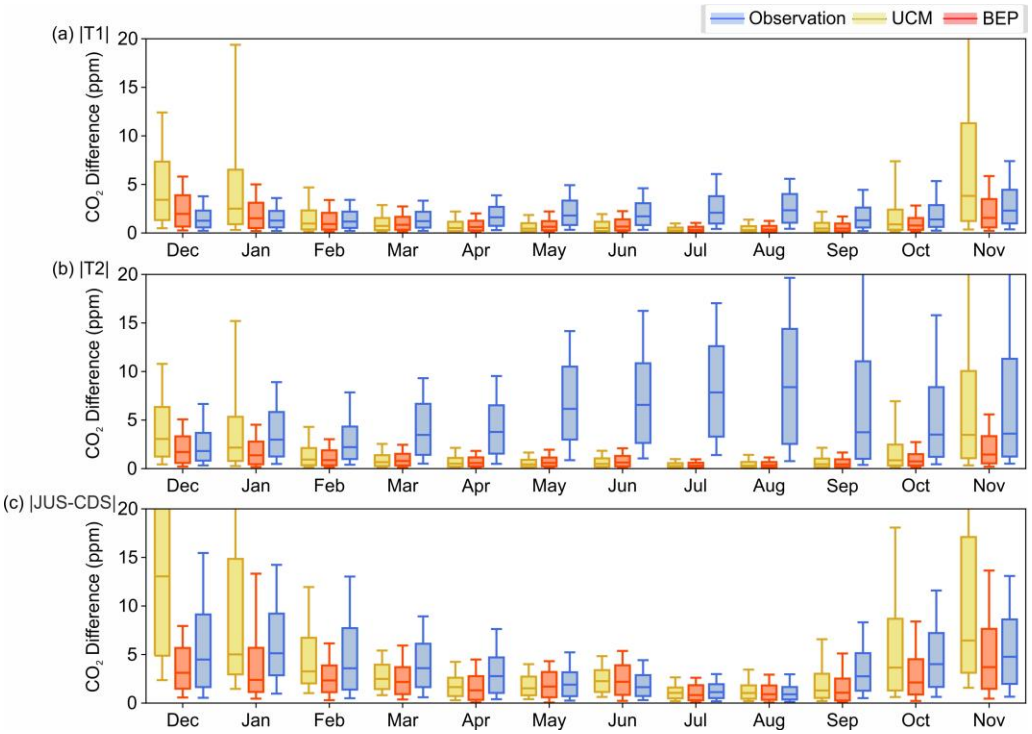

Figure 6: Distribution of the GreenLITE™ observed and modeled absolute CO_2 concentration differences between all pairs of chords for (a) T1 and (b) T2 from December 2015 to November 2016. (c) Distribution of the observed and modeled absolute CO_2 concentration differences between JUS and CDS from December 2015 to December 2018. The midpoint, the box and the whiskers represent the 0.5 quantile, 0.25/0.75 quantiles, and 0.1/0.9 quantiles respectively. Note that only the afternoon data (11-16 UTC) are used in the analysis.

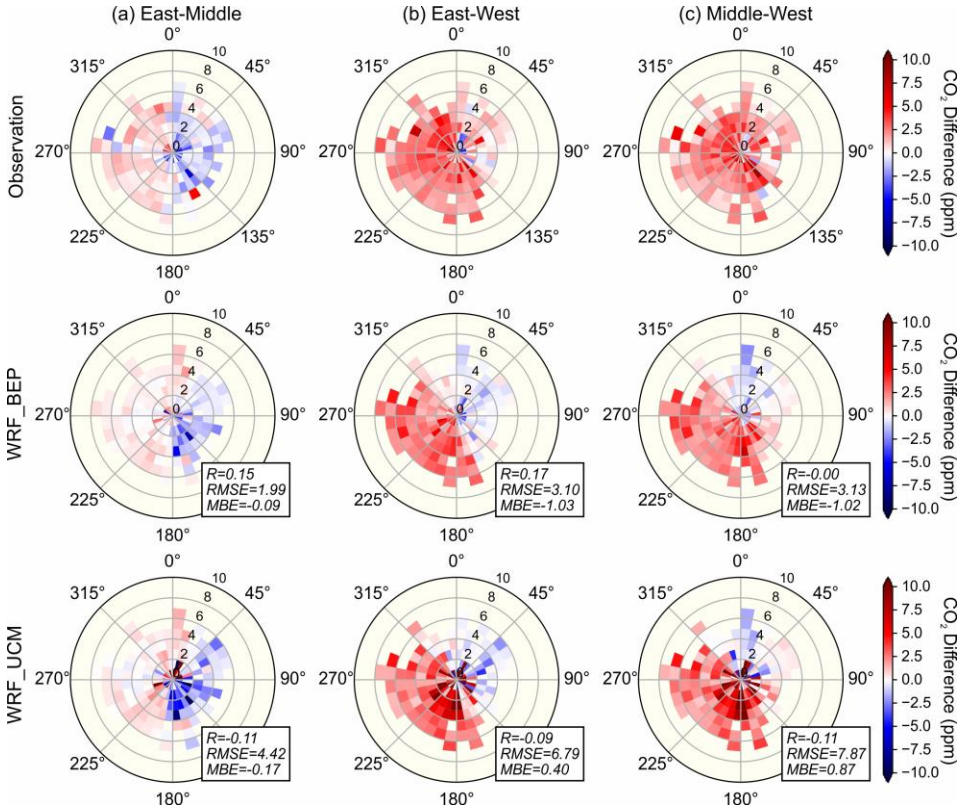

Figure 7: Spatial differences in CO_2 concentration between (a) east-middle, (b) east-west and (c) middle-west parts of the GreenLITE™ T1 measurement, averaged accounting for wind speed and direction. Only the afternoon (11-16UTC) data are used. The top row shows the observations, whereas the other two rows show the two simulations (UCM, BEP). The statistics of hourly values of observed and modeled CO_2 concentration difference are shown in the box.

