Reply to comments by Jocelyn Turnbull

Review of A first year-long estimate of the Paris region fossil fuel CO2 emissions based on atmospheric inversion (Staufer et al)

This paper uses atmospheric CO2 observations in a Bayesian inversion to evaluate urban CO2 emissions for Paris. It builds on earlier work by Breon et al (2015) that first developed the inversion framework, using the innovative concept of inverting for differences between two observing sites (rather than absolute CO2 mole fractions). This work expands the dataset to a full year, which allows meaningful evaluation of how well the framework works. The inversion framework already shows some useful outcomes that can inform/improve the bottom-up inventory priors. That is, the observations imply a stronger seasonal cycle in emissions than is represented in the bottom-up data products, and this stringer seasonal cycle is also consistent with expectations. They also discuss the challenges and limitations of their inversion framework. A major challenge is that the inversion result is still strongly dependent by the prior (bottom-up) emission estimate. They show that there is still much work to be done to provide detailed information from this type of regional inversion and they provide useful suggestions as to how the inversion could be improved.

This is a very nice paper and an excellent contribution to the (still very small) urban greenhouse gas literature. It is entirely appropriate for publication in ACP with some minor revisions as noted below.

We thank Dr. J. Turnbull for this positive assessment of our study, for her comments, and for having highlighted our objective of discussing the challenges, limitations and potential improvement of the current framework as well as the first successes obtained with it. We consider that it is a critical aspect of our analysis (see answers to reviewer 2).

General comments: There is very little attention paid to the contribution of the biosphere to the urban CO2 observations and it's contribution to uncertainties in the inversion. The biosphere fluxes used as priors are described only very briefly, but there is no information about the quality of that prior and how much biases and uncertainties in it might contribute to biases/uncertainties in the inversion. Some discussion of this is needed in the paper.

We agree that the topic of the natural fluxes is critical for city scale inversions and we agree that we need to discuss it in this paper. However:

- Paris is a particular case in the sense that it is a very dense urban area with limited vegetation inside it.
- According to the model simulations, the computation of gradients succeeds in cancelling the impact of biogenic fluxes in the data that are analyzed (it was too rapidly mentioned in the last section of the paper).
- The weak impact of biogenic fluxes on the inversion of anthropogenic CO2 (according to the modelling framework) has already been addressed in more detail in Bréon et al. (2015), with which we try to avoid redundancies. In addition, the specific topic of the

> cancelling of the biogenic component in the simulated CO2 gradients has also been analyzed in the Boon et al. 2015, paper.
>
> - The new computation of the gradients proposed in this study further decreases the potential impact of the biogenic fluxes in the inversion.

Therefore, we expand the introduction by discussing the cancelling of the natural component in the CO2 gradients according to the model. We will indicate in the result section that the amplitude of the natural component of the simulated CO2 gradients is further decreased when narrowing the wind ranges for the gradient selection. Furthermore, we will expand the discussion on this topic based on such an analysis and carefully remind that actual gradients in the measurements may bear a larger signature of natural fluxes than the simulated gradients. See also the answer to the specific comment on C-TESSEL.

Reference:

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 CO2 and CH4 in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model, Atmos. Chem. Phys., 16, 6735-6756, doi:10.5194/acp-16-6735-2016, 2016.

Specific comments: Pg 2 line 14. "two-month" not "two-months"

It will be corrected.

Pg 2 line 26. Presumably air parcels pass over the city, rather than through it?

Yes. It will be corrected in the revised manuscript.

Pg 3. Lines 24-29. This paragraph is hard to follow. It transitions abruptly from a description of preliminary tests to describing what is in specific sections of the paper.

We will add 1 or 2 sentences explaining that this revision consists in narrowing the wind range for the gradient selection to avoid situations during which the air leaving the "upwind site" or reaching the "downwind site" does not overpass a significant part of the city and the vicinity of the other site.

Pg 3 section 2. It would be helpful if the inversion parameters were referred to by what they are, rather than by the algebra term in the equations. For the reader who is not a specialist in Bayesian inversions, it is hard to remember what y yo, x etc are referring to.

We will add the name of the variable before such labels throughout section 2 wherever it does not alter the reading of the text.

Pg 5 lines 1-4. How good is C-TESSEL for the urban area? See also my general comment above.

This model is definitely not perfect for modeling ecosystems within and in the vicinity of urban

areas. It has a relatively low resolution compared to that of the transport model. This should, in principle, increase the challenge of dealing with the impact of uncertainties in the natural fluxes since significant NEE is simulated on grid cells with high emissions at the edges of the urban areas. However, the fact that the signature on the simulated gradients from such fluxes is low demonstrates that these drawbacks from the C-TESSEL product do not have critical impacts on the inversion. In the revised manuscript, this will be discussed based on references to Bréon et al. 2015 and Boon et al. 2015 when we present the C-TESSEL product.

Pg 5 lines 15-18. What is the measurement quality for the 1 hour means that are used in the inversion?

Based on the regular analysis of target gases, and intercomparison of side by side measurements (at Gif sur Yvette) the accuracy of the hourly averages at the three sites is better then 0.4 ppm which makes the measurement errors for the hourly averages negligible compared to the signals and other sources of errors discussed in this study.

Pg 6 lines 3-4. There are other studies that show that ATMs do poorly at night. It would be better to reference some work other from outside your own research group.

We will reference to Geels et al., ACP, 2007 in the revised manuscript.

References:

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe – Part 1: mapping the atmospheric CO2 signals, Atmos. Chem. Phys., 7, 3461-3479, doi:10.5194/acp-7-3461-2007, 2007.

Pg 6 lines 30-32. The justification for discarding the GON-MON gradients needs to be given. Is it that the sites are too close together so that emissions are not large enough to create consistent enhancements? Or is there a major flaw in the methodology?

The explanation was given at Pg 6 lines 11-12: we definitely consider that the sites are too close from each other so that in the GON-MON gradients, we face uncertainties in the high resolution emission mapping rather than uncertainties in the city-cale budget of the emissions. Furthermore, the Roissy Charles de Gaulle airport and its very specific type of emissions (with very specific problems for simulating their injection heights) is located between the two sites. The text at this place will be slightly expanded in the revised manuscript.

Pg 6 line 26. "BR2015" should be "Breon et al 2015" I think.

Yes, it will be corrected.

Pg 7 plines 10-14. One of the gradients had far more impact than the rest, so was excluded? This seems important - what is the justification for excluding it? How different was the inversion when this gradient was included?

We provide a detailed answer to this technical topic below. However, as explained at the end of this answer, it has a negligible impact on the study.

We based our judgment on an objective computation of the "observation impact" which consists in indicating how much a given data, or more precisely a model-data misfit, drives the increment applied to the fluxes given a set of data assimilated (it is given by the product between the gain matrix K and $y0\_i - H\_i x\_b$ where y_i is the corresponding observation). Through the least square minimization of the misfits to the data, it can happen that one data point out of tens of data points is the dominating driver of the inversion results. It happens, for instance, if the corresponding model-data misfit is far larger than the other ones, or if at the corresponding time, the atmospheric transport is such that the sensitivity of this data point to the fluxes is far larger than that of the other data points (typically if the PBL is very shallow at the corresponding time). The large model-data misfit can reveal a very high observation error and giving too much weight to a single data is dangerous (we prefer to work with situations where there is a more balanced fit to most of the data). For similar reasons, inverse modelers generally filter out data corresponding to model data misfits that are more than several times the standard deviation of the whole set of model-data misfits they use (e.g., Chevallier et al 2010). Looking at extreme observation impacts can be viewed as an alternative way to detect situations where we could give an excessive weight to large observation errors.

We made such an analysis when running a preliminary *ini* inversion. We actually removed two data points in November (instead of 1 data point in November and 1 data point in December as erroneously said in the manuscript). The data points removed had both an impact of nearly -0.3MtCO2 on the emission budget for November (i.e. ~ -0.6MtCO2 in total). The impact of other data during this month were not more negative than -0.1MtCO2 (such a situation was extreme as demonstrated by the fact that we removed 2 data points only). In both cases, this was driven by high prior model-data misfits in combination with relatively low vertical mixing conditions in November.

Opposed to what the manuscript could have let think, these data correspond to wind directions for which the gradients are not selected in the reference configurations of the observation vector. Therefore, the impact of this removal applies to the *ini* experiment only, and it is quite negligible for our study.

We will expand a bit the text to discuss this but given the weak impact it has for the results, we will keep it short.

Reference:

Chevallier, F., P. Ciais, T. J. Conway, T. Aalto, B. E. Anderson, P. Bousquet, E. G. Brunke, L. Ciattaglia, Y. Esaki, M. Fröhlich, A.J. Gomez, A.J. Gomez-Pelaez, L. Haszpra, P. Krummel, R. Langenfelds, M. Leuenberger, T. Machida, F. Maignan, H. Matsueda, J. A. Morguí, H. Mukai, T. Nakazawa, P. Peylin, M. Ramonet, L. Rivier, Y. Sawa, M. Schmidt, P. Steele, S. A. Vay, A. T. Vermeulen, S. Wofsy, D. Worthy, 2010: CO2 surface fluxes at grid point scale estimated from a global 21-year reanalysis of atmospheric measurements. J. Geophys. Res., 115, D21307,- doi:10.1029/2010JD01388

Pg7 lines 20-31. There are a number of minor typos in this section.

We will correct them.

Pg 7 llines 20-31. Ylag experiment. How does the ylag experiment account for the evolving boundary layer during the day?

The boundary layer (BL) does not evolve only in time, but also in space. The concern about a varying BL for the gradients was raised in a more general way during the review of the study of Bréon et al., 2015 (see the discussion with reviewer 1 http://www.atmos-chem-phys.net/15/1707/2015/acp-15-1707-2015-discussion.html). Because of the variability of the boundary layer, the gradients by themselves cannot be a perfect representation of the enrichment of CO_2 in the air when crossing the city, even though we expect it to be a good proxy for it. But, in any case, the assimilation of gradients better characterizes it than the assimilation of individual CO_2 measurements for any wind direction, especially since the signature of remote fluxes on the atmospheric CO_2 concentrations is expected to be well mixed and does not to evolve much with BL variations over short durations and distances.

Above all, our atmospheric transport model simulates this variability of the boundary layer and therefore the inversion accounts for it when assimilating the gradients. The uncertainty in the modelling of the boundary layer is part of the model errors. Such a model error is a traditional source of error in inverse modelling systems, which we do not overcome by assimilating gradients. The role of setting-up the R matrix in the inverse modelling system is to account for such errors.

This topic will now be discussed earlier when we recall the concept of gradients in the introduction but we will mainly refer to the similar discussions in Bréon et al. 2015.

Pg 8 lines 16-21. This paragraph should come before rather than after the preceding one.

We will reverse the order of the paragraphs.

Pg 8 lines 29-31. Are these the emissions that are in the model domain but outside Ile-de-France? Not clear.

It is a combination of the influence of these emissions, and that of the model boundary and initial conditions (any component of the model that is not controlled by the inversion). The beginning of the sentence at line 31 and the phrase "emissions outside Île de France" instead of "emissions outside Île-de-France but within the modeling domain" are indeed misleading.

We will clarify these sentences in the revised manuscript.

Pg 10 line 1. (first sentence). Is the conclusion you state from your work or elsewhere? If the latter, please reference.

It comes from our own work. We will clarify it in the revised manuscript.

Pg 11 lines 17-19. Please reference the independent analysis that shows the temperature dependence.

We will clarify the fact that this analysis has been conducted by one of the co-authors of this paper (François-Marie Bréon) so that we cannot cite it as a personal communication.

To complement our answer, here are some details about the analysis he produced and which has not been published. The figure below shows the variations of the electric consumption in the Île-de-France region as a function of the temperature in 2013 at 15:00. Similar figures for the other hours of the day demonstrate a similar threshold of the temperature at ~19° below which the consumption is highly (negatively) correlated to the temperature (with similar slopes for each hour). The electricity consumed in Île-de-France is mainly generated by nuclear power plants. Given their clear dependence on atmospheric temperature, these variations, however, very likely reveal variations in heating behavior. And, a large part of heating in Île-de-France is based on gas consumption, which is responsible for a large part of the CO_2 emissions in the Paris area. Assuming that users of gas and electric heaters display similar heating behavior, this result supports the assumption of the temperature dependence described in the manuscript.
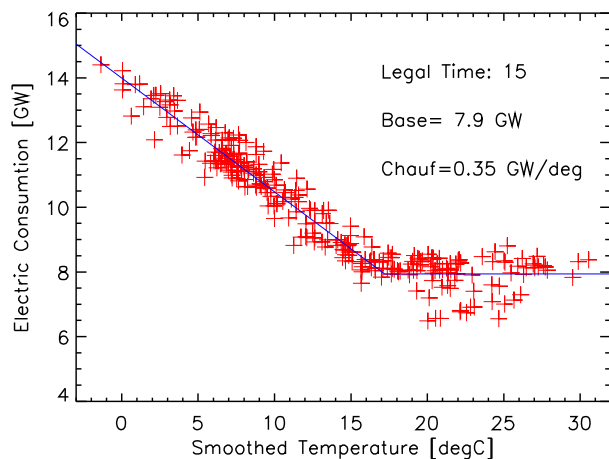
Pg 11 lines 27-32. This paragraph seems spurious.

We will shorten it. We will remind how the uncertainties shown in the figures are derived and will not warn about the fact the uncertainties will not be analyzed in detail (this will be explained in the discussion only). A sentence will be added on the uncertainties arising from the inversions at the end of section 3.2.

Pg 12 lines 22-24. Can a problem with a bias in one observation site during the month of December be ruled out as a cause of this December anomaly?

We do not think so since both types of gradients GIF-MON and GIF-GON drive the inversion in the same way (see the next sentence at l.25). We thus cannot connect it to a bias at MON or GON. And a local bias at a given site would impact SW gradients too and would probably impact the site during other months.

Pg 13 "r2" not "R2" in several places.

We will apply this correction.

Pg 16 line 22. You have not demonstrated that GON-MON gradients are "not evidently related to the whole city emissions". See also my earlier comment.

We think the small distance between these two sites and the fact that emissions between two sites are significantly impacted by the Roissy Charles-de-Gaulles airport is sufficient to assume that the gradients between these two sites cannot be representative of the city-scale emissions.

We will remind in this sentence that the improvement is assessed through the analysis of the inverted emissions.

Pg 16 lines 24-31. nd pg 17 lines 18-20. Have you tried inverting with the AIRPARIF2010 prior? Does it pull the posterior values down even further? Or not? If every inversion pulls the values lower, does this imply a fundamental flaw in the inversion?

We did not have access to the spatialized inventory corresponding to the AIRPARIF (2013) report on the emissions for 2010 (in which we extracted the so-called AIRPARIF2010 annual data) for this study. Still, we have conducted tests with prior values lower than that of AIRPARIF2008 in the FLAT_3.0 experiments and it is one of the topic of the paragraph on p17 line 18 to p18 line 2.

In the case of FLAT_3.0H, the inversion increases the emissions, so that our system does not systematically pulls the values lower. Our argument regarding the decrease of the annual emissions in the FLAT_3.0M experiment is that this decrease, in practice, does not arise from a systematic decrease of the monthly emissions, and that it is quite small (so that we have a strong convergence from the ensemble of prior values to the ensemble of posterior values obtained in the set of experiments).

In the paragraph p17 line18 to p18 line2 we will more clearly separate the discussions on a.) the potential impact of using relative prior uncertainties in the convergence of the results, and b.) the decrease of the annual emissions when using the day-to-day variations from AIRPARIF2008.

Pg 17 lines 1-2. "a large fraction of the Paris emissions are due to domestic and commercial heating". Add "believed to be" or something similar.

We will change the parenthesis (43%) into (43% according to the AIRPARIF2008 inventory) in this sentence.

Pg 17 lines 29-31. This is a very awkward sentence.

We will reformulate it.

Pg 18 lines 15-18. Be clear here that your inversion solves only for mid-pm observations, and

your analysis does not exclude that the poor representation of the diurnal cycle could have a strong impact on nighttime emission estimates.

Actually, we must acknowledge an error when running the FLAT_mD experiments. The actual results from FLAT_mD are far closer to those from FLAT_mM than to those from FLAT_mH, opposed to what was indicated in the manuscript. This is far more logical given that the inversion can control 6-hour mean fluxes and thus, to some extent, the day to day variations. As highlighted by the reviewer, it, however, cannot control nighttime emissions except through the indirect extrapolation of the information driven by the correlations in the prior error covariance. Still, of note is that the hourly mean emissions for the 11h-16h window is only ~30% above the hourly mean emissions for the whole day in the AIRPARIF2008 inventory.

We will correct the results of experiment FLAT_mD and make a clear conclusion regarding the impact of the description of diurnal cycle in the prior estimate of the 6-hour mean budgets.

Pg 20 lines 13-15. Presumably actual nighttime measurements would be useful to constrain nighttime emissions as well!

Initially, the sentence assumed that, in principle, it should be easier to model parts of the PBL's transitional phases in late morning and late afternoon, during which we have some level of mixing, than to model the PBL at nighttime. But, such a discussion would be too loose at this point and we agree with this comment by the reviewer. The sentence will be rewritten to encompass nighttime data.

Pg 20 line 18. The mention of Recife seems irrelevant.

We will remove it.

Pg 20 line 25. What is GB?

This refers to Gregoire Broquet.

Figure 3. Caption is inconsistent with labelling on graph (a, b, etc). Also, on the right hand panels, the numbers at the top are hard to follow (they are fine on the left panel).

Thanks for pointing to the inconsistent figure labels. It will be corrected.

Reply to comments by Anonymous Reviewer #2:

We thank reviewer 2 for her/his technical review of our manuscript. It gives us the opportunity to improve the clarity of our text, since most of the major comments were already discussed in the submitted version of the manuscript and seem to come from a misunderstanding of the main concept underlying our assimilation of gradients. We hope that the following additional explanations will clarify the concept and results of our study.

1 Overview:

Review of "*A first year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion*" by Staufer et al.

Staufer et al. present a year-long estimate of CO_2 emissions in Paris using the "Gradient Method" developed by Breon et al., ACP (2015). This manuscript is largely just an extension of the Breon et al. (2015) paper.

Our manuscript represents much more than "just an extension" of the Bréon et al. (2015) paper. As its title implied ("An attempt at…"), Bréon et al. (2015) showed first results of Paris-area CO_2 emission estimation from atmospheric measurements during a few weeks, based on a gradient approach. The present study brings deeper analysis of the concept of assimilating cross-city gradients and evaluates the inversions strengths and weaknesses with the results from a full year worth of measurements, an improved method, and a series of sensitivity tests to the main components of the inverse modeling system. Technically, it represented a large amount of work (for instance, we had to rebuild the **H** matrix in the inverse modeling system with the Meso-NH/TEB-CHIMERE transport configuration). We also paid a lot of attention not to be redundant with Bréon et al. (2015) in terms of analysis and discussions in the manuscript. The two papers should rather be read as two complementary studies.

They now use a longer record of observations but they include fewer sites and use more stringent data selection. The manuscript is fairly well written. However, I have some serious concerns with the manuscript. In particular, the authors need to justify some of the crucial assumptions in the "Gradient Method".

As detailed by the answers below, in principle, there is no reason to think that our gradient method is more prone to errors than a more straight-forward inversion approach that would avoid adapting the observation and control vectors to the weaknesses of the models. Both this paper and Bréon et al. (2015) actually produce a series of analyses to verify that the gradient approach definitely helps to overcome some practical limitations of the traditional inverse modelling framework.

This will be better stated in the discussion section of the revised manuscript.

2 Major comments:

I have some major concerns with assumptions made in the "Gradient Method".

2.1 Spatio-temporal offset between upwind and downwind measurements

The authors assume, as in Breon et al. (2015), that the difference between measurements made at two of their surface sites are directly comparable and that the difference between them can be related to the city-scale emissions. However, it's not clear to me that this is a valid assumption. My issues with this assumption are: (1) a temporal lag between the measurements, (2) a spatial offset between the measurements, and (3) a poor choice of model to evaluate this.

These issues were already discussed in the paper. (1) was addressed by the introduction of a time-lag between upwind and downwind measurements that did not improve the results (Figure 5, Section 3.2). (2) was addressed by the use of a new strict wind selection (Section 3.2). (3) was addressed by changing the meteorological product (from ECMWF to Meso-NH/TEB) for the forcing of CHIMERE and the gradient selection (Section 3.3.3).

The critical question is: "Did the downwind measurement actually originate near the upwind site?"

This is not a critical question for our concept of relating the cross-city gradients to the city-scale emissions. As explained in the introduction and section 2.2 of the manuscript, this concept assumes that we cancel or at least decrease the impact of fluxes outside the city by assimilating cross-city gradients instead of individual measurements (because we assume that this impact has a relatively large spatial and temporal scale), and that the information about the city emissions, that is contained in the assimilated cross-city gradients, corresponds to a large part of the city rather than to just small portions of it.

Ensuring that the downwind measurement correspond to an air mass that actually originates near the upwind site at the time of the upwind measurement would help to verify the first assumption (which is why we have conducted the test case "*lag*" with a time-lag between the upwind and downwind data in the gradients). However, this is not a requirement since:

- the impact of the boundary conditions, if not also that of a major part of fluxes outside the city, can be assumed to be quite smooth in time and thus quite similar over a 2-h window (see the debate about this specific time-lag below) due to atmospheric mixing

 - we do not need a perfect canceling of the impact of boundary conditions when assimilating gradients instead of individual measurements to improve the inversion behavior.

Both the Bréon et al. 2015 paper and our study demonstrate that even though our approach is not perfect (in the sense that there likely remains some impact of the remote fluxes in the measured gradients) it yields better results than if assimilating individual measurements for any wind direction.

This discussion will be expanded in the new manuscript (in the introduction, method and

discussion sections).

This reply and that to the general comment 1 answer in a general way the following questions within comment 2.1. Our answers to these following questions focus on the specific technical points that they raise, without systematically reminding what is said above.

Temporal Lag: The authors touch on this issue in the final paragraph of Section 3.2.

Actually the discussion on this topic starts in the last paragraph of Section 2.2 and ends at the beginning of Section 4.3.

The upwind and downwind sites are quite far apart in space and it will take a few hours for the airmass to travel from the upwind site to the downwind site. I did a quick back of the envelope calculation for the transport time using distances from Breon et al. (2015; Table 1). GIF is 23 km from the Paris centre, GON is 16 km from the Paris centre and the EIF site is directly between them. This would suggest these sites are about 40 km apart (Google Earth puts the distance at about 39km). A 3 m s^{-1} windspeed (the minimum windspeed criteria) would give a transit time of about 4 hours, quite a bit longer than the temporal lag of 2 hours reported in Staufer et al. (although it's not clear how they estimate their temporal lag).

3 m/s is the minimum wind speed near the surface (i.e. at station height) for the assimilated gradients. Over the year, the average wind speed near the surface is about 4.5 m/s. Looking at the wind speed near the surface is relevant when aiming at avoiding too much influence of sources near the measurement sites. The transport through the city, however, should be better characterized by the average wind speed in the PBL. We actually selected a 2-hour time-lag because the average wind speed 100m above the ground is about 7 m/s over the year.

It will be clarified in the revised manuscript.

This means that the upwind measurement should be compared with an observation that is at least 2-4 hours later in order to relate it to the city-scale emissions (assuming there were no changes in wind direction over a 2-4 hour period).

However, a 4 hour difference in the observed airmass will make it difficult, if not impossible (without a model), to relate the two measurements because CO_2 has a strong diurnal cycle (see, for example, Fig. 1 from McKain et al., 2012 for the Salt Lake City diurnal cycle).

Here, we use an inventory with diurnal temporal profiles and an atmospheric transport model to relate the measurements and to account for such a diurnal cycle. These models are not perfect. However, uncertainties in the modeled diurnal cycle of the emissions and in the atmospheric transport (which relates emissions at a location and time and concentrations at another location and a later time, i.e. the basic concept of atmospheric inversion) impact any of the existing inverse modelling approaches. There is no reason to assume that it would be a more critical issue for the gradient simulation.

It would be incorrect to compare measurements made at different hours because they are

influence by external processes (like boundary layer growth).

The reviewer use the term "compare" in a misleading way. The difference between the two measurements is interpreted with the help of a transport model in the same way that emissions at a given time are connected to measurements at later times and at other locations with the help of a transport model in the basic concept of the inversion.

See also the answer to the first reviewer who focuses on the boundary layer:

"The boundary layer does not evolve only in time, but also in space. The concern about a varying BL for the gradients was raised in a more general way during the review of the study by Bréon et al. (2015) (see the discussion with reviewer 1 http://www.atmos-chem-phys.net/15/1707/2015/acp-15-1707-2015-discussion.html). Because of the variability of the boundary layer, the gradients by themselves cannot be a perfect representation of the enrichment of CO2 in the air when crossing the city even though we expect it to be a good proxy for it. But, in any case, the assimilation of gradients better characterizes it than the assimilation of individual CO2 measurements for any wind direction, especially since the signature of remote fluxes on the atmospheric CO2 concentrations is expected to be well mixed and not to evolve much with BL variations over short durations and distances.

Above all, our atmospheric transport model simulates this variability of the boundary layer and therefore the inversion accounts for it when assimilating the gradients. The uncertainty in the modelling of the boundary layer is part of the model errors. Such a model error is a traditional source of error in inverse modelling systems, which we do not overcome by assimilating gradients. The role of setting-up the R matrix in the inverse modelling system is to account for such errors.

In the revised manuscript, this topic will be discussed earlier when we recall the concept of gradients in the introduction. We, however, mainly refer to similar discussions by Bréon et al. 2015."

The authors do touch on this issue in the last paragraph of Section 3.2 and show that it is a problem (inversion is not consistent with their other inversions) but disregard it anyway.

We think that our results (and analysis in the corresponding text) show that it is not a problem of inconsistency, but rather a problem of a weak constraint from the reduced datasets when a time-lag in the gradients is considered. Qualitatively, the corrections are consistent between experiments *ref* and *lag* but the amplitude of the corrections is smaller in the latter. In the revised manuscript, the text will be slightly modified to explain this in a clearer way.

Spatial Offset (vertical mixing): It is unlikely that the airmass measured by the downwind instrument was actually in the direct vicinity of the upwind instrument, there is almost certainly some spatial offset.

Such a (random) offset should not be an issue as long as the variations in time and space of the impact of fluxes outside the Paris area on CO2 concentrations are relatively small over periods and distance that correspond to the typical "space and time offsets" that are tolerated by the

method. If these variations were large, the gradient computation would only fail to decrease this impact compared to using a traditional observation vector.

The author's data selection criteria are designed to minimize the horizontal offset between the downwind airmass and the upwind instrument

Indeed, the choice of a significant range of wind directions for the gradient selection acknowledges the acceptance for such an offset, even if the amplitude of the horizontal atmospheric diffusion should be considered in such a debate. This offset is not only taken into account by the study, but its amplitude is also discussed in the manuscript through discussions on the range of wind directions to be used for the gradient selection.

but there is also probably a vertical offset. As mentioned above, the downwind airmass will take a few hours to travel from the upwind site to the downwind site and there was almost certainly some vertical motion during this period. It's possible that the downwind airmass was not even in the boundary layer when it passes the upwind instrument.

See all previous answers above and in particular the one concerning the variations in the boundary layer.

This leads into my next point.

Poor choice of model to evaluate this:

The authors are using an Eulerian model, CHIMERE, for their atmospheric transport. However, a Lagrangian Particle Dispersion Model (LPDM), such as FLEXPART or WRF-STILT, would be more appropriate for this work and would allow them to answer the critical question from above ("Did the downwind measurement actually originate near the upwind site?").

Eulerian and Lagrangian models are two types of atmospheric transport models, with both having pros and cons. To our knowledge, there has not been any study demonstrating that one is "more appropriate" than the other one at the spatial scale and resolution of our study. Note that the Eulerian CHIMERE model has been selected to produce the official operational forecasts of air pollution in the Paris area and a large number of studies on the chemistry-transport of pollutants in the Paris region have been based on this model.

Technically, Lagrangian models are also not exclusively required for computing so-called emission and concentration "footprints". The adjoint of the CHIMERE Eulerian model (e.g. Broquet et al., 2011) could be used for the computation of the atmospheric "footprints" of the downwind measurements.

Whatever tool is used, answering the reviewer's question ("Did the downwind measurement actually originate near the upwind site?") is not straightforward because of the atmospheric diffusion and because of model uncertainties. There is not a clear threshold of the sensitivity of the downwind measurements to the concentration at the upwind site above (below) which we should say that the corresponding air mass travelled (did not travel, respectively) from the upwind to the downwind site. And, the uncertainties in the state-of-the-art meteorological

forcing is such (with typically more than 10° errors on the hourly wind direction at a given location; Lac et al. 2013, discussion on Bréon et al. 2015) that it would be vain to expect, based on footprint computations, a precise metric of whether "the downwind measurement actually originate near the upwind site". (Accounting for such modeling uncertainties motivated some of the sensitivity studies that we have presented and the analysis of the selected gradients when using different models as meterological forcing).

An LPDM will simulate trajectories for individual measurements which would allow the authors to easily determine if the downwind measurement actually originated near the upwind measurement.

The reviewer assumes that LPDM provides perfect simulations of the actual transport and we disagree on this point (see above).

This would greatly simplify the "data selection criteria" because the authors would just need to find a minimum distance between the downwind trajectory and the upwind measurement site. It would also provide an appropriate time-lag for each measurement.

By speaking about a "minimum distance" (with a "distance" which would likely be difficult to define precisely), the reviewer acknowledges the fact that it would be quite impossible to gather a significant number of gradients for which "the downwind measurement actually originates near the upwind site". This further decreases the legitimacy of the reviewer's assumption that the use of adjoint transport simulations would be far better than wind analysis to check this. The minimum distance proposed by the reviewer corresponds to the "offsets" that she/he criticizes earlier.

Instead, the authors have chosen to develop a set of "data selection criteria" that are extremely difficult to evaluate because they do not actually know the trajectories of their measurements (e.g., they spend a lot of the text arguing that the Breon et al. criteria are not stringent enough).

On the contrary, the data selection criteria can be evaluated in terms of the cancelling of the components of the natural fluxes and boundary conditions in the simulated CO_2 gradients, and through the evaluation of the inversion results. The discussions on the fact that the Breon et al. criteria is not stringent enough rely on the simple fact that this one does not ensure that air masses reaching the downwind site have crossed a major part of the city at all.

 It also means that they end up throwing out most (92%) of their data...

The filtering would be even larger if we would be willing to ensure that "the downwind measurement actually originate near the upwind site" whatever transport model is used to evaluate it. And, if "offsets" and "minimum distances" are tolerated, the rate of selected data would be correlated to the corresponding level of tolerance.