Table 1. Information about CO_2 observation stations used in this study.

(a) In-situ stations

Site	Acronym	Latitude (°)	Longitude (°)	Height AGL (m)
Jussieu	JUS	48.8464	2.3561	30
Cité des Sciences	CDS	48.8956	2.3880	34
Andilly	AND	49.0126	2.3018	60
Coubron	COU	48.9242	2.5680	30
OVSQ	OVS	48.7779	2.0486	20
Saclay	SAC	48.7227	2.1423	15 and 100

(b) The GreenLITE™ system

		R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	R14	R15
Chord Length (km)	T1	2.80	2.67	3.17	4.02	3.81	4.84	4.59	4.53	5.06	4.72	4.88	4.93	4.94	4.93	4.71
	T2	5.11	4.91	5.00	5.17	4.30	5.00	4.59	4.38	4.28	3.40	3.37	3.30	2.90	2.74	2.39
Height AGL (m)	R	50.4	41.7	18.3	28.1	19.7	20.8	24.5	25.9	16.9	28.8	29.7	24.7	21.8	16.8	23.6
	T	T1: 50.3; T2: 86.8														

Table 2. A summary of WRF-Chem configurations used in this study.

Option		Setting
Simulation Periods		2015.09.01~2016.11.30
Horizontal Resolution		25 km (Domain 01), 5 km (Domain 02), 1 km (Domain 03)
Boundary & Initial Conditions	Meteorology	ERA-Interim reanalysis data (0.75 °×0.75 °, 6 hourly)
	CO_2 concentration	LMDZ_CAMS (3.75 °×1.895 °, 3 hourly)
Nudging		Grid nudging + Surface nudging + Observation nudging (NCEP operational global observation surface data (ds461.0) and upper-air data (ds351.0))
Flux	Anthropogenic emissions	IER inventory for 2005 (5 km, outside IdF) + AirParif inventory for 2010 (1 km, within IdF) rescaled for 2015-2016 using national budgets from Le Quéré et al. (2018)
	Biogenic NEE	VPRM (online coupling)
Physics Schemes	Microphysics	WSM6 scheme
	Cumulus convection	Grell 3D ensemble scheme only in Domain 01
	Longwave radiation	RRTM scheme
	Shortwave radiation	Dudhia scheme
	PBL	MYJ scheme
	Surface layer	Eta Similarity scheme
	Vegetated land surface	Unified Noah land-surface model
	Urban land surface	UCM (34 vertical levels wherein 15 below 1.5 km)
		BEP (44 vertical levels wherein 25 below 1.5 km)

Table 3. Seasonal statistics for observed and modeled hourly afternoon CO_2 concentrations for two urban canopy schemes (UCM, BEP) from December 2015 to November 2016. DJF denotes December-January-February, MAM denotes March-April-May, JJA denotes June-July-August and SON denotes September-October-November. The color highlights the value in the cell with the minimum in blue, the median in white and the maximum in red. All other cells are colored proportionally.

(a) Correlation coefficient

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
Hourly afternoon (11-16 UTC)	DJF	0.79	0.83	0.70	0.79	0.68	0.65	0.65	0.59	0.65	0.86	0.65	0.86
	MAM	0.67	0.81	0.69	0.79	0.51	0.60	0.71	0.78	0.77	0.81	0.81	0.82
	JJA	0.46	0.47	0.45	0.46	NA	NA	0.52	0.55	0.57	0.63	0.49	0.49
	SON	0.73	0.83	0.71	0.82	0.55	0.73	0.65	0.75	0.77	0.83	0.74	0.82

(b) Root-mean-square error (RMSE. Unit: ppm)

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
Hourly afternoon (11-16 UTC)	DJF	31.82	5.98	23.79	6.68	42.31	10.08	33.75	9.61	8.14	5.33	7.08	4.92
	MAM	7.84	4.47	6.69	5.12	9.17	6.11	7.27	4.79	5.75	4.55	5.11	4.47
	JJA	7.07	5.99	7.51	7.25	NA	NA	7.26	5.46	5.86	4.06	5.04	4.56
	SON	31.87	9.57	28.39	10.45	42.50	13.09	32.29	12.01	9.72	6.50	8.20	6.46

(c) Mean bias error (MBE. Unit: ppm)

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
Hourly afternoon (11-16 UTC)	DJF	17.37	0.99	12.99	-0.90	13.55	-5.24	19.61	2.69	3.51	1.74	0.59	0.21
	MAM	2.59	0.59	-0.72	-2.71	0.58	-2.36	2.91	0.52	3.22	1.59	2.08	0.46
	JJA	0.66	-0.89	-2.65	-4.09	NA	NA	1.85	0.06	3.14	1.62	1.13	0.17
	SON	14.01	-0.86	8.65	-4.36	12.84	-4.47	11.29	-0.92	4.88	1.14	2.60	0.02

Supplement

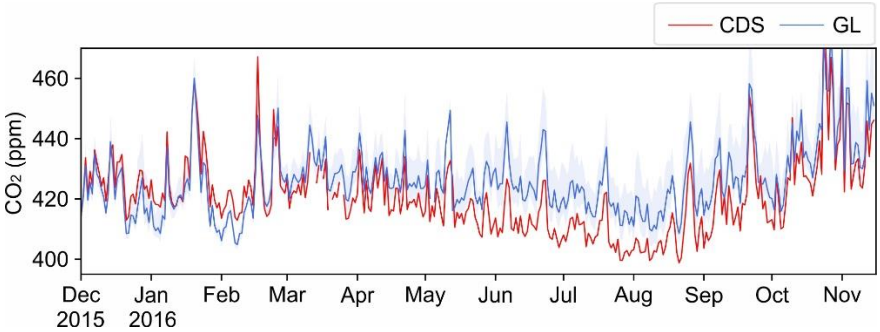

Figure S1: Daily average CO_2 concentrations between the original GreenLITE™ data and the CDS measurement. The shaded area indicates the 30 inter-chord range.

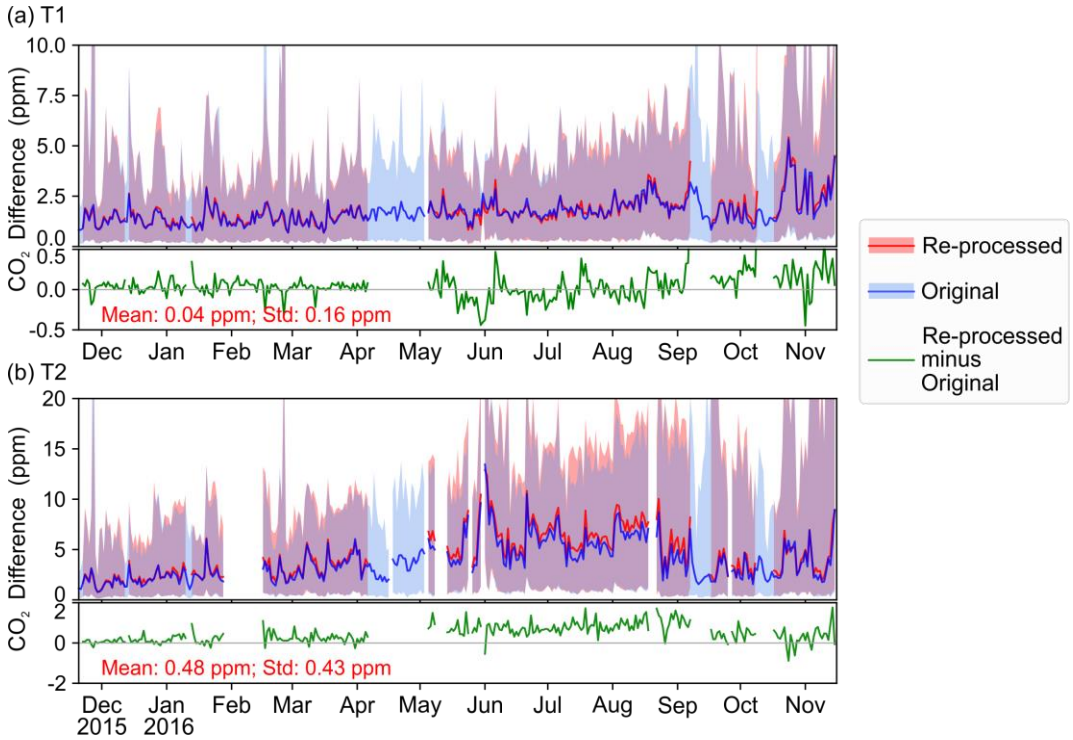

Figure S2: Distribution of the original and re-processed GreenLITE™ absolute CO_2 concentration differences between all pairs of chords for (a) T1 and (b) T2 from December 2015 to November 2016. The solid lines in top panels of (a) and (b) indicate the 0.5 quantile, and the shaded areas represent the 0.1 and 0.9 quantile intervals for original data in blue and re-processed data in red. The green line in bottom panels of (a) and (b) indicates the differences between the median values of the re-processed and original inter-chord range.

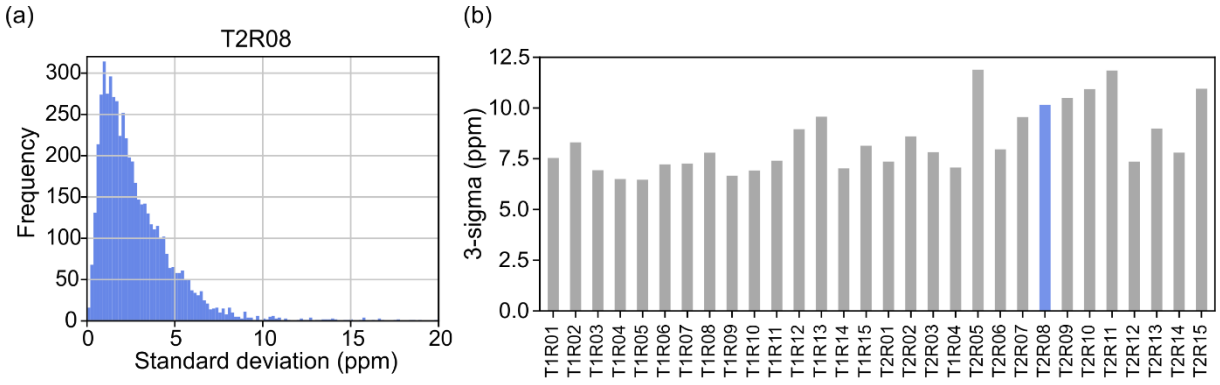

Figure S3: (a) Frequency distribution of the standard deviations of 4-minute CO_2 concentrations measured within one hour for one chord (e.g. T2R08); (b) Three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord.

The outlier detection for the 4-minute GreenLITE™ data is mainly based on the 3-sigma rule, which is used to remove the data outside three standard deviations from a mean in the positive direction. Figure S3 (a) shows the frequency distribution of the standard deviations of the 4-minute CO_2 concentrations measured within one hour for one given chord (e.g. T2R08). Figure S3 (b) shows the three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord. In general, the threshold varies between 6.5 ppm and 11.9 ppm from chord to chord. We therefore choose to use a uniform threshold value of 10 ppm to remove the outliers for all chords.