2.2 Aggregation Error and Design of the Control Vector

The authors use a very crude control vector. It is a single scaling factor for the entire Paris region (with some temporal resolution). This means that they assume the gradient between their

two sites is representative of the ENTIRE Paris region and that the entire Paris region should be scaled up or down. As the authors mention (Page 6, Line 25), this will induce large aggregation errors because different parts of the city can no longer vary independently.

The authors acknowledge that this is a problem

Rather than naming it "problem" we prefer to remind that it is a source of uncertainty (such as transport errors) which is accounted for in both the set-up of the R matrix (see section 2.5) and when conducting a test of sensitivity to the spatial distribution of the emissions.

in Section 4.2 when they say: "The inversion results, however, are significantly affected by changes in the emission distribution. It reveals the need to rely on robust, high resolution emission maps such as those produced by local agencies like AIRPARIF." It seems more likely that this issue is due to aggregation error.

The two points are strongly connected. By looking at the mismatch when using two different maps of the emissions, we actually characterize the aggregation error. While the sensitivity to these maps is significant (as highlighted by the citation of our discussion section by the reviewer), the differences between the results from *ref* and INV_mapIER is not dramatic enough to form the basis for a major concern regarding our control vector.

For example, the authors could deduce a large underestimate in emissions if the downwind airmass happened to pass over some point source that is missing in the inventory. Because the authors only have a single scaling factor for the entire Paris region, the whole city would be scaled up and the emissions would now be overestimated.

Unlike what is sometimes assumed, solving for the emissions at the grid scale does not prevent from bearing aggregation errors. The use of relatively long spatial (and temporal) correlations in the prior uncertainty covariance matrix for the uncertainty in grid-scale emissions results in errors that correspond to aggregation errors. Such a use of long correlations, however, is necessary to regularize the inversion and to extrapolate the information from very scarce observation networks.

In the example given by the reviewer, given our observation network, a grid-scale inversion would not detect and solve for the missing point source. It would scale up the emission in an area corresponding to the typical spread of the measurement emissions footprints inflated by the prior error correlation length scale.

Our choice of the control vector implies an aggregation over the whole Ile de France region. But, almost all of the emissions in this region concentrate within the ~40 km x 40 km Paris urban area.

More importantly, our choice of the control vector is in line with our selection of the observation vector (which is an important aspect of our study): by assimilating "cross-city gradients" we minimize the impact of aggregation errors (see below) while we still account for them in the diagnostic or the **R** observation error covariance matrix from Bréon et al., who use gradients that were representative of smaller parts of the city. From this point of view, the

example given by the reviewer does not fit: the characteristics of the emissions in the Paris area (dominated by widespread sources) and the assimilation of cross-city gradients prevent the mis-location of a given point source from impacting the inversion in the way described by the reviewer.

We will slightly expand our discussions on the aggregation errors in the revised manuscript by adding part of the above answer in the introduction and discussion sections.

So the question is, "Are these gradients actually representative of the ENTIRE Paris region?".

The question is rather, "are these gradients actually representative of the Paris urban area ?" (see above). It is difficult to give an objective answer to this question. Still, the comparison of the results from experiments *ref* and INV_mapIER demonstrate that we can answer in a rather positive way.

This is why groups typically solve for fluxes at high-resolution.

See above for the answer regarding groups solving fluxes at high resolution but use long spatial correlation scales for the prior uncertainty in the emissions. We assume that the reviewers refer to inverse modeling groups in general since there have been very few attempts at conducting city-scale inversions. To our knowledge, there has been only one attempt at solving for the emissions at "high resolution" (i.e. Lauvaux et al., JGR, 2016).

From our experience, the city scale inversion activity requires adapting the traditional framework of the inversion at larger scales. As discussed in the final section of the paper, we propose some strategies to overcome the main issues that we first encountered, and showed first promising results. We hope that future observation frameworks and modeling techniques will allow us for solving the emissions at a higher resolution and for assimilating much more data. But we assume that it will require time, resources and research.

References:

Lauvaux, T., et al. (2016), High-resolution atmospheric inversion of urban CO2emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), J. Geophys. Res. Atmos., 121, 5213–5236, doi:10.1002/2015JD024473.

2.3 Boundary Layer Heights (PBLH)

Accurate simulation of the boundary layer heights is crucial for modelling urban CO_2. It controls the size of the box over which the emissions are mixed (e.g., CO_2 concentrations are lower during the daytime even though emissions are peaking). However, there is no discussion of the boundary layer heights in the model. Are they reasonable?

Lac et al. (2013) gave a positive assessment of the PBLH of their Meso-NH simulations (which

are used in our study) based on comparisons to PBLH measurements in the Paris area. Given the lack of PBLH observations, we did not resume such an analysis, neither in the Bréon et al. (2015) paper nor in this study. However, through our comparison of the results between using ECMWF and Meso-NH as forcing of CHIMERE, we provide an assessment of the impact of uncertainties in the meteorological simulation, which include uncertainties in the modeling of the PBLH (see Section 3.3.3). Finally, the diagnostic of the observation error by Bréon et al. (2015), which is used to set-up our inversion system, should also encompass uncertainties in the PBLH.

This will be better discussed in the revised manuscript.

2.4 Resolution and representativeness of the measurements, prior, and meteorology

The authors use a fairly coarse resolution model (CHIMERE at 2km

The use of the adjective "coarse" is highly subjective in this comment. In principle, Eulerian meteorological or transport models like WRF and CHIMERE can hardly be used at the subkm scale resolution. The type of models used for solving the meteorology-transport at higher resolution face the difficulty of simulating turbulent patterns which are quite impossible to fit.

with meteorology at 15km resolution)

In the INV_MNH experiment, we actually use a meteorological product at 2km resolution.

even though previous work (e.g., McDonald et al., 2014) has shown that 1km is necessary to resolve highways. From the abstract of McDonald et al. (2014): "High CO_2 emission fluxes over highways become apparent at grid resolutions of 1 km and finer." Why is 2km sufficient here?

Over distances, the representation errors vanish due to atmospheric mixing. We do not need to resolve precisely the local impact of individual highways or chimneys (or individual cars and buildings) on concentrations when we target the emissions at the city-scale using peri-urban sites at the edges of the urban area.

Along a similar vein, the authors use surface observations from 4 to 7 meters above ground level. Are these measurements actually representative of a 2km × 2km region (which is assumed since their grid is 2km × 2km)?

These measurements are made in 2km × 2km "regions" at the edge of the urban area without any major CO2 source, and under wind conditions where the CO2 signal should be dominated by the signature of upwind fluxes (so that the representation error should be limited). Campaigns of dense arrays of measurements at high frequency across these 2km×2km "regions" would be needed to verify such an assumption. However, we do not have such campaigns to support this assumption. Still, the Desroziers diagnostic of the observation error used by Bréon et al. (2015) encompasses the potential representation error.

2.5 Separating fossil fuel and biosphere fluxes

You don't have measurements that would distinguish between the two. It seems that most of this separation is from assumptions you've made, not data. Further, it seems that there could be compensating errors because that separation is unconstrained. As such, it seems like the title might be overstepping. "Fossil fuel" should probably be removed from the title.

See the answer to a similar comment by the first reviewer:

"We agree that the topic of the natural fluxes is critical for city scale inversions and we agree that we need to discuss it in this paper. However:

- Paris is a particular case in the sense that it is a very dense urban area with limited vegetation inside it.
- According to the model simulations, the computation of gradients succeeds in cancelling the impact of biogenic fluxes in the data that are analyzed (it was too rapidly mentioned in the last section of the paper).
- The weak impact of biogenic fluxes on the inversion of anthropogenic CO2 (according to the modelling framework) has already been addressed in more detail in Bréon et al. (2015). In addition, the specific topic of the cancelling of the biogenic component in the simulated CO2 gradients has also been analyzed in the Boon et al. 2015, paper.
- The new computation of the gradients proposed in this study further decreases the potential impact of the biogenic fluxes in the inversion.

Therefore, we expand the introduction by discussing the cancelling of the natural component in the CO2 gradients according to the model. We will indicate in the result section that the amplitude of the natural component of the simulated CO2 gradients is further decreased when narrowing the wind range for the gradient selection. Furthermore, we will expand the discussion on this topic based on such an analysis and carefully remind that actual gradients in the measurements may bear a larger signature of natural fluxes than the simulated gradients."

Also, what do the biosphere fluxes look like in the different inversion cases? Are they unchanged or could they be compensating for some of the changes you're seeing?

By assimilating cross-city gradients, we impose a very weak constraint on the natural fluxes (which are really weak within the Paris urban area, see above). This topic was already analyzed by Bréon et al. 2015 and this is emphasized here by the use of narrower wind ranges for the gradient selection. This explains why we prefer not to analyze it in this manuscript. We will clarify it in the introduction of the revised manuscript.

3 Minor comments:

3.1 Dimensions of your matrices?

What are the dimensions of the different matrices (e.g., are H^{map}, H^{trans}, and H^{samp} all the same dimension? How are they multiplied together?) Even just reporting them in terms of

something like: "H is an $n_y \times n_x$ matrix" (can use other notation) would be helpful.

The dimension of the different matrices will be provided in the revised manuscript.

3.2 Clarify how you are constructing the H matrix

From Section 2.3.1, it seems that you're doing a brute-force construction of the H matrix (e.g., perturbing each element in your control vector in a separate simulation), is that correct?

We compute the "response function" to each control variable. In the revised manuscript, we will detail this computation in a separate subsection of 2.3.

4 Specific comments:

Page 1, Line 4: Can you reliably estimate 6-h emissions with just 4 hours of obs?

This question is the main topic of our paper and of the Bréon et al. paper. In the absence of a more specific question, we can only refer the reviewer to these two papers.

Page 4, Line 5: Are these errors unbiased if the PBL height were wrong?

Yes, if the errors in PBL do not bias the simulation of CO2 gradients (typically if the modeled PBLH is unbiased).

Wouldn't that induce a systematic bias? Was that tested for?

We did not conduct sensitivity tests of the simulated CO2 as a function of different scenarios for errors in the PBL (if this is what the reviewer refers to). But we think that this is out of the scope of this study. See the answers to the generic questions above regarding errors in the simulated PBLH.

Page 4, Eq. 1: This form assumes that $n_y > n_x$, is that true?

Not necessarily. It just assumes that nx is small enough. Of note is that in the *ini* experiment, there are months during which ny>nx. The order of magnitude of ny and nx is the same in this study, so this is really a detail (see below).

Given the small number of observations, I would have guessed the other form would be more computationally efficient.

Since nx is small, it is extremely efficient as it is. We test different configuration of y in this study and as explained in the discussion section, there were plans to increase the network at the time of this study (and we have actually run inversions with more sites since the end of this study). This is why we preferred using the formulation of Eq.1 to write the code for the inversion system.

Page 4, Lines 26-27: However, the diurnal cycle is probably largely unchanged because you're

only using afternoon observations. I doubt there's much change to any other time periods.

Two of the 6-hour windows are directly constrained by the observations. There are some indirect constraints from the correlations in the B matrix. This was analyzed in the Bréon et al. 2015 paper and we will remind their conclusion regarding it.

Page 5, Lines 28-31: What about vertical gradients? These sites are all in the boundary layer so the airmasses might be fundamentally different. . .

The impact of remote fluxes diffuses horizontally and vertically upwind of the city. There is no reason to think that they will generate strongly different vertical gradients between the PBL and the free troposphere upwind vs downwind the city.

Page 6, Lines 17-26: Why not solve at a finer spatial scale? This would greatly reduce the aggregation error.

See the answer to the general question above about this topic.

Page 7, Lines 27-28: How do you estimate this? I get 3-4 hours using the cutoff windspeed of 3 m s^{-1} (see Major Comment 2.1).

See the answer to this major comment above.

Page 7, Lines 30-31: Why not just do a more traditional inversion with finer spatial resolution? You could jointly solve for the background concentration (see, for example, Henne et al., 2016).

See the answer to the general comment above about this topic. In addition to this general answer: we preferred not to rush into controlling emissions at fine spatial resolution and parameters associated to the "background concentration" and turn a blind eye to the information content in the observations. Simple inverse modelling frameworks as the one proposed here help understanding and evaluating the issues and challenges associated with an inversion problem. City-scale inversion definitely bears a large number of challenges that we do not claim to solve at once (see our discussion). The blind faith in piling up control variables faces challenging practical issues if we do not run inversion systems as black boxes: how to define spatial correlations for uncertainties in anthropogenic emissions at 2km resolution without using abusive assumptions? How to reduce the number of control variables for the boundary conditions to a "manageable" number that could be constrained with the few sites available; this would easily raise large aggregation errors on these boundaries, whose impact can have large variations from day to day and over the distances of the boundaries. From our understanding, the method proposed by Henne et al. (2016) for controlling the boundary conditions as a smooth baseline would hardly cope with the day-to-day variations of the impact from the boundary conditions shown in Bréon et al (2015).

Page 9, Lines 25-31: Confusing. A plot of your covariance would be useful. Even just plotting a single row (would just be a simply x-y line plot) would be really helpful here.

The Kronecker product between the day-to-day and 6-hour window to 6-hour window would be impossible to understand in such a plot. We will rather rewrite the paragraph for clarification.

Page 13, Lines 34-35: Misleading. It sounds like you're using a Lagrangian Particle Dispersion Model, but I'm pretty sure you aren't since there was no mention of one.

Lagrangian properties do not belong to Lagrangian transport models only. But we agree that other readers could be confused by the term "Lagrangian inversion" and we will modify it in the revised manuscript.

Pages 14, Lines 1-4: This is what you should be doing though!

We assume that this comment is equivalent to the generic comment 2.1. Please see the answers to this comment.

Pages 19, Lines 30-35: Again, I think this is the correct way to use the gradient method. I don't think a lack of data is a good justification for not using it. I think that you should either: (a) use the gradient method and the lag approach, (b) provide a better justification for why you don't need to consider temporal lags, or (c) use a traditional inversion with finer spatial resolution (with time-dependent footprints).

Please see the answers to comment 2.1.

Pages 20, Lines 11-12: This is what most groups already do. . .

The sentence means that once we will be able to exploit urban measurements, we can plan to solve for–in the sense of targeting the estimation of–the spatial distribution of the emissions. As long as relying on peri-urban sites, we consider that we cannot rely on estimates that would be derived for sub-part of the city. In the revised manuscript, it will be rewritten for clarity.

A more traditional Bayesian inversion with an LPDM would allow you to solve at high spatial resolution without inducing large aggregation errors.

This crude assumption propagates the idea that atmospheric inversion is an objective technique which does not need to be precisely adapted to the specific challenges of a real case study, and that running after higher resolution always solves for the problems. We do not support this vision. Atmospheric inversion is a tool whose core parameters are informed by the user, and whose theoretical framework and technical limitations can make it extremely weak if the observation and control vectors are not well adapted.

Figure 3: (b) and (c) labels should be flipped because they don't agree with the caption.

Thank you for pointing to the inconsistent figure labels. It will be corrected.

Measurement Sites: I don't think the paper lists the height of the measurement sites. I had to go back to Breon et al. (2015) to find it.

We will provide these heights in the revised manuscript.