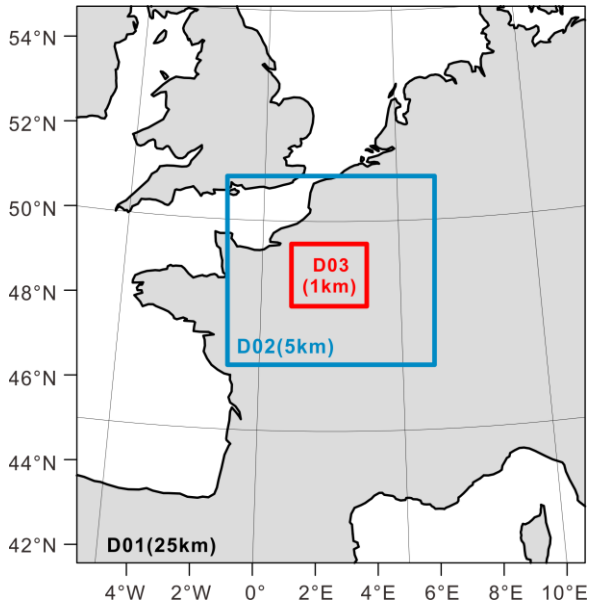

Figure S4: WRF-Chem domain settings.
(Data source: the continent-ocean boundary is from the NCAR Command Language (NCL) Ncarg4_1 database)

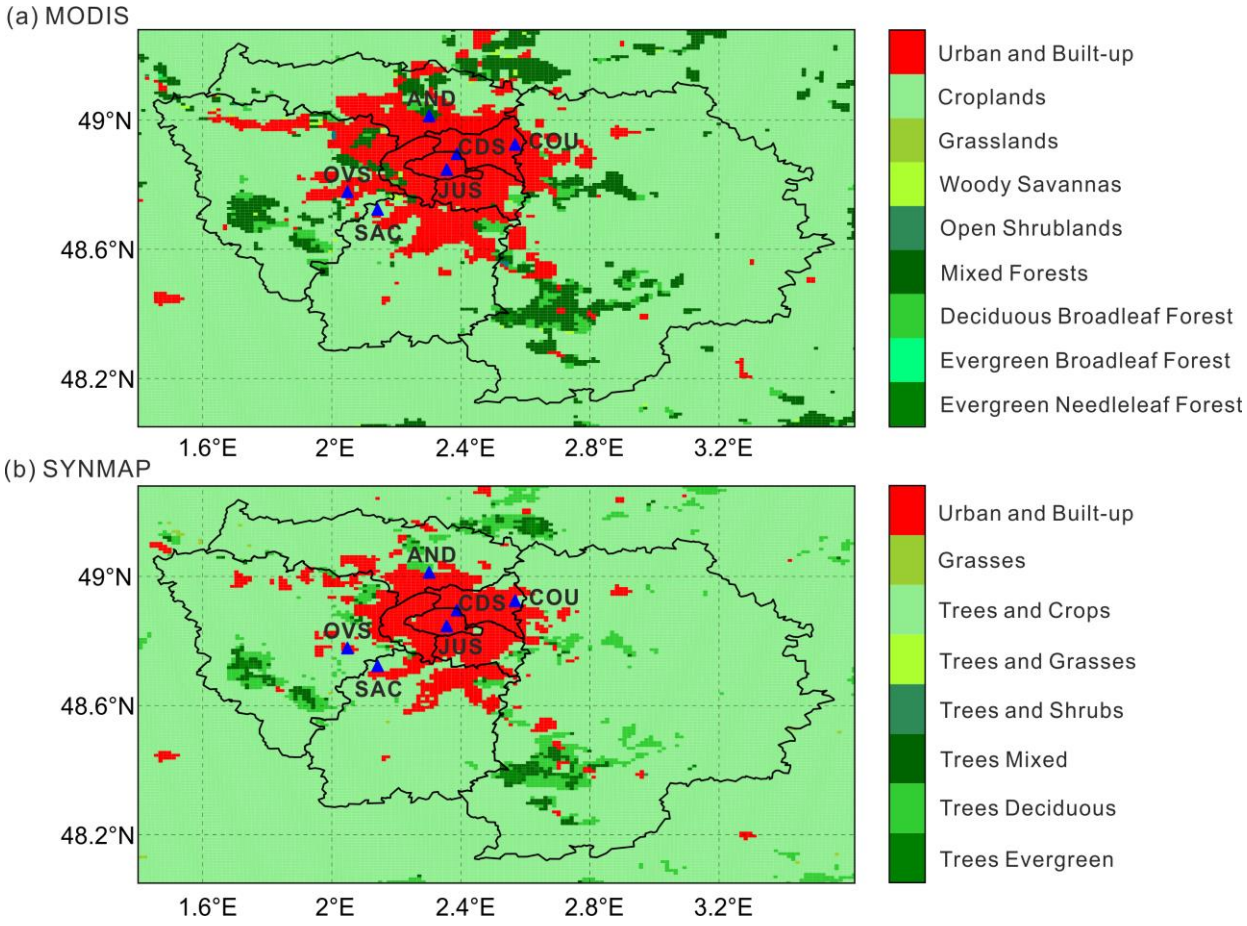

Figure S5: Dominant land use categories over the IdF region of Domain 03 from (a) MODIS database used in the WRF model and (b) SYNMAP database used in the VPRM model. Note that in the VPRM, it is the fractional vegetation coverage for a given grid cell covered by the respective land cover classes.

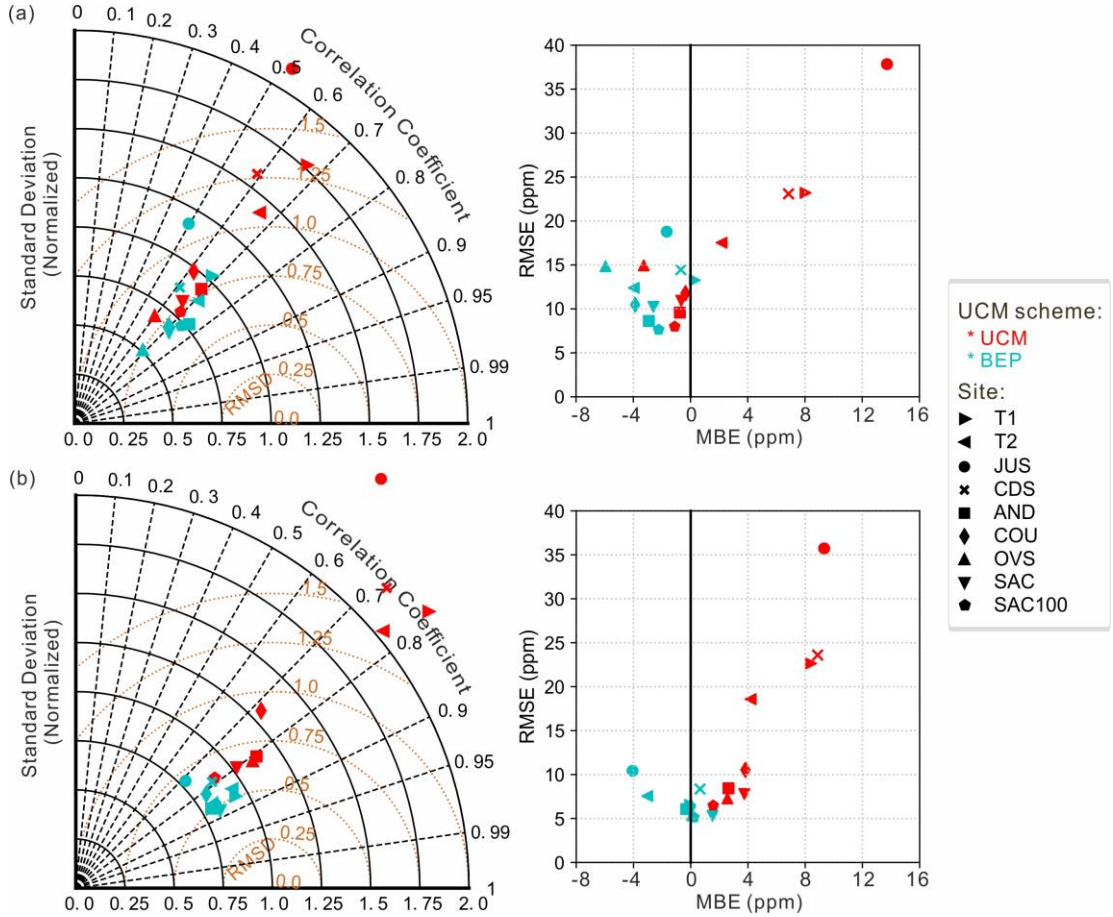

Figure S6: Normalized Taylor diagram (left) and RMSE versus MBE (right) for simulated (a) all hourly and (b) hourly afternoon (11-16 UTC) CO_2 concentrations from December 2015 to November 2016. The colors of markers indicate 2 different urban canopy schemes with UCM in red and BEP in cyan. In the Taylor diagram, normalized standard deviation is on the radial axis; Correlation coefficient is on the angular axis; Orange dashed lines indicate RMSD.

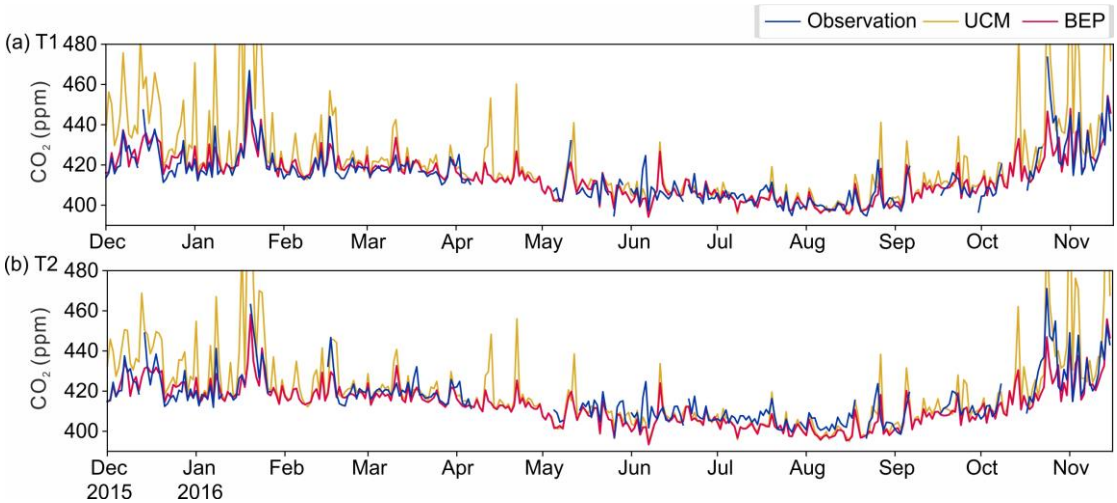

Figure S7: Time series of the GreenLITE™ observed and modeled averaged CO_2 concentrations during afternoon (11-16 UTC) for the (a) T1 and (b) T2 chord ensembles.