A first year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion

Johannes Staufer[1], Grégoire Broquet[1], François-Marie Bréon[1], Vincent Puygrenier[1],
Frédéric Chevallier[1], Irène Xueref-Rémy[1,2], Elsa Dieudonné[1,2], Morgan Lopez[1,3], Martina Schmidt[1,4],
Michel Ramonet[1], Olivier Perrussel[5], Christine Lac[6], Lin Wu[1], 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]now at Laboratoire de Physico-Chimie de l'Atmosphère, Université du Littoral, Côte d'Opale, Dunkerque, France
[3]now at Environment and Climate Change Canada, Toronto, Canada
[4]now at Institut für Umweltphysik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
[5]AIRPARIF Surveillance de la Qualité de l'Air en Île-de-France, Paris, France
[6]CNRM-GAME (CNRS-Météo-France), UMR3589, Toulouse, France

Correspondence to: J.Staufer (johannes.staufer@lsce.ipsl.fr)

Abstract. The ability of a Bayesian atmospheric inversion to quantify the Paris region's fossil fuel CO_2 emissions on a monthly basis, based on a network of three surface stations operated during one year as part of the CO2-MEGAPARIS experiment (August 2010–July 2011), is analysed. Differences in hourly CO_2 atmospheric mole fraction between the near-ground monitoring sites (CO_2 gradients), located at the north-eastern and south-western edges of the urban area, are used to estimate the 6- h mean fossil fuel CO_2 emission. The inversion relies on the CHIMERE transport model run at $2\,km \times 2\,km$ horizontal resolution, on the spatial distribution of fossil fuel CO_2 emissions in 2008 from a local inventory established at $1\,km \times 1\,km$ horizontal resolution by the AIRPARIF air quality agency, and on the spatial distribution of the biogenic CO_2 fluxes from the C-TESSEL land surface model. It corrects a prior estimate of the 6- h mean budgets of the fossil fuel CO_2 emissions given by the AIRPARIF 2008 inventory. We found that a stringent selection of CO_2 gradients is necessary for reliable inversion results, due to large modelling uncertainties. In particular, the most robust data selection analysed in this study uses only mid-afternoon gradients if wind speeds are larger than $3\,ms^{-1}$ and if the modelled wind at the upwind site is within $\pm 15^o$ of the transect between downwind and upwind site. This stringent data selection removes 92 % of the hourly observations. Even though this leaves few remaining data to constrain the emissions, the inversion system diagnoses that their assimilation significantly reduces the uncertainty in monthly emissions, by 9 % in November 2010 to 50 % in October 2010. The inverted monthly mean emissions correlate well with independent monthly mean air temperature. Furthermore, the inverted annual mean emission is consistent with the independent revision of the AIRPARIF inventory for the year 2010, which better corresponds to the measurement period than the 2008 inventory. Several tests of the inversion's sensitivity to prior emission estimates, to the assumed spatial distribution of the emissions, and to the atmospheric transport modelling demonstrate the robustness of the measurement constraint on inverted fossil fuel CO_2 emissions. The results, however, show significant sensitivity to the description of the emissions' spatial distribution in the inversion system, demonstrating the need to rely on high-resolution local inventories such as that from AIRPARIF. Although the inversion constrains emissions through the assimilation of CO_2 gradients, the results are

1

hampered by the improperly-modelled influence of remote CO_2 fluxes when air masses originate from urbanised and industrialised areas north-east of Paris. The drastic data selection used in this study limits the ability to continuously monitor Paris fossil fuel CO_2 emissions: the inversion results for specific months [..[1]]such as September 2010 or November 2010 are poorly constrained by too few CO_2 measurements. The high sensitivity of the inverted emissions to the prior emissions' [..[2]]diurnal variations highlights the limitations induced by assimilating data during [..[3]]afternoon only. Furthermore, even though the inversion improves the seasonal variation and the annual budget of the city's emissions, [..[4]]the assimilation of data during a limited number of suitable days does not necessarily yield robust estimates for individual months. These limitations [..[5]]could be overcome through a refinement of the data processing for a wider data selection, and through the expansion of the observation network.

1 Introduction

There is a high political and scientific interest in developing methods for improving and verifying estimates of fossil fuel and cement CO_2 emissions. Consequently, there is an increasing deployment of urban CO_2 monitoring networks with the objective of quantifying city emissions through the atmospheric inversion approach [..[6]](Boon et al., 2016; Duren and Miller, 2012; Lauvaux et al., 2013, 2016; Kort et al., 2013; McKain et al., 2012; Strong et al., 2011; Turnbull et al., 2015). Bréon et al. (2015), upon which this study builds, recently reported first estimates of fossil fuel CO_2 emissions of the Paris urban area during a two-month period. They used three ground-based CO_2 measurement sites at the north-eastern and south-western edge of the area and an inversion system based on a $2\,km \times 2\,km$ horizontal resolution transport model. The monitoring stations in Gonesse (GON), approximately $15\,km$ north of Paris' city centre, and in Montgé-en-Goële (MON), 35[..[7]]km north-east (NE) of Paris' city centre, were deployed by the CO2-MEGAPARIS project and operated from August 2010 to July 2011. The monitoring station in Gif-sur-Yvette (GIF), 20[..[8]]km south-west (SW) of Paris' city centre, is part of the Integrated Carbon Observation System-France long-term network. [..[9]]

The main principle of the atmospheric inversion proposed by Bréon et al. (2015) consists in constraining CO_2 emission budgets of the urban area by assimilating atmospheric CO_2 mole fraction gradients between pairs of sites located upwind and downwind of the city.

---

[1]removed: like

[2]removed: day-to-day

[3]removed: a limited number of suitable days. Therefore

[4]removed: it

[5]removed: ,

[6]removed: (Duren and Miller, 2012; Lauvaux et al., 2013; Kort et al., 2013; McKain et al., 2012; Strong et al., 2011; Turnbull et al., 2015; Lauvaux et al., 2016)

[7]removed: km

[8]removed: km

[9]removed: The

The use of cross-city gradients, rather than individual mole fractions, aims at eliminating the variability of CO_2 caused by the transport of remote and natural fluxes outside the urban area. [..[10]]It assumes that the signal from these fluxes has a relatively large spatial and temporal scale compared to the distance and transport duration between the measurement sites. These signals and the potential signal from natural fluxes within the urban area cannot be sufficiently well controlled by the monitoring network[..[11]], in particular because their large day-to-day variations cannot be filtered as a smooth baseline in the time series of CO_2 concentrations at individual sites (an approach frequently used in regional atmospheric inversions, e.g., in Henne et al., 2016). On the contrary, such signals can be as high as the signal caused by the emissions within the urban area [..[12]](Bréon et al., 2015; Kort et al., 2013; Nordbo et al., 2012). Uncertainties in remote and natural fluxes can thus highly impact the skill for inverting the urban emissions. In the simulations by Bréon et al. (2015), the ratio between the signal from the natural and remote fluxes and the signal from the urban emissions is high when analysing individual measurements. It, however, strongly decreases when analysing gradients. This weak impact of natural fluxes on inversions is on one hand due to the fact that the dense and compact Paris urban area exhibits little vegetation within its bounds. On the other hand, it is due to the fact that, upwind the city, the signal from fluxes outside this urban area is sufficiently diffused in space so that it is relatively homogeneous over the Paris urban area and constant during the duration of the transport over this area.

The selected cross-city gradients [..[13]]also provide a characterization of the increase in the CO_2 mixing ratios of air parcels that pass [..[14]]over the city. [..[15]]It is assumed that these gradients represent emissions from the entire city [..[16]]and are not highly sensitive to the distribution of the emissions. This assumption is line with the inversion system of Bréon et al. (2015) [..[17]]that controls the city-scale emissions budgets and the temporal variation of fossil fuel CO_2 emissions, but not their spatial distribution. [..[18]]However, this method should not be seen as a sort of mass balance, given that, in practice, the inversion is not set up to ensure that the upwind and downwind concentrations corresponds to the same air masses that travelled from the upwind to the downwind site. Furthermore, since the atmospheric boundary layer evolves significantly in space and time and due to the atmospheric diffusion during the transport over the Paris area, such cross-city gradients cannot perfectly represent the CO_2 [..[19]]enrichment of air parcels passing over the urban area. In addition, temporal variations of the emissions during the transport of air masses over the Paris area prevent from relating a given gradient to the emissions at a given time. The gradients need to be interpreted using a transport model and knowledge on the spatio-temporal variations of the emissions at hourly scale. In that sense, the assimilation of gradients is affected

---

[10]removed: This signal

[11]removed: and it

[12]removed: (Bréon et al., 2015; Kort et al., 2013).

[13]removed: correspond to an

[14]removed: through

[15]removed: This increase is assumed to represent the

[16]removed: . Therefore,

[17]removed: only controls the temporal variations of the

[18]removed: It uses

[19]removed: gradients between downwind and upwind stations during the

by transport modelling uncertainties and by uncertainties in the variations of the emissions at high spatial and temporal resolution, such as it the case for any inverse modelling approaches.

The inversion assimilates cross-city CO_2 gradients during afternoon to correct prior estimates of 6-h fossil fuel CO_2 emissions budgets of the Paris metropolitan area (Île-de-France administrative region). These prior estimates are derived from the AIRPARIF inventory for the year 2008 (AIRPARIF, 2012). AIRPARIF is a non-profit agency that is accredited by the French Ministry of Environment to monitor the air quality in Île-de-France. [..[20]]Even though they have a limited impact on the inversion when gradients are assimilated, the system of Bréon et al. (2015) also inverts biogenic fluxes [..[21]][..[22]]and corrects prior estimates of the biogenic fluxes from C-TESSEL, the land-surface component of the ECMWF (European Centre for Medium-Range Weather Forecasts) numerical weather forecasting system (Boussetta et al., 2013). In order to model the CO_2 gradients, the inversion uses an estimate of the fossil fuel CO_2 emission and biogenic flux distribution at $2\,km \times 2\,km$ and hourly resolution, coupled to a $2\,km \times 2\,km$ resolution configuration of the chemistry transport model CHIMERE (Menut et al., 2013).

Bréon et al. (2015) developed and tested this inversion set-up for two months in autumn 2010. The values of the AIRPARIF 2008 inventory were used to derive the prior estimates for the corresponding dates in 2010. Bréon et al. (2015) reported a significant improvement of the fit between modelled and measured CO_2 gradients by the inversion and reasonable patterns of corrections applied to prior emission values. The small number of monitoring sites and the stringent criteria for selecting gradients leads to a high number of periods, ranging from one to several days, during which the inversion does not assimilate any atmospheric CO_2 data. As a consequence, averages of the inverted emissions over one month were found to be more reliable than 1-day to 1-week mean results.

The aim of this study is to derive a full year of monthly mean emission estimates for the Paris area, based on the inversion system described by Bréon et al. (2015) and on the availability of measurements at MON, GON and GIF during the period mid-2010 to mid-2011. The 1-yr long inversion allows a better evaluation of the method by analysing the seasonal variation and the annual budget of the inverted emissions. In particular, the annual budget can be compared to the AIRPARIF emission assessment for 2010 (AIRPARIF, 2013). This assessment is based on an inventory model that has been improved since the release of the 2008 inventory. The 2010 inventory applies to a time period which better corresponds to the inversion period than the 2008 inventory used for the inversion. Therefore it provides some independent information to check the corrections applied by the inversion to the prior estimate of the annual budget derived from the 2008 inventory.

Preliminary tests of the inversion during the 1-yr period, however, revealed that the selection of gradients, as proposed by Bréon et al. (2015), do not conform fully with the underlying assumptions of having gradients dominantly influenced by urban emissions. Notably, negative [..[23]]CO_2 gradients between downwind and upwind sites were frequently [..[24]]measured when using the gradient selection criteria of Bréon et al. (2015). This led us to revise the selection of CO_2 data to form gradients.

---

[20]removed: The

[21]removed: in the region, since the urban area, while being dense, is surrounded by green areas which can cause variations in the selected

[22]removed: gradients. It thus

[23]removed: measured

[24]removed: found

The revision consists primarily on a tighter filtering of wind directions to select gradients in order to avoid situations when air parcels leaving the upwind site or reaching the downwind site do not overpass a significant part of the city and the vicinity of the other site. Section 2 presents a summary description of the inversion configuration and the [..[25]]revised gradient selection. Section 3 analyses the inversion results for different configurations. In particular, it assesses the impact of the stricter gradient selection, and the sensitivity of the results to the prior emission estimates, to the emissions' spatial distribution, and to the atmospheric transport modelling, so as to evaluate how robustly the emissions are constrained by atmospheric CO_2 data. These results are discussed in Sect. 4.

2 Inversion configuration

The inversion method described by Bréon et al. (2015) is based on the Bayesian approach. The control vector x gathers the CO_2 flux budgets. x_b is the vector of the prior estimates of these budgets, independent of atmospheric observations. The observed CO_2 mole fraction gradients selected for the inversion are assembled into y_0, which defines the observation space y. The linear observation operator $\mathcal{H} : x \mapsto y = \mathbf{H}x + y^{\mathrm{f}}$ projects the control vector x into the observation space y through the linear operator $\mathbf{H}$ (combining the description of the fluxes' spatial distribution and the atmospheric transport model) and the addition of CO_2 gradients y^{f} caused by fluxes that are not controlled by the inversion, such as the remote fluxes characterized by the CO_2 boundary conditions of the regional transport model. The uncertainties in x_b and the observation errors, i.e., errors in the measurements y_0 and from the observation operator $\mathcal{H}$, are assumed to have unbiased Gaussian distributions and are characterized by the prior uncertainty covariance matrix $\mathbf{B}$ and the observation error covariance matrix $\mathbf{R}$, respectively. x_a, the optimal posterior estimate of x, knowing its prior estimate x_b and measurements y_0, can be obtained from (e.g., Rodgers, 2000):

$$x_a = x_b + (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(y_0 - y^{\mathrm{f}} - \mathbf{H}x_b). \tag{1}$$

The uncertainty in x_a has unbiased Gaussian distribution and is characterized by the posterior uncertainty covariance matrix $\mathbf{A}$:

$$\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}. \tag{2}$$

The inversion uses measurements during a given 30-day period to derive fluxes during the same 30-day period. Independent inversions are made for twelve consecutive 30-days periods starting on 1 August 2010 to cover the entire observation period from August 2010 to July 2011. The 6- h mean inverted emissions during each period serve as the basis for the analysis of emissions in the Paris area at the monthly scale. Even though these 30 day periods do not correspond exactly to the [..[26]]calendar months, the names of the calendar months are used to label them.

---

[25]removed: revision of

[26]removed: month in the year

We briefly recall descriptions of the components of Eq. (1)–(2) as laid out by Bréon et al. (2015) in the next Sections (Sect. 2.1-2.4). As detailed in Sect. 2.2, two modifications, however, are brought to the definition of the observation space y and thus to the observation operator $\mathcal{H}$. The modifications result in two new inversion configurations that are denominated *initial* [..[27]](i.e., close to Bréon et al., 2015) and *reference* configuration hereafter. Section 2.6 presents the set-up of the sensitivity tests, where the prior estimates of the control variable, x_b, and components of the observation operator $\mathcal{H}$ are modified with respect to the reference configuration.

2.1 Control vector x and the prior estimate of the flux budgets x_b

x contains 6-h mean fossil fuel CO_2 emission budgets for windows 0-6 h, 6-12 h, 12-18 h, 18-24 h (local time is used hereafter) for each day for the [..[28]]Île-de-France region. Most of the emissions in this region are concentrated in the urban agglomeration of Paris. Thus, this choice of x approximately consists in controlling the emission budget of this urban area. x also contains 30-day mean biogenic CO_2 fluxes for each of the four 6-hour windows of the day (0-6 h, 6-12 h, 12-18 h, 18-24 h) for 9 areas that make up the Northern France modelling domain, including one that encompasses the Paris region (see Fig. 1). The inversion optimises the diurnal cycle of both the fossil fuel CO_2 emissions and biogenic fluxes through resolving these fluxes for the different [..[29]]6-h windows of the day. However, it controls the day-to-day variability of the fossil fuel CO_2 emissions but not the one of the biogenic fluxes. The inversion controls scaling factors of the flux budgets provided by the emission inventories and the ecosystem model simulations through the [..[30]]linear part of the observation operator ($\mathbf{H}$, see below). For the sake of simplicity [..[31]]we state hereafter that it controls the flux budgets themselves.

The initial and reference inversion configurations use our best available knowledge on the flux budgets—the AIRPARIF 2008 inventory (since the AIRPARIF 2010 monthly mean budgets were not available for this study) and the C-TESSEL simulation— to define the prior estimate x_b. The sensitivity tests, described in Sect. 2.6, investigate the impact of using different prior estimates for the Paris fossil fuel CO_2 emissions.

2.2 Configuration of the observation vector y and measurement vector y_0

The specific definition of the observation space y and of the corresponding measurement vector y_0 depends on the measurement availability, on the range of wind directions used to select gradients, and on the meteorological forcing of the CO_2 transport model. Two different meteorological products are used to define the wind direction for the gradient selection (see forward Sect, 2.3.1).

The three monitoring sites are located roughly along a NE-SW direction at edges of the urban area in mixed urban-rural environments (Fig. 1) at heights of 9 m (MON), 4 m (GON) and 7 m (GIF) above ground. The NE-SW direction corresponds to the dominant wind directions in Île-de-France. Technical details about the measurements are given in Xueref-Remy et al.

---

[27]removed: (i.e., close to Bréon et al. (2015))

[28]removed: whole Paris region. It

[29]removed: 6-hour

[30]removed: operator $\mathbf{H}$ (

[31]removed: however,

(2016), Bréon et al. (2015) and Lac et al. (2013). Here we briefly summarise the main aspects. The CO2-MEGAPARIS sites GON and MON were equipped with a ring-down cavity analyser from Picarro (model G1302), while an automated gas chromatograph analyser (Agilent HP6890, see Gibert et al., 2007) has been used at GIF. All measurements are quality controlled and calibrated against the World Meteorological Organisation mole fraction scale WMO-X2007 (Zhao and Tans, 2006). The

5 instrumental reproducibility of the Piccaro 5 min averages is better than 0.17 ppm, while measurement accuracy is estimated at 0.38 ppm [..[32]](Bréon et al., 2015; Xueref-Remy et al., 2016). The precision of the chromatograph analyser in GIF is estimated at 0.05 ppm for 5 min averages (Lac et al., 2013). In our study, we binned measured CO_2 data into 1-h means. The accuracy for these hourly means is better than 0.4 ppm at the three sites which is negligible compared to the modelling uncertainties (see forward Sect. 2.5).

10 Figure 2 and Fig. A1–A2 illustrate the temporal coverage of the measurements available during the CO2-MEGAPARIS period (August 2010–July 2011) at each measurement site. They also show which data are finally used to form gradients. Some significant data gaps can be noticed, e.g., during June 2010 and 2011 at GON, September 2010 at MON, January, November and December 2010 at GIF. The regular 1-day gaps correspond to instrument calibrations.

CO_2 at the measurement sites is significantly influenced by both the Paris urban emissions and the remote fluxes (i.e., by

15 fluxes outside the [..[33]]modelling domain, whose influence is simulated by the transport of the CO_2 conditions imposed at the model boundaries, and by biogenic and fossil fuel fluxes within the modelling domain but outside the Paris urban area). It is assumed, that, due to atmospheric diffusion, the signature of the remote fluxes upwind the city on the concentrations in our domain has horizontal and vertical spatial scales and a temporal scale of variability that are large enough so that it does not evolve during the transit of an air parcel above the city. In other words, it is assumed that the remote fluxes do not cause

20 CO_2 gradients between downwind and upwind stations when the wind blows from the upwind to the downwind sites. This critical assumption is supported by the fact that the simulated CO_2 gradients, caused by remote fluxes, are negligible. However, this does not necessarily imply that the measured gradients are not influenced by the actual fluxes (Bréon et al., 2015). This assumption is also supported by the much better fit between observed and modelled CO_2, when observations are defined by cross-city gradients instead of CO_2 mixing ratios at individual sites (Bréon et al., 2015). By assimilating CO_2 gradients rather

25 than individual CO_2 mole fractions, we thus expect to prevent the inversion from being sensitive to the uncertainties in the estimate of the remote fluxes.

Local sources in the vicinity of the measurement sites are difficult to represent in the model. In order to limit their impact, Bréon et al. (2015) selected gradients only if the wind speed is above a given threshold of $2\,\text{ms}^{-1}$. Similar to most inversion studies that used rural measurement sites [..[34]](e.g., Broquet et al., 2011; Geels et al., 2007), Bréon et al. (2015) assimilated

30 data during the afternoon only, since the model seemed to poorly represent vertical transport during other periods of the day. Specifically, Bréon et al. (2015) used differences in simultaneous hourly-averaged CO_2 measurements between the peri-urban stations during the afternoon (12-16 h) to define the measurement vector $\boldsymbol{y}_0$. When (at a given hour) the wind at GIF, given

---

[32]removed: (Bréon et al., 2015)

[33]removed: modeling

[34]removed: (e.g., Geels et al., 2007; Broquet et al., 2011)

7

by the meteorological simulation (see below Sect. 2.3.1), is from SW, i.e. from 160° to 260°, and is above $2\,\mathrm{ms}^{-1}$, GIF is the upwind site and Bréon et al. (2015) assimilate hourly CO_2 mole fraction differences between MON and GIF and between GON and GIF. When the simulated wind at MON is from NE, i.e. from 0° to 135° and exceeds $2\,\mathrm{ms}^{-1}$, MON is the upwind site and Bréon et al. (2015) assimilate the CO_2 differences between GON and MON and between GON and GIF.

5   Using this configuration, Bréon et al. (2015) assimilated CO_2 gradients between GON and MON. GON and MON [..[35]]are separated by only a short distance. The enhancement of CO_2 between these two sites therefore rather reflects the emissions from a small portion of the North-Eastern suburbs of Paris [..[36]]than emissions from the entire urban area. Model-data misfits for such gradients relate far more to the uncertainties in the high resolution mapping of the emissions than to uncertainties in the budget of the city emissions. In addition, these gradients are strongly affected by emissions

10   from the Charles-de-Gaulle airport which is located between the two sites and an important local source of CO_2 that is not representative of the main CO_2 sources in the Paris urban area. Thus, gradients between GON and MON are not adapted to constrain city-scale emissions. Furthermore, in order to retain a [..[37]]significant fraction of measurements in the inversion, Bréon et al. (2015) used a loose range of wind directions to define upwind and downwind conditions. This loose range could allow the assimilation of gradients when air masses leaving the upwind site or reaching the downwind site hardly

15   cross a significant portion of the Paris urban area, or, more generally, when air masses are not really transported from the upwind to the downwind site. This loose selection of gradients for constraining fluxes was not identified as a major source of systematic error. Through this configuration, [..[38]]Bréon et al. (2015) primarily aimed at decreasing the impact of remote fluxes on CO_2 mole fractions while keeping a large amount of data for the inversion. [..[39]]Both choices, the assimilation of GON-MON and the loose wind filtering to select gradients , however, lead to estimates of spatially integrated emissions

20   of the city constrained by measurements that are influenced only by emissions from a small fraction of the city. This would not be an issue if the spatial distribution of emissions provided by AIRPARIF was perfectly accurate. On the other hand, any significant error in the emissions' spatial distribution may induce a large error on the city-wide emission inversion. Indeed, if the assumed spatial distribution of the emission bears significant errors (which is likely the case), the inversion corrections, driven by model-data misfits due to errors in emissions from a small part of the city, will become inconsistent with the errors

25   at the city-scale, raising large so-called aggregation errors (Kaminski et al., 2001).

As mentioned in the introduction, preliminary tests of inversions using the [..[40]]configuration of Bréon et al. (2015) for the period August 2010–July 2011 demonstrated the need for an improved configuration where the selection of CO_2 conforms better with our assumptions on gradients. In this study, two critical changes are applied. They consist in assimilating GIF-GON and GIF-MON and in discarding GON-MON gradients when the wind is from NE. Furthermore, a stricter (narrower) range

30   of wind directions to select CO_2 gradients is used. Discarding GON-MON gradients suppresses the large amount of negative gradients in the measurement vector $\boldsymbol{y}_0$. The impact of discarding GON-MON gradients on the inversion results is not analysed

---

[35]removed: , however,

[36]removed: . Gradients between these two sites do not relate to the whole

[37]removed: lot of data

[38]removed: they

[39]removed: These choices,

[40]removed: BR2105 configuration

deeper in the following. It relates to specific details of the Paris network configuration. Here, we focuses on the impact of using a narrower range of wind directions for the gradient selection. The stricter selection of wind directions consists in assimilating a gradient between two sites only if the modelled wind at the upwind site is within $\pm 15^o$ of the transect between the downwind and upwind site. The specific choice of $\pm 15^o$ is somewhat arbitrary. On the one hand it ensures the selection of a significant

5   number of gradients. On the other hand it ensures that air masses leaving the upwind site or reaching the downwind site are transported over a large part of the urban area and in a direction that is close to the transect between downwind and upwind sites. Thus, the gradients GIF-GON, GIF-MON, MON-GIF and GON-GIF are assimilated only if the wind is from 20° to 50°, 35° to 65°, 215° to 245°, and 200° to 230°, respectively. We use the term *SW gradients* for the gradients GON-GIF and MON-GIF and *NE gradients* for the gradients GIF-MON and GIF-GON.

10   We apply other significant changes to the gradient selection criteria of Bréon et al. (2015). First, we increase here the minimum wind speed threshold at the upwind site from 2 to $3\,\mathrm{ms}^{-1}$. This change is driven by the fact that, as noticed by Bréon et al. (2015), large model-data misfits persist after inversion for wind speeds close to $2\,\mathrm{ms}^{-1}$. This suggests that a threshold of $2\,\mathrm{ms}^{-1}$ was not sufficient to avoid a large contamination of the measurements by poorly-modelled local sources. Furthermore, in this study, a single valid 1- h mean gradient during a given afternoon is not selected for the inversion. This avoids constraining

15   the emissions of a given day based on a single observation that potentially bears a large transport model error. At last, [..[41]]an analysis of the impact of individual observations on the corrections applied by the inversion to the prior [..[42]]monthly flux estimates (i.e., impact of the product between the gain matrix $\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$ and the model-data misfit for each individual gradient, see for more details Moore et al., 2011) was conducted for the initial and reference inversion experiments. It revealed that, for the initial inversion, during November, two [..[43]]gradients had far more impact

20   [..[44]]on the correction to the emissions budget of this month than the other gradients. The gradients removed had both an impact of approximately -0.3 $\mathrm{MtCO_2}$ on this budget (i.e. approximately -0.6 $\mathrm{MtCO_2}$ in total). In both cases, these high impacts were connected to high prior model-data misfits during weak vertical mixing episodes. Again, similar to many inversion experiments (e.g, Chevallier et al., 2010), in order to avoid giving too much weight to individual measurements [..[45]]the two corresponding gradients [..[46]]were removed from the initial inversion experiment. However, such gradients

25   are not selected by the tighter wind direction filtering of the reference inversion.

Three configurations of the observation space $\boldsymbol{y}$ are used in this study: $\boldsymbol{y}_{ini}$, $\boldsymbol{y}_{ref}$, and $\boldsymbol{y}_{lag}$. The first one, $\boldsymbol{y}_{ini}$, corresponds to the initial inversion configuration. It includes all the new options discussed above, except the narrowing of the wind direction ranges for the gradient selection. The selection of GIF-GON, GIF-MON, MON-GIF and GON-GIF gradients in $\boldsymbol{y}_{ini}$ is based on the wind direction ranges at GIF and MON as proposed by Bréon et al. (2015). The second one, $\boldsymbol{y}_{ref}$, corresponds to the

30   reference inversion configuration. It includes all the new options and selects gradients based on the new wind direction ranges at GIF, GON and MON defined in this section. The comparison between the initial and reference inversions is used to assess

---

[41]removed: the

[42]removed: flux estimates (analysis not shown in this study) revealed thatduring

[43]removed: 30-day periods (November, December), one of the

[44]removed: than the rest of the gradients. Again,

[45]removed: ,

[46]removed: have not been assimilated in any inversion experimentof this study

the impact of using tight wind direction ranges on retrieved emissions and to evaluate if the selected gradients now conform better with our assumptions (see Sect. 3.1-3.2). $\boldsymbol{y}_{lag}$ is only used for a single experiment whose results are briefly discussed in Sect. 3.2. This [..[47]]observation vector consists in spatio-temporal gradients[..[48]], i.e., mole fraction differences between a downwind [..[49]]site at a given time and [..[50]]an upwind site 2 h [..[51]]before. Given a mean wind speed of 7 ms^{-1} in the lower planetary boundary layer of the Paris area during the afternoon over the 1-yr period and a distance of about 40 km between the upwind and downwind site the typical time for air being transported from the upwind to the downside site is approximately 2 h[..[52]]. The wind selection in this experiment is similar to that of the reference experiment. It uses simulated wind fields at the time of the upwind mole fraction measurement involved in the gradient. At a given site the assimilation window is also reduced so that a given gradient does not involve any measurement outside the 12-16 h window, despite the use of a time lag in the gradient. The [..[53]][..[54]]use of spatio-temporal gradients instead of spatial gradients appears more in line with [..[55]]the concept of a mass balance approach which constrains emissions based on [..[56]]mole fraction variations in air parcels that are transported [..[57]]over the Paris area. Due to atmospheric diffusion and variations in the planetary boundary layer, the spatio-temporal gradients still need to be interpreted using a high resolution transport model. However, with such a configuration of the observation vector, the number of data that can be assimilated is further decreased as the assimilation window at both the upwind and downwind sites is reduced.

2.3 Observation operator $\mathcal{H}$

This section describes the observation operator $\mathcal{H} : \boldsymbol{x} \mapsto \boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{y}^{\mathrm{f}}$. The linear operator $\mathbf{H}$ [..[58]]can be decomposed into three operators ($\mathbf{H} = \mathbf{H}^{\mathrm{samp}}\mathbf{H}^{\mathrm{trans}}\mathbf{H}^{\mathrm{map}}$) consisting in the fluxes' spatio-temporal distribution ($\mathbf{H}^{\mathrm{map}}$), the atmospheric transport simulated using CHIMERE ($\mathbf{H}^{\mathrm{trans}}$), and the sampling of simulated 4D-CO_2 field like the observations ($\mathbf{H}^{\mathrm{samp}}$). $\boldsymbol{y}^{\mathrm{f}}$ gathers influences on the gradients which are not controlled by the inversion such as the signature of the model boundary conditions. In the following, we present the implementation of these operators and vectors used in Bréon et al. (2015) and the initial and reference inversions of this study, respectively, as well as alternative options used for sensitivity tests.