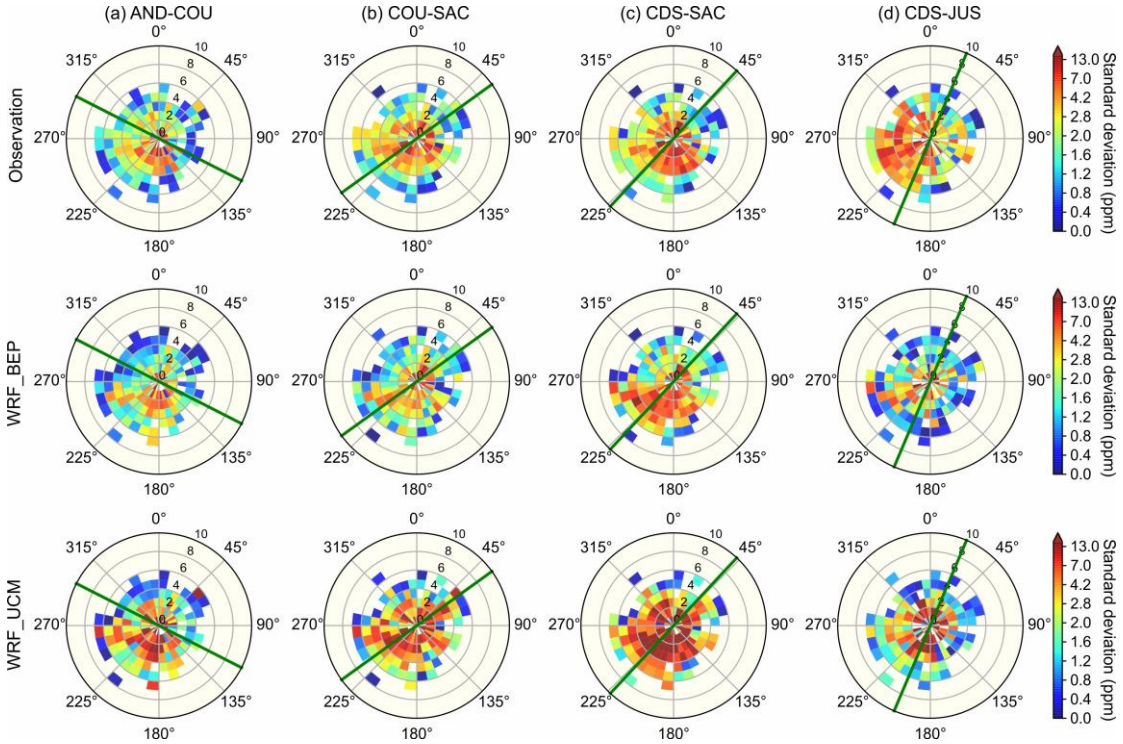

Figure S8: Standard deviations of spatial differences in CO_2 concentration between two stations of the in-situ network accounting for wind speed and direction. Only the afternoon (11-16UTC) data are used. The top row shows the observations, whereas the other two rows show the two simulations (UCM, BEP). The green line indicates the direction defined by two in-situ stations.

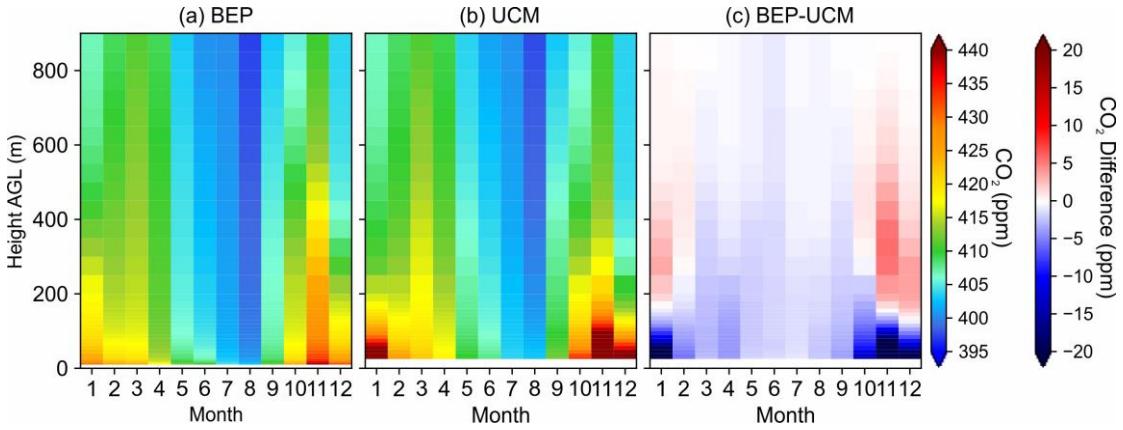

Figure S9: Vertical distributions of CO_2 concentrations during afternoon (11-16 UTC) at transceiver T2 (also JUS) for 12 calendar months for two simulations (a) BEP, (b) UCM, and (c) their differences.