2.3.1 Atmospheric transport modelling and sampling

---

[47]removed: Lagrangian experiment

[48]removed: between upwind and downwind sites

[49]removed: sites

[50]removed: upwind sites 1-

[51]removed: or 2-

[52]removed: before

[53]removed: typical time for air being transported from the upwind to the downside site is 2

[54]removed: , justifying the time lag that have been tested. The

[55]removed: our approach to constrain

[56]removed: the

[57]removed: above

[58]removed: combines

The atmospheric transport $\mathbf{H}^{\text{trans}}$ and the signature of sources and sinks on CO_2 concentrations that are not controlled by the inversion ($\boldsymbol{y}^{\text{f}}$) are modelled using a Northern France configuration of CHIMERE. It has a a $2\,\text{km}\times2\,\text{km}$ spatial resolution for the Paris region, and a $2\times10\,\text{km}$ and $10\,\text{km}\times10\,\text{km}$ spatial resolution for the surroundings (see Fig. 1). It has 20 vertical hybrid pressure-sigma (terrain-following) layers that range between surface and the mid-troposphere, up to $500\,\text{hPa}$. In the initial and reference inversion of this study, as in Bréon et al. (2015), CHIMERE is driven by operational analyses of ECMWF's Integrated Forecasting System, available at approximately $15\,\text{km}\times15\,\text{km}$ spatial resolution and $3\,\text{h}$ temporal resolution. In this case we will denote $\mathbf{H}^{\text{trans}}=\mathbf{H}^{\text{trans}}_{ECM}$ and $\boldsymbol{y}^{\text{f}} = \boldsymbol{y}^{\text{f}}_{ini-ECM}, \boldsymbol{y}^{\text{f}}_{ref-ECM}$ or $\boldsymbol{y}^{\text{f}}_{lag-ECM}$, depending on the type of gradient selection used.

[..[59]]Lac et al. (2013) conducted meteorological simulations on our modelling domain using a $2\,\text{km}\times2\,\text{km}$ [..[60]]

[..[61]]resolution configuration of the non-hydrostatic mesoscale model Meso-NH. Meso-NH, jointly developed by Météo-France and Laboratoire d'Aérologie (Lafore et al., 1998), is coupled to 3-hourly analysed meteorological fields from AROME (Application of Research to Operations at Mesoscale)-France (Seity et al., 2010) and to the land-surface-atmosphere interaction model SURFEX (Masson et al., 2013). SURFEX includes the urban and vegetation scheme TEB (Masson, 2000). Therefore, in contrast to the ECMWF meteorological forcing, Meso-NH/TEB includes some urban parametrisation, which may have a large impact on the transport [..[62]]over the city. Lac et al. (2013) showed, by comparison to Lidar systems operated on a short-term basis in the Paris area, that Meso-NH/TEB captures relatively well the diurnal cycle of the boundary layer height as well as the differences in this height between peri-urban and urban locations.

A test of sensitivity is conducted to assess the impact of the uncertainties in the meteorological product on the inversions (in particular the uncertainties in the wind and in the boundary layer height). The meteorological product is used to drive the atmospheric transport model and to select the cross-city gradients. This test consists in using hourly mean outputs of Meso-NH/TEB to drive CHIMERE and to select gradients. Meso-NH/TEB simulations, originally conducted over a slightly different grid (see Lac et al., 2013, their Fig. 1a.), are interpolated onto the CHIMERE grid. When using Meso-NH/TEB $\mathbf{H}^{\text{trans}}$ and $\boldsymbol{y}^{\text{f}}$ are denoted by $\mathbf{H}^{\text{trans}}_{MNH}$ and $\boldsymbol{y}^{\text{f}}_{ref-MNH}$, respectively.

In order to build the linear part of the observation operator $\mathbf{H}$ and $\boldsymbol{y}^{\text{f}}$, the operator $\mathbf{H}^{\text{samp}}$ is applied. $\mathbf{H}^{\text{samp}}$ extracts the selected gradients between the monitoring sites from the simulated 4-D CO_2 mole fraction fields as described in Sect. 2.2. The underlying selection of the horizontal and vertical positioning of the monitoring sites in the CHIMERE grid is the same as in Bréon et al. (2015). Because the gradient selection depends on modelled wind speed and direction, the observation space $\boldsymbol{y}$ and thus $\boldsymbol{y}^{\text{f}}$ and $\mathbf{H}^{\text{samp}}$ depend on the meteorological simulations (ECMWF or Meso-NH/TEB). We denote $\mathbf{H}^{\text{samp}}$ by $\mathbf{H}^{\text{samp}}_{ini-ECM}, \mathbf{H}^{\text{samp}}_{ref-ECM}, \mathbf{H}^{\text{samp}}_{lag-ECM}$ or $\mathbf{H}^{\text{samp}}_{ref-MNH}$, depending on the inversion cases.

---

[59]removed: For sensitivity tests (see Sect. 3.3.3), CHIMERE is alternatively driven by the Meso-NH/TEB meteorological simulation at

[60]removed: and 1-hour resolution (Lac et al., 2013). This Meso-NH/TEB simulation, originally conducted over a slightly different grid (see Lac et al., 2013, their Fig. 1a.), were interpolated onto the CHIMERE grid. When using Meso-NH/TEB $\mathbf{H}^{\text{trans}}$ and $\boldsymbol{y}^{\text{f}}$ are denoted by $\mathbf{H}^{\text{trans}}_{MNH}$ and $\boldsymbol{y}^{\text{f}}_{ref-MNH}$, respectively.

[61]removed: The

[62]removed: through the cityand thus

2.3.2 [..[63]]Emissions outside Île-de-France and model boundary conditions

y^f encompasses the signature of [..[64]]fossil fuel CO_2 emissions outside the Paris region but within the modelling domain, that of the modelling domain's CO_2 boundary conditions and that of the 30-day simulations initial conditions. [..[65]]The signature of the emissions outside the Paris region but within the modelling domain are simulated by CHIMERE using fossil fuel CO_2 emissions [..[66]]from the EDGAR database (Janssens-Maenhout et al., 2012). Daily CO_2 mole fraction fields provided by the global inversion of Chevallier et al. (2010) are used as CO_2 boundary conditions at the lateral and top edges of the modelling domain and as initial conditions for the CO_2 mole fraction fields at the beginning of each 30-day period. The global inversion of Chevallier et al. (2010) is based on the simulation of the CO_2 transport by the LMDZ model (Hourdin et al., 2006) and on the assimilation of ground-based measurements from a global network.

2.3.3 Mapping of the Paris fossil fuel CO_2 emissions and biogenic fluxes

$\mathbf{H}^{map}$ is built on hourly biogenic flux and emission maps at the horizontal resolution of the CHIMERE transport model. In both the initial and reference inversions, as in Bréon et al. (2015), the description of the fossil fuel CO_2 Paris emissions at 1- h and $2 \, km \times 2 \, km$ resolution in $\mathbf{H}^{map}$ is based on the hourly AIRPARIF 2008 inventory. The temporal profiles and spatial distributions of this inventory are analysed in Bréon et al. (2015). We just recall that emissions are available at 1 h and $1 \, km \times 1 \, km$ resolution for three typical days (weekday, Saturday, Sunday) of 5 typical months (January, April, July, August, October) of the year 2008. In order to build hourly estimates for the 1-yr period August 2010-July 2011, we follow AIRPARIF's recommendation and use January emissions for all five months from November to March, April data for all three months from April to June, and October data for both September and October and, for a given day in 2010 or 2011, we use the values from the same day in 2008.

For sensitivity tests (see Sect. 3.3.2), the emission component of $\mathbf{H}^{map}$ is alternatively built based on a national emission inventory for 2005 compiled by the Institut Für Energiewirtschaft Und Rationelle Energieanwendung (IER) of the University of Stuttgart, Germany. Latoska (2009) disaggregated reported emission totals for France for 2005 into a 1×1 arc minute grid with the use of various proxies for the distribution of emitting activities such as population census, traffic intensity and land cover. We used monthly, weekly and hourly temporal profiles for different emissions sectors from the IER inventory for Europe as described by Vogel et al. (2010) to disaggregate annual emissions to hourly emissions. The IER and AIRPARIF emission inventories are two largely independent datasets.

In all experiments, the component in $\mathbf{H}^{map}$ that corresponds to the biogenic control variables is based on Net Ecosystem Exchange simulated by C-TESSEL at 3 hourly and $15 \, km \times 15 \, km$ resolution. The simulated Net Ecosystem Exchange is interpolated hourly onto the CHIMERE grid (at 2 km to 10 km resolution). The C-TESSEL model does not have a specific implementation for urban ecosystems and due to its moderate horizontal resolution, it is not expected to provide a precise

---

[63]removed: Fluxes

[64]removed: anthropogenic

[65]removed: These signatures

[66]removed: outside the Paris region

representation of biogenic fluxes within the urban area and in its vicinity. However, as reminded in the introduction , the signal from C-TESSEL in the CO_2 gradients between the peri-urban sites simulated by Bréon et al. (2015) is low. Therefore, the natural fluxes are not expected to critically affect the inversion of fossil fuel CO_2 emissions in our study (see forward Sect. 3). We denote $\mathbf{H}^{map}$ by $\mathbf{H}_{AP}^{map}$ if the hourly fossil fuel CO_2 flux maps are built using AIRPARIF 2008; by $\mathbf{H}_{IER}^{map}$ if the hourly fossil fuel CO_2 flux maps are built using IER.

2.3.4   Building the $\mathbf{H}$ matrix

In order to apply eq. (1) and (2), $\mathbf{H}$ is built based on the different operators described above. Each column of $\mathbf{H}$ corresponds to the response of the selected CO_2 gradients to a control variable. Each column of this matrix is computed by applying the $\mathbf{H}$ operator (i.e. the series of operators described above) to a control vector containing only zeros except for the corresponding control variable which is set to 1.

Let n_x denote the number of control variables (156 elements) for a given month of inversion, n_f the dimension of the 3D flux field in the input of the CHIMERE model (i.e. the number of model horizontal grid cells times the number of hours during one month of inversion, i.e, $118 \times 118 \times 720$), n_c the dimension of the 4D field of CO2 in output of the CHIMERE model (i.e. number of model grid cells times the number of hours during one month of inversion, i.e.,$118 \times 118 \times 20 \times 720$), and, at last, n_y the number of gradients selected for a 1-month inversion. The dimension of $\mathbf{H}$ is $n_x \times n_y$, while the dimension of $\mathbf{H}^{map}$ is $n_x \times n_f$, of $\mathbf{H}^{trans}$ $n_f \times n_c$ and of $\mathbf{H}^{samp}$ $n_c \times n_y$. n_x application of $\mathbf{H}^{samp}\mathbf{H}^{trans}\mathbf{H}^{map}$ are needed to build the $\mathbf{H}$ matrix. Once the $\mathbf{H}$ is built, since both n_x and n_y are relatively small, we can easily afford the computations in eq. (1) and (2) which involve the inversion of matrices of size $n_x \times n_x$ or $n_y \times n_y$ and multiplication of such matrices with $\mathbf{H}$.

2.4   Prior error covariance matrix B

[..[67]]We set-up the prior error covariance matrix $\mathbf{B}$ [..[68]]as in Bréon et al. (2015). Assuming that there is no correlation between the uncertainties in the fossil fuel CO_2 emissions and the uncertainties in the biogenic fluxes, $\mathbf{B}$ is modelled as a diagonal block matrix with two blocks: one corresponds to the uncertainties in the Paris fossil fuel CO_2 emissions, and the other one to the Net Ecosystem Exchange in the modelling domain. For each block, we make separate assumptions on the variance of the uncertainty in the individual control variable in one hand and on the temporal and spatial correlations between these uncertainties on the other hand.

Regarding the Paris fossil fuel emissions, we assume a 50[..[69]]% relative uncertainty (in terms of standard deviation) in the prior estimates of individual 6- h [..[70]]emission budgets. [..[71]]We assume that we can decompose these prior uncertainties for a given month into uncertainties in the mean diurnal cycle of the emissions and into uncertainties in

---

[67]removed: As in Bréon et al. (2015), we model

[68]removed: by assuming

[69]removed: %

[70]removed: fossil fuel

[71]removed: The temporal correlations of

the day-to-day variations of the emissions. Therefore, we compute the temporal autocorrelations of the prior uncertainties in the 6- h [..[72]][..[73]]emission budgets, $c_t(t_1, t_2)$ (where t_1 and t_2 are two 6- h windows of 2 days of the month of inversion), as the product of the correlations of the uncertainties in the mean diurnal cycle between the four 6- h windows of [..[74]]the day, $c_w(w_1, w_2)$ (where w_1 and [..[75]]w_2 are the two 6- h windows of the day corresponding to t_1 and t_2 respectively),

5   and correlations of uncertainty in the day-to-day variations between different days[..[76]], $c_d(d_1, d_2)$ (where d_1 and d_2 are the two days corresponding to t_1 and t_2). We assume that the correlations of the uncertainty in the mean diurnal cycle between the 6- h windows of the day [..[77]]are positive: $c_w(w_1, w_2)$=0.4 for two consecutive windows (for example, w_1=0-6[..[78]] h and w$_2$=6-12 h) and $c_w(w_1, w_2)$=0.2 for two non-consecutive ones (for example, w_1=0-6[..[79]] h and w_2=12-18 h). The [..[80]]correlations of uncertainty in the day-to-day variations between different days are modelled using an exponentially

10   decaying function with a characteristic time of 7 days: $c_d(d_1, d_2) = \mathrm{e}^{\frac{|d_2 - d_1|}{7}}$.

     [..[81]][..[82]]The standard deviation of the prior uncertainty in the 30-day budgets of Net Ecosystem Exchange for a given area and 6- h window of the day is assumed to be about 75[..[83]]% of the prior estimate of this budget from C-TESSEL. In practice, it appears from our computations that the resulting value of this uncertainty decreases when the surface of the corresponding area increases. Spatial and temporal correlations between the uncertainties for the various [..[84]]6- h windows of

15   the day and areas are assumed to be negligible due to the large size of the corresponding areas and due to the differences in the processes dominating the ecosystem exchanges between daytime and night-time. [..[85]][..[86]]

2.5   [..[87]]Observation error covariance matrix **R**

The observation errors encompass instrumentation errors and errors in the observation operator $\mathcal{H}$[..[88]]. The latter combines transport model errors, representation errors, aggregation errors, errors from the boundary conditions [..[89]]and errors from the

20   emissions in the modelling domain but outside the Paris area[..[90]]. One of the main source of transport errors is linked to errors in the wind and planetary boundary layer height in the meteorological forcing of the transport model. Representation

---

[72] removed: fossil fuel

[73] removed: emission budgetsare calculated

[74] removed: a same day

[75] removed: correlations

[76] removed: . Positive correlations are used

[77] removed: :

[78] removed: and

[79] removed: and

[80] removed: day-to-day correlations of prior uncertainties in emission budgets

[81] removed: For a given area and 6-

[82] removed: window of the day, the

[83] removed:

[84] removed: 6-hour

[85] removed: It is also assumed that there is no correlation between the uncertainties in the fossil fuel

[86] removed: emissions and uncertainties in the biogenic fluxes.

[87] removed: Matrix

[88] removed: (such as

[89] removed: ,

[90] removed:).

errors are associated with the variations of CO_2 within the $2\,\text{km} \times 2\,\text{km}$ horizontal resolution grid cell of the model which encompass the peri-urban sites. They should be relatively small since there is no major CO_2 source in these grid cells. Aggregation errors are mainly associated with uncertainties in the spatial and temporal distribution of the emissions within the Paris area and 6-h windows. Aggregation errors are critical to account for in our inverse modelling system given that

5   for a given 6-h window, we control one scaling factor for the emissions over the whole Paris area.

Bréon et al. (2015) used the diagnostics of Desroziers et al. (2005) to estimate the variances of the observation error. It [..[91]]was assumed that these errors have the same statistics for any hourly gradient and that there is no correlation between errors for different hourly gradients. Their corresponding estimate of the standard deviation of the observation error was 3 [..[92]]ppm. Here, since a similar inverse modelling framework is used, and even though the revised gradient selection

10   should decrease the aggregation errors (see Sect. 2.2), we assume that our misfits between the measured and modelled gradients bear the same observation error as in this study. $\mathbf{R}$ is thus modelled as a diagonal matrix with a $(3\,\text{ppm})^2$ variance for all elements in the diagonal.

2.6   Principles of the sensitivity tests

Several tests are conducted to check the reference inversion results' sensitivity to changes in different components: (a) the prior

15   estimate of the Paris fossil fuel CO_2 emissions, (b) the spatio-temporal distribution of the fossil fuel CO_2 emissions within the Paris region and within a given 6-h window, (c) the meteorological forcing driving both the atmospheric transport model and the selection of the observations. These [..[93]]changes are representative of typical uncertainties in these components. Most of these uncertainties are, in principle, accounted for in the configuration of the prior uncertainty covariance matrix $\mathbf{B}$ and observation error covariance matrix $\mathbf{R}$, respectively. Their impact on the robustness of the inversion results should

20   be given by the posterior uncertainty covariance matrix $\mathbf{A}$. However, they may not be correctly reflected by the statistical representation that is based on Gaussian and unbiased distributions and by the rather simple models used to set up the covariance matrices. Therefore, the sensitivity tests provide a useful alternative evaluation of the robustness of the inversion results.

Regarding the prior estimate of the Paris fossil fuel CO_2 emission budgets, as an alternative to the AIRPARIF 2008 budgets,

25   we use what is called hereafter [..[94]]flat priors, i.e., prior fossil fuel CO_2 emission estimates that are not informed about month to month variations. Three sets of flat priors are built by rescaling the AIRPARIF 2008 budgets using monthly, daily or 6-h scaling factors. In the first case, the flat priors have constant monthly values, but retain the relative temporal variations of the [..[95]]6-h budgets within a month. In the second case the flat priors have constant daily values, but retain the relative temporal variations of the 6-h budgets within a day. In the third case, the flat priors have constant 6-h mean values. This change of

30   prior estimate can potentially have a large impact on the results since the system assimilates data during the afternoon

---

[91] removed: is

[92] removed:

[93] removed: tests

[94] removed: *flat priors*

[95] removed: 6-hour

only. Consequently, such a change imposes a direct constraint on two 6-h windows of the day only (the 6-12 h and the 12-18 h windows) while the constraint on the two other windows (the 0-6 h and the 18-24 h windows) relies indirectly on the description of the temporal correlations in the prior uncertainty covariance matrix $\mathbf{B}$.

For each set, different flat priors are tested by taking different values for the monthly budgets. These values cover a case of

5   relatively high emissions ($5\,\mathrm{MtCO_2 month^{-1}}$), a case of relatively low emissions ($3\,\mathrm{MtCO_2 month^{-1}}$), as well as an interme-
diate case corresponding to the annual budget from the AIRPARIF 2008 inventory ($4.3\,\mathrm{MtCO_2 month^{-1}}$). A prior estimate
based on the budgets from the IER inventory is also used for the sensitivity tests. As explained in Sect. 2.3.3, $\mathbf{H}^{\mathrm{map}}_{IER}$ is used
as alternative $\mathbf{H}^{\mathrm{map}}$ to $\mathbf{H}^{\mathrm{map}}_{AP}$, while $\mathbf{H}^{\mathrm{trans}}_{ref-MNH}$, $\boldsymbol{y}^{\mathrm{f}}_{ref-MNH}$ and $\mathbf{H}^{\mathrm{samp}}_{ref-MNH}$ are used as alternative $\mathbf{H}^{\mathrm{trans}}$, $\boldsymbol{y}^{\mathrm{f}}$ and $\mathbf{H}^{\mathrm{samp}}$ to
$\mathbf{H}^{\mathrm{trans}}_{ECM}$, $\boldsymbol{y}^{\mathrm{f}}_{ref-ECM}$, $\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$. Table 1 summarizes the acronyms and settings of the different sensitivity tests.

10   # 3   Results

Bréon et al. (2015) analysed the skill of the inversion by comparing the fit between measured and modelled CO_2 gradients, which is a first indicator of the reliability of the inverted emissions. In Tab. 2, statistical comparisons between the selected [..[96]]measured gradients and results from the initial and reference inversion, respectively, are provided. It demonstrates that both inversions strongly increase the consistency between model and measurements compared to the prior simulations (Tab. 2). Of

15   note is that the statistics of both the prior and posterior model-data misfits are smaller for the initial inversion than for the reference inversion. This is explained by the fact that the initial inversion selects gradient for which the signal (mainly the impact of the city emissions) is smaller than for the gradients that both inversions select. However, we avoid a more detailed analysis of the model-data misfit in the following.

[..[97]]Using the loose wind ranges of the initial inversion to select gradients, the root mean square of the biogenic

20   signal in these hourly gradients, averaged over the 1-yr CO2-MEGAPARIS period, is 1.1 ppm, which is, as indicated in introduction, much lower than the signal from the Paris emission. When using the tighter wind ranges of the reference inversion, the root mean square of the biogenic signal in these gradients is even smaller (0.8 ppm). Therefore, the changes in the inversion configuration proposed in our study decrease the impact of the uncertainty in the natural fluxes. This weak impact was already demonstrated by Bréon et al. (2015). We thus do not analyse further the results that correspond to

25   the control of the natural fluxes in the following.

The presentation of the results rather focuses on the estimates of the monthly fossil fuel CO_2 emission budgets from mid-2010 to mid-2011, expressed in $\mathrm{MtCO_2 month^{-1}}$ (strictly speaking $\mathrm{MtCO_2}$ per 30 days, see Sect. 2). [..[98]]The uncertainties (in terms of standard deviation) in prior and posterior estimates of the monthly emissions are based on the modelling of the $\mathbf{B}$ matrix, described in Sect. 2.4, and on the derivation of the $\mathbf{A}$ matrix (eq. (2)). However, the robustness of

30   the results is evaluated using independent knowledge on the emissions rather than using these theoretical indicators. AIRPARIF (2013) reports an estimate of $41.8\,\mathrm{MtCO_2 yr^{-1}}$ for the annual emission budget of Île-de-France in 2010. This

---

[96]removed: measurement

[97]removed: Here, the

[98]removed: AIRPARIF (2013)

16

number is used as an indicator for the evaluation of the 1-yr budget of the estimates from the inversion. According to AIRPARIF (2013), the residential and the service sector account for 43% of the Paris fossil fuel CO_2 emissions in 2010. These emissions are almost entirely linked to heating. Heating in the industry sector contributes also significantly to Paris' emissions. The heating is mostly dedicated to ambient air in the buildings and to sanitary water. Therefore we expect a large increase of the emissions from summer to winter and a high correlation between these emissions and the temperature during the cold season. An independent analysis of both daily gas use and hourly electric consumption within Île-de-France indicate a heating energy use that is highly correlated to the daily-mean temperature when this temperature is below 19° C, and essentially independent of the daily mean temperature when this temperature is above 19° C (unpublished analysis led by one co-author of this study, François-Marie Bréon). For the evaluation of the results, our emission estimates are thus compared with monthly averages of an independent measure, which we call heating degrees hereafter. It is defined as the positive difference between the daily mean temperature and 19° C (set to 0 for days when the temperature is higher than 19° C). The ratio between January and July emission estimates from the AIRPARIF 2008 inventory seem surprisingly low given these considerations. Furthermore, the prior estimates based on this inventory make use of a single emission value from November to March, which does not account for the large temperature variations during this period. Therefore, an amplified seasonal cycle of the emissions that better correlates with heating degrees is expected through the atmospheric inversion.

[..[99]][..[100]]

3.1   Initial inversion using loose wind direction criteria for the gradients selection

Figure 3a shows prior and posterior estimates of the Paris emissions from the initial inversion (experiment *ini*, Tab.1). Monthly mean heating degrees for the centre of Paris, derived from the temperature given by the ECMWF's operational analysis, are also shown on this figure. The posterior estimates are lower than the prior ones for all months. The inversion decreases the annual emissions from $51.9\,\mathrm{MtCO_2yr^{-1}}$ to $37.4\pm2.1\,\mathrm{MtCO_2yr^{-1}}$. This number is smaller but closer to the AIRPARIF inventory 2010 used for evaluation ($41.8\,\mathrm{MtCO_2yr^{-1}}$, see Tab. 1).

Posterior fluxes are lowest for August ($1.6\,\mathrm{MtCO_2month^{-1}}$), increase steadily until peaking in February ($4.6\,\mathrm{MtCO_2month^{-1}}$), drop strongly in March ($2.8\,\mathrm{MtCO_2month^{-1}}$), and vary between 2 and $3\,\mathrm{MtCO_2month^{-1}}$ from April to July. Compared to the prior estimate, the inversion yields larger emissions in winter and increases the amplitude of the seasonal variations. For the period analysed, monthly mean heating degrees were highest in December (Fig. 3a). Hence, fossil fuel CO_2 emissions are expected to be the highest in December rather than in February. There is no clear correlation between monthly heating degrees and emission estimates during the November-March period.

The number of assimilated gradients varies considerably from one month to another, which influences the month-to-month variations of the inverted emissions. For instance, 163 observations are assimilated in March, compared with only 34 in Novem-

---

[99]removed: The uncertainties (in terms of standard deviation) in prior and posterior estimates of the monthly emissions are based on the modelling of the **B** matrix, described in Sect. 2.4, and on the derivation of the **A** matrix (Eq.

[100]removed:). This section provides these uncertainties estimates and the corresponding uncertainty reduction (i.e., the relative difference between the posterior and prior uncertainties) but it does not analyse them. The reliability in these uncertainty estimates is actually commented in Sect. 4 in the light of the assessment of the robustness of the inversions based on comparisons to AIRPARIF 2010 and to the monthly mean heating degrees.

ber. Figure 3a also shows that, for most months, the numbers of selected gradients are not apportioned equally amongst the NE and SW wind directions. For instance, there are no NE gradients to constrain August emissions, while less than half of the gradients in March are SW gradients. The different upwind conditions for NE and SW gradients could play a role in the month-to-month variability of the inverted emissions, in case the gradient approach does not remove entirely the influence of remote fluxes.

We investigate the impact of assimilating these two different gradient types on monthly fossil fuel CO_2 flux estimates by conducting inversions based on NE gradients, SW gradients, or even GIF-MON, GIF-GON, MON-GIF or GON-GIF only (Fig. 3b). The difference in inverted December emissions when assimilating only SW gradients compared to assimilating only NE gradients is large, even though a large number of both types of gradients is available during this month. Compared to the prior estimate, the inversion of SW gradients increases the December emissions. The opposite, however, is true for the inversion of NE gradients. This behaviour seems to be driven by both the assimilation of GIF-MON and GIF-GON gradients. An analysis of the average temperature in Paris (not shown) shows lower temperatures for NE wind conditions than for SW wind conditions. The heating emissions in Paris should thus be higher for NE wind conditions. Therefore, the temperature variations cannot explain the differences in December emissions between assimilating SW gradients and assimilating NE gradients.

The differences between the results when using NE gradients or SW gradients are not as large for other months as in December. However, they can still be significant, e.g., April (Fig. 3b). These differences cannot be explained by a lack of data for a given type of gradient, except in August, when there are no NE gradients. For January and February, differences of 1 to $1.5 \, \mathrm{MtCO_2}$ (i.e., about 35 % relative differences) are obtained between inversions assimilating only MON-GIF compared to assimilating only GON-GIF, although both are SW gradients and gather more than 40 observations during each month. This large mismatch between the different inversions when using different data subsets undermines the reliability of the inversion results, in particular in December. The seasonal profile of the retrieved emission when assimilating only SW gradients is far better correlated with heating degrees than when the inversion uses both SW and NE gradients, as emissions reach their maximum in December. Results seem nearly as sensible when using MON-GIF or GON-GIF as when using all SW gradients.

Figures 4a and 4c illustrate that, even if we discard GON-MON and increase the threshold of the wind speed, there are episodes when measured gradients show negative values. They, however, should be positive owing to the city's emissions. Most of the negative gradients are found when the wind is from NE, such as in December (Fig. 4a), suggesting that such gradients may not represent the emissions of the entire city.

3.2 Reference inversion

The estimates of the monthly Paris fossil fuel CO_2 emissions by the reference inversion (experiment *ref*, Tab.1) are given in Fig. 3c. All negative observed gradients outlined above were obtained at the limit of the wind direction range proposed by Bréon et al. (2015). As illustrated for December and May by comparing Fig. 4a to Fig. 4b and Fig. 4c to Fig. 4d, respectively, the reference inversion, which uses a stricter range of wind directions (Sect. 2.2), removes the negative gradients. Despite the

loss of 65 % of the data compared to the initial inversion, the reference inversion still predicts a large uncertainty reduction for monthly fossil fuel CO_2 emission estimates, from 9 % in November to 50 % in October.

The Paris monthly fossil fuel CO_2 emission estimates from the reference inversion correlate well with monthly heating degrees ($r^2 = 0.67$ for the whole period, $r^2 = 0.45$ for November–February), which was not the case of the initial inversion ($r^2 = 0.54$ for the whole period, $r^2 = 0.07$ for November–February). In general, the reference inversion decreases the fossil fuel CO_2 budget from the prior estimate, except in December, which becomes the peak of emissions (Fig. 3c). The emissions decrease from February to March, which does not correspond to a relative change in heating degrees, is significantly smaller in the reference than in the initial inversion. The seasonal variations of the reference inversion are strongly improved compared to initial inversions. The annual budget from the reference inversion ($40.9\,MtCO_2$) is close to that from AIRPARIF 2010 ($41.8\,MtCO_2$, see Tab. 1).

The stricter gradient selection further leads to a much better agreement between emission estimates when using different subsets of gradients (compare Fig. 3d with Fig. 3b). Although significant differences in December and April emissions are still apparent between using NE gradients or using SW gradients, and in January and February emissions between using GON-GIF or MON-GIF, these differences are smaller than in the initial inversion. Now, even when assimilating only NE gradients, the four months with largest inverted emissions correspond to the four coldest ones with the highest heating degrees of the year (November to February), though the assimilation of NE gradients still leads to smaller emissions in December than in November, January and February. One may argue that the improvements of the reference over the initial inversion reflect the assimilation of a smaller dataset, and therefore are due to smaller corrections. However, results from the reference and initial inversion are closer to each other if only SW gradients are assimilated than if only NE gradients are assimilated. Highest correlation with the heating degrees are obtained when only SW gradients are assimilated.

The theoretical posterior uncertainties in the monthly budgets of the emissions are generally much lower than the prior uncertainties (with more than 30% uncertainty reduction for most of the months) in the reference inversion. However, even though our analysis above gives more confidence in the results from the reference inversion than in the results from the initial inversion, the inverse modelling system diagnoses smaller posterior uncertainties in the latter than in the former.