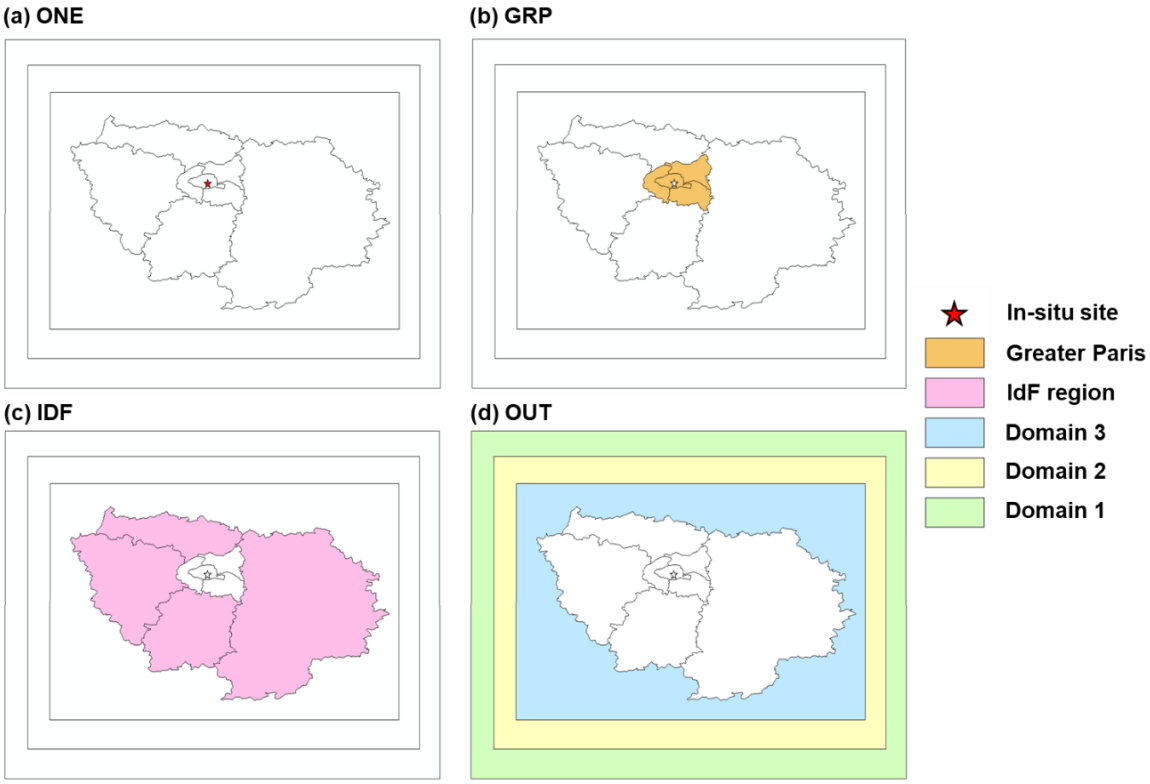

Figure S10. Four experiments are carried out for the JUS station with the assignment of emissions to: (a) ONE: one grid cell that contains an in-situ station; (b) GRP: all grid cells within the Greater Paris except the one where the station is located; (c) IDF: all grid cells within the IdF region except those of the Greater Paris; (d) OUT: all grid cells outside the IdF region. Another four experiments are carried out for the COU station.

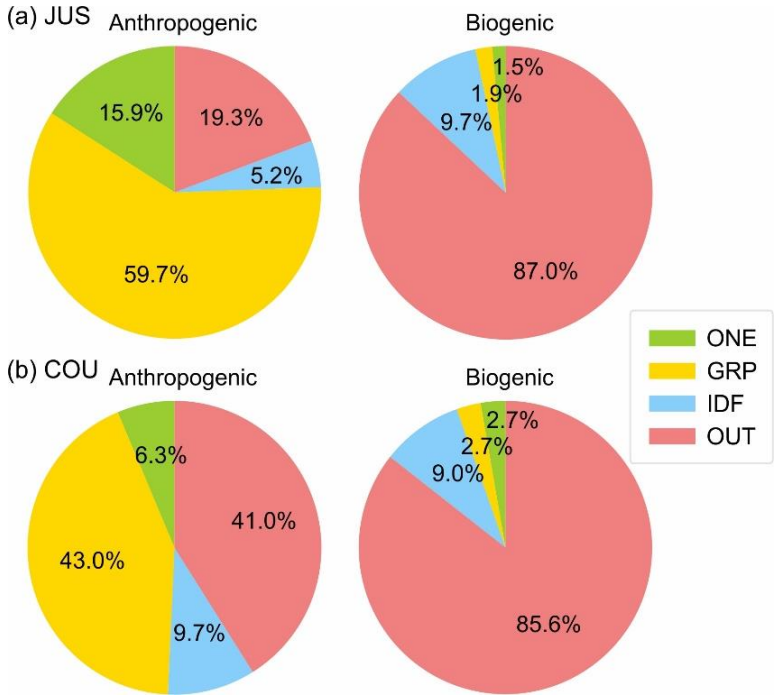

Figure S11. Relative contributions (in percentages) of each component flux to the modeled total anthropogenic and biogenic CO_2 concentrations for (a) urban site JUS and (b) suburban site COU. Note that only the afternoon data (11-16 UTC) are used in the analysis.

In order to determine respective contributions of various areas/sectors to the simulated CO_2 concentrations at a certain measurement site, we carried out a set of sensitivity experiments for the one-month period of March 2016 with anthropogenic and biogenic emissions limited to a given area within the simulation domain. This set of experiments includes the assignment of emissions to: 1) ONE: one grid cell that contains an in-situ station, 2) GRP: all grid cells within the GReater Paris except the one where the station is located, 3) IDF: all grid cells within the IdF region except those of the Greater Paris, 4) OUT: all grid cells outside the IdF region, as shown in Figure S10 (a) (b) (c) (d) respectively. The contribution from sources outside the model domain is small enough so that its influence is negligible. Figure S11 shows the relative contributions (in percentages) of each component to the modeled total anthropogenic and biogenic CO_2 concentrations for one urban site JUS and one suburban site COU respectively. The simulated monthly mean concentrations of anthropogenic CO_2 are 11.0 ppm at JUS and 5.4 ppm at COU, which are much larger than those of biogenic CO_2 (0.6 ppm at JUS and 0.7 ppm at COU). In general, an urban station like JUS is under a strong influence of the anthropogenic emissions within the IdF region. The contributions of anthropogenic emissions in the vicinity of the station (ONE) and from the Greater Paris (GRP) areas to the simulated anthropogenic CO_2 concentrations are around 16% and 60% respectively, whereas the remote anthropogenic emissions account for less than 20%. For a suburban station like COU, the Parisian emissions (GRP) and the remote ones (OUT) have a comparable influence (~40%) on the simulated anthropogenic concentrations, with very large variations depending on the wind direction (downwind or upwind of the city). Note that in these experiments, the emission inventory and the WRF-Chem modeling cannot describe the CO_2 patterns (both emission and concentration) at a scale finer than 1 km, and the simulation shows that the "local" contribution is significant. The unresolved spatial distribution of the emission can therefore be a significant contribution to the uncertainty. The biogenic CO_2 concentrations mainly come from outside of the IdF region (~86%).