Results from the [..[101]]inversion using a 2-h lag between the upwind and the corresponding downwind measurement are shown in Fig.5. The [..[102]]corrections applied to the prior estimate from AIRPARIF 2008 by this inversion are qualitatively consistent with those of the reference inversion. The amplitudes of the corrections, however, are much smaller as the large decrease of the number of assimilated gradients in this inversion compared to the reference configuration (only 4% of available measurement [..[103]]are used by reducing the time window of eligible upwind or downwind measurements; see Sec.2.2) [..[104]]clearly limits the weight of the observational constraint. [..[105]]

---

[101]removed: Lagrangian

[102]removed: further and

[103]removed: data

[104]removed: compared to the reference configuration

[105]removed: The results of this inversion is between that from the reference inversion and the prior estimates from AIRPARIF 2008 (often closer to AIRPARIF 2008 than to the reference inversion), without any qualitative improvement.

3.3 Sensitivity Tests

3.3.1 Sensitivity of the Paris emission budgets to prior estimates x_b

Monthly fossil fuel CO_2 emissions estimates using flat priors for x_b (all experiments *FLAT_* in Tab. 1) are reported in Fig. 6. Although differences in prior monthly budgets between the FLAT_3.0, FLAT_4.3 and FLAT_5.0 experiments amount to

5    $2\,\mathrm{MtCO_2 month^{-1}}$, posterior differences between their monthly budgets are generally much lower. In addition, the posterior monthly emissions when using flat priors are comparable to those from the reference inversion—with a very similar month-to-month variation. The differences between posterior monthly emissions from the FLAT_mM and FLAT_mD (m=3.0, 4.3 or 5.0) inversions and the reference inversion are generally smaller than $1\,\mathrm{MtCO_2 month^{-1}}$, except during September, November, May and July when very few (4 to 24) gradients are assimilated (Fig. 6a and Fig. 6b). Larger differences are obtained between

10   the reference inversion and FLAT_m[..[106]]H (m=3.0, 4.3 or 5.0) inversions, which use prior estimates that are flat at [..[107]]6- h scale (Fig. 6[..[108]]c). The FLAT_m[..[109]]H experiments yield larger posterior monthly budgets than the FLAT_mD and FLAT_m[..[110]]M experiments.

Posterior annual budgets from FLAT_mM and FLAT_mD inversions range between 33 and $45.3\,\mathrm{MtCO_2 yr^{-1}}$, encompassing the budgets from the reference inversion and AIRPARIF 2010 (Tab. 1). In particular, the inverted annual budget from

15   FLAT_4.3M and FLAT_4.3D, whose prior estimate [..[111]]have the same annual budget as the prior estimate from the reference inversion, is equal to $41.1\,\mathrm{MtCO_2}$. This is very close to the annual budgets from the reference inversion and AIRPARIF 2010. However, the annual emissions budgets from the FLAT_m[..[112]]H inversions range from [..[113]]33 to $52.2\,\mathrm{MtCO_2}$, which is biased compared to both the reference inversion and AIRPARIF 2010.

3.3.2 Sensitivity of the Paris emissions budgets to [..[114]]the mapping and variations at hourly scale

20   Figure 7 compares the estimates from the reference inversion, which uses $\mathbf{H}^{\mathrm{map}} = \mathbf{H}_{AP}^{\mathrm{map}}$, to the estimates from the sensitivity test with $\mathbf{H}^{\mathrm{map}} = \mathbf{H}_{IER}^{\mathrm{map}}$ (INV_mapIER and INV_IER, see Tab. 1). Thus, this experiment also includes results when using the IER inventory to build both the 6- h [..[115]]budgets in x_b and $\mathbf{H}^{\mathrm{map}} = \mathbf{H}_{IER}^{\mathrm{map}}$ (INV_IER). It provides estimates when the inversion relies entirely on the IER inventory to define these parameters and ignores the existence of the AIRPARIF inventory. This situation is similar to that in cities, where no local inventory is available. We have less confidence in the posterior estimates

25   from such an inversion, since the IER inventory does not rely on the same amount of local data as the AIRPARIF inventory.

---

[106]removed: D and FLAT_m
[107]removed: the daily or 6-hour
[108]removed: b and Fig. 6
[109]removed: M
[110]removed: H
[111]removed: has
[112]removed: D and FLAT_m
[113]removed: 36.9
[114]removed: $\mathbf{H}^{\mathrm{map}}$
[115]removed: mean

INV_mapIER regularly predicts lower monthly budgets than the reference inversion, except in June, July, September and November. The corresponding differences are relatively small and do not exceed $0.5 \, \mathrm{MtCO_2 month^{-1}}$, except in January and February. Similar to the reference inversion, INV_mapIER predicts highest emissions in December. However, its estimates for January and February fluxes are particularly low, e.g., January estimates ($3.9 \, \mathrm{MtCO_2 month^{-1}}$) roughly equal that for May ($3.9 \, \mathrm{MtCO_2 month^{-1}}$) or October ($3.8 \, \mathrm{MtCO_2 month^{-1}}$), and are smaller than that for September ($4.1 \, \mathrm{MtCO_2 month^{-1}}$). This results in an annual budget of $39 \, \mathrm{MtCO_2 yr^{-1}}$ that is still closer to the one from AIRPARIF 2010 than to the one from AIRPARIF 2008 (Tab. 1).

The monthly prior emissions from AIRPARIF 2008 and IER differ substantially. In particular, from November to May, the IER inventory estimates up to $3 \, \mathrm{MtCO_2 month^{-1}}$, (approximately 40%) higher fossil fuel CO_2 emissions for the Paris region than AIRPARIF 2008. At the annual scale, estimates differ by $8.2 \, \mathrm{MtCO_2 yr^{-1}}$ (Tab. 1). The differences between the two inventories are due to both the differences in the emission model and the driving activity data used. The two inventories correspond to two different years (2008 versus 2005). However, this hardly explains the amplitude of the difference between the two inventories by itself. The decrease of the total emission in France between 2005 and 2008 was approximately 5 % (see, e.g., CITEPA, 2015). Here, the difference in total emissions in Île-de-France between the IER and AIRPARIF 2008 inventory, however, is about 14%. Results of the inversion using IER for both the prior emission budgets and the emissions' spatial distribution (INV_IER) reflect these large prior discrepancies. Indeed, monthly and annual budgets of Paris' fossil fuel CO_2 emissions estimated by INV_IER are larger than that from the reference inversion and from INV_mapIER (Fig. 7). The differences in posterior February emissions from IER_INV and the reference inversion exceed $2 \, \mathrm{MtCO_2 month^{-1}}$. The discrepancies are even larger, when comparing the monthly emission estimates from INV_mapIER and IER_INV, since the change of $\mathbf{H}^{\mathrm{map}}$ from $\mathbf{H}^{\mathrm{map}}_{AP}$ to $\mathbf{H}^{\mathrm{map}}_{IER}$ has a tendency to decrease the posterior emission estimates.

The IER inventory indicates higher emissions in March than in November. Posterior estimates from INV_IER still indicate that the highest emissions are in November-February. Due to a residual influence from the IER prior estimate, INV_IER predicts highest emissions in February. The December emission estimate is close to February emission estimate, and is the second highest 1-month mean estimate from INV_IER. Finally, the annual posterior emission from INV_IER is closer to that from AIRPARIF 2010 than to that from the 2008 inventory, despite the far higher prior annual estimate from the IER inventory (Tab. 1).

3.3.3 Sensitivity to $\mathbf{H}^{\mathrm{samp}}$, $\mathbf{H}^{\mathrm{trans}}$ and y^{f}

INV_MNH is compared to the reference inversion to analyse the impact of using Meso-NH/TEB instead of ECMWF as meteorological simulation for both the gradient selection in the observation vector ($\mathbf{H}^{\mathrm{samp}} = \mathbf{H}^{\mathrm{samp}}_{ref-MNH}$ versus $\mathbf{H}^{\mathrm{samp}} = \mathbf{H}^{\mathrm{samp}}_{ref-ECM}$) and the forcing of CHIMERE ($\mathbf{H}^{\mathrm{trans}} = \mathbf{H}^{\mathrm{trans}}_{MNH}$ versus $\mathbf{H}^{\mathrm{trans}} = \mathbf{H}^{\mathrm{trans}}_{ECM}$ and $y^{\mathrm{f}} = y^{\mathrm{f}}_{ref-MNH}$ versus $y^{\mathrm{f}} = y^{\mathrm{f}}_{ref-ECM}$). Meso-NH/TEB data are available up to June only which explains that the analyses here are restrained to the period August 2010 to June 2011.

The time series of the gradients that are selected for the assimilation using $\mathbf{H}^{\mathrm{samp}}_{ref-MNH}$ and $\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$, respectively, are shown in Fig. 2 (and Fig. A1–A2). The significant differences in selected gradients apparent in Fig. 2 (and Fig. A1–A2)) are

driven by small differences in simulated wind fields between ECMWF and Meso-NH/TEB. Small differences in wind direction and speed are often sufficient to cross the thresholds defining the gradient selection (Fig. A1–A2). This differences result in a significantly different set of assimilated gradients and in a different apportionment according to prevailing NE or SW wind directions (Fig. 8).

5   Despite this, Fig. 8 reports similarities in inverted monthly emissions from INV_MNH and the reference inversion. Differences in monthly posterior emission estimates are less than $0.5\,\mathrm{MtCO_2month^{-1}}$ when assimilating all selected gradients (Fig. 8). The four highest emitting months are still November to February for INV_MNH. However, larger differences to the reference inversion estimates are found for December and May, resulting in the loss of the peak in December and in an unexpected peak in May in INV_MNH (Fig. 8a). This disagreement is related to the assimilation of NE gradients. As shown in

10   Fig. 8b, emissions estimates from INV_MNH and the reference inversion are very similar when only SW gradients are assimilated. By contrast, large differences are obtained in December and May when only NE gradients are assimilated (Fig. 8c). The larger fraction of selected NE gradients compared to selected SW gradients when using Meson-NH/TEB instead of ECMWF could explain the loss of the emission peak in December. There is no peak in December when using either Meso-NH/TEB, or ECMWF and NE gradients only. Nevertheless, when assimilating SW gradients, the consistency between INV_MNH and the

15   reference inversion is surprising, given the significantly different SW gradient selection.

4   Discussion and Conclusions

4.1   Summary and general analysis of the results

We have analysed estimates of monthly mean fossil fuel CO_2 emissions from the Paris urban area from August 2010 to July 2011 using continuous CO_2 measurements from three stations and a city-scale atmospheric inverse modelling framework

20   derived from Bréon et al. (2015). The inversion [..[116]]modelling is based on a mesoscale configuration of CHIMERE, on the AIRPARIF high-resolution CO_2 emission inventory for 2008, on the C-TESSEL simulation for the biogenic fluxes in Northern France, and on the principle of constraining the [..[117]]6- h city-scale budget of the emissions using cross-city CO_2 gradients. [..[118]]As demonstrated by the analysis of the inversion results, this study has critically improved configuration of Bréon et al. (2015) by (i) discarding GON-MON gradients since that are not [..[119]]related to the [..[120]]emission of the entire city,

25   and (ii) by using stricter criteria on the wind direction and wind speed for the selection of gradients.

   The analysis suggests an improvement of the city's seasonal to annual emission budget from the reference inversion compared to the prior estimate that is based on the AIRPARIF 2008 inventory. The inversion derives an annual emission budget (for August 2010–July 2011) of $40.9\,\mathrm{MtCO_2yr^{-1}}$, which is closer to the independent estimate from the AIRPARIF 2010 inventory ($41.8\,\mathrm{MtCO_2yr^{-1}}$) than to the prior estimate ($51.9\,\mathrm{MtCO_2yr^{-1}}$). Although the reported estimate from the AIRPARIF 2010

---

[116]removed: modeling

[117]removed: 6-hour mean

[118]removed: This

[119]removed: evidently

[120]removed: whole cityemissions

inventory does not exactly correspond to the mid-2010 to mid-2011 period, changes between the 2008 and 2010 inventories reflect improvements in the inventory model and actual changes of the Paris emissions. Therefore, the fact that the corrections applied by the inversion to the prior estimate from AIRPARIF 2008 are consistent with the differences between the AIRPARIF 2008 and 2010 inventories gives confidence in the inversion.

5   The seasonal variations of the monthly inverted emissions also appear more realistic than that of the prior emission estimates. The seasonal amplitude of the emissions revealed by the reference inversion is higher than that of the prior estimate of the emission, derived from the 5 typical months of the AIPARIF 2008 inventory. This increase in amplitude makes sense, given that a large fraction (43 % according to the AIRPARIF 2008 inventory) of the Paris emissions are due to domestic and commercial heating. It is supported by the fact that the seasonal variations in the AIRPARIF 2010 inventory are higher than that derived

10 from the AIRPARIF 2008 inventory. The inverted seasonal cycle of the emissions correlates well (r^2 =0.45) with the heating degrees in fall-winter (November–February). The four months with highest inverted emissions correspond to the four coldest months (November to February)–with a peak in both the emissions and the heating degrees in December. By contrast, the prior estimate of the emissions derived from AIRPARIF 2008 does not differentiate monthly budgets from November to March.

     The sensitivity tests indicate that the uncertainties assigned to the prior estimates of the 6- h mean emissions, to the spatio-

15 temporal distribution of the emissions within the Paris area and 6- h windows, and to the meteorological simulations (for the cross-city gradient selection and for the forcing of CHIMERE) have a moderate impact on the monthly mean emission estimates once the inversion is driven by the most stringent selection of the measurements. This weak sensitivity of the inverted emissions to the uncertainties assigned to the inverse [..[121]]modelling components is important for the credibility of the inversion approach in view to apply this approach as an independent method to verify inventories. Here, the inverted emission

20 budgets are sensitive to each of the above-mentioned components. However, even though we assimilate a relatively small number of data, this sensitivity is generally much smaller than the differences between inverted and prior estimates at monthly to annual scale. Furthermore, the plausible seasonal variations of the emissions revealed by the reference inversion is robust to most sensitivity tests.

     The inversions generally return smaller emissions than the prior estimates. This is even the case when using a prior estimate

25 that is flat at the monthly scale only, and that has an annual emission budget of $36 \, \mathrm{MtCO_2 yr^{-1}}$, i.e., a budget that is smaller than that from AIRPARIF 2010. The inversion decreases the annual emission budget when using the diurnal cycle of the emissions from AIRPARIF 2008 as prior estimate of the 6- h mean emissions. In contrast, the annual emission budget is increased when a flat diurnal cycle and a prior estimate of the annual emission budget that is smaller than that from AIRPARIF 2010 (i.e. of $36 \, \mathrm{MtCO_2 yr^{-1}}$) is used. This can reveal an error in the mean diurnal cycle of the emissions from

30 AIRPARIF 2008, which the inversion could not correct for since data are assimilated during afternoon only. Moreover, we define the uncertainties in the prior emission estimates in terms of relative rather than absolute uncertainty. Consequently, using the diurnal cycle of the emissions from AIRPARIF 2008 in the prior estimate of the 6- h mean emissions and higher (smaller) prior emissions at monthly to annual scale leads to higher (smaller) prior uncertainties, and thus to a stronger (weaker) constraint from the atmospheric measurements, resulting in a stronger (weaker) decrease of the emissions. One could

---

[121]removed: modeling

23

argue that this artificially helps getting a robust convergence of the sensitivity tests using different prior estimates and it likely plays a role at the annual scale. This could be problematic, since having a fixed value for the relative uncertainty in the prior estimates is not suitable when these estimates become very small. However, for some months, the convergence between inversions utilising different flat priors is obtained by both positive and negative corrections. This is the case in January and

5 February 2011 for FLAT_mM experiments (Fig.6a). The convergence can also be obtained with positive corrections that are larger when prior uncertainties are smaller