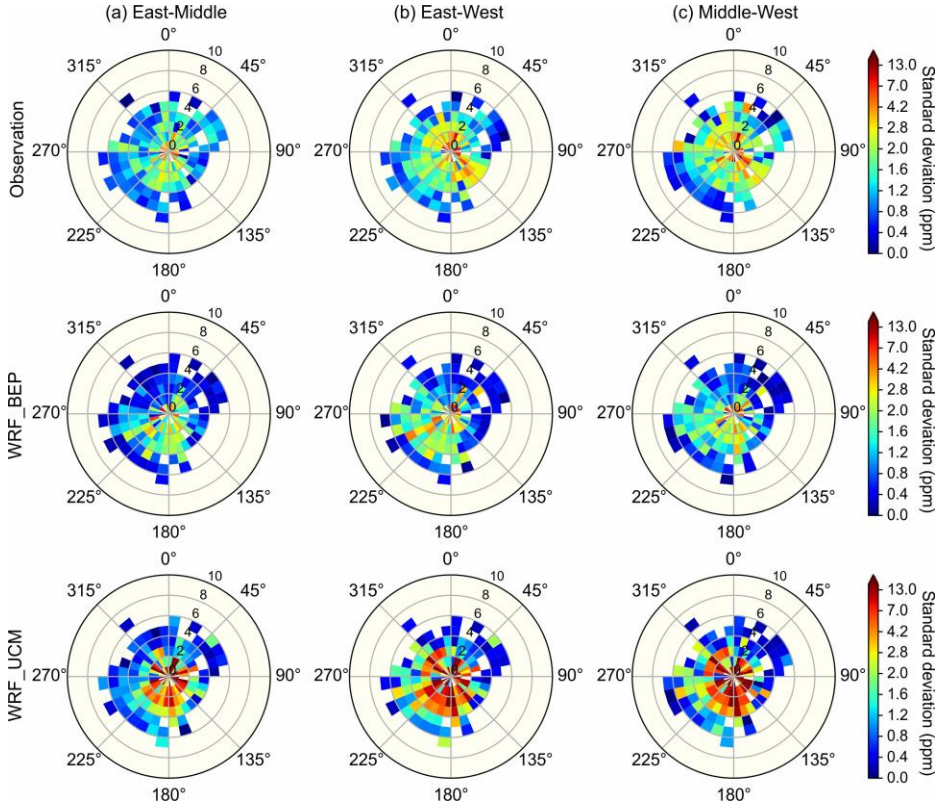

Figure S12: Standard deviations of spatial differences in CO_2 concentration between (a) east-middle, (b) east-west and (c) middle-west parts of the GreenLITE™ T1 measurement accounting for wind speed and direction. Only the afternoon (11-16UTC) data are used. The top row shows the observations, whereas the other two rows show the two simulations (UCM, BEP).

Table S1. National CO_2 emissions from fossil-fuel combustion and cement production for the countries within the WRF-Chem domain used in this study (unit: $MtCO_2$/yr). The data in the following table are taken from Le Quéré et al. (2018), available at https://www.icos-cp.eu/GCP/2018, last access: August 2019. (The use of data is conditional on citing the original data sources: data in black are from the CDIAC inventory (Boden et al., 2017), data in red are from the UNFCCC national inventory reports (UNFCCC, 2018), data in purple are from the BP Statistical Review of World Energy (BP, 2018). Cement emissions are updated from Andrews (2018))

	Austria	Belgium	France (including Monaco)	Germany	Italy (including San Marino)	Luxembourg	Netherlands	Spain	Switzerland	United Kingdom
2005	79.37	125.64	432.64	867.22	495.23	12.05	177.53	368.96	45.78	570.00
2010	72.38	113.58	397.90	833.68	424.87	11.15	182.18	283.88	45.05	512.21
2015	66.70	100.23	348.16	797.08	355.48	9.26	165.03	271.73	38.74	422.66
2016	67.40	100.24	350.10	801.75	350.32	9.00	165.52	260.99	39.20	398.55

Table S2. Seasonal statistics for observed and modeled all hourly CO_2 concentrations for two urban canopy schemes (UCM, BEP) from December 2015 to November 2016. DJF denotes December-January-February, MAM denotes March-April-May, JJA denotes June-July-August and SON denotes September-October-November. The color highlights the value in the cell with the minimum in blue, the median in white and the maximum in red. All other cells are colored proportionally.

(a) Correlation coefficient

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
All hourly	DJF	0.68	0.67	0.58	0.67	0.62	0.63	0.49	0.51	0.68	0.77	0.59	0.73
	MAM	0.52	0.56	0.52	0.61	0.39	0.39	0.48	0.53	0.64	0.71	0.64	0.67
	JJA	0.63	0.61	0.60	0.60	NA	NA	0.52	0.55	0.68	0.72	0.59	0.63
	SON	0.55	0.58	0.57	0.63	0.47	0.46	0.55	0.54	0.57	0.65	0.64	0.70

(b) Root-mean-square error (RMSE. Unit: ppm)

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
All hourly	DJF	28.26	11.23	19.38	11.03	40.96	14.43	28.84	13.63	8.82	7.42	7.47	6.64
	MAM	18.91	11.77	14.85	9.84	25.89	14.42	18.24	12.23	8.78	7.86	7.85	7.74
	JJA	9.98	10.33	10.13	10.09	NA	NA	12.11	11.00	11.48	11.49	7.14	7.20
	SON	32.94	20.06	25.23	18.11	43.50	24.22	29.57	20.27	13.82	13.20	9.46	8.97

(c) Mean bias error (MBE. Unit: ppm)

		T1		T2		JUS 30m		CDS 34m		SAC 15m		SAC 100m	
		UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP	UCM	BEP
All hourly	DJF	12.99	-0.36	6.75	-2.97	14.24	-3.85	12.09	-0.11	0.96	-0.89	-1.09	-1.62
	MAM	6.28	1.21	1.11	-3.32	8.65	-0.12	4.94	-0.62	0.03	-1.53	-1.30	-2.59
	JJA	1.25	0.97	-2.50	-3.68	NA	NA	1.77	0.74	-3.71	-4.38	-1.69	-2.72
	SON	14.06	-0.83	5.33	-6.20	17.70	-1.39	8.99	-3.05	-0.64	-3.98	-0.27	-2.01

References in the supplement:

Andrew, R. M.: Global CO_2 emissions from cement production, Earth System Science Data, 10, 195-217, https://doi.org/10.5194/essd-10-195-2018, 2018.

Boden, T. A., Marland, G., and Andres, R. J.: Global, Regional, and National Fossil-Fuel CO_2 Emissions, available at: http://cdiac.ornl.gov/trends/emis/overview_2014.html (last access: July 2017), Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA, 2017.

BP: BP Statistical Review of World Energy June 2018, available at: https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf, last access: June 2018.

UNFCCC, 2018. National Inventory Submissions 2018. United Nations Framework Convention on Climate Change. Available at: http://unfccc.int/process/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/national-inventory-submissions-2018; accessed June 2018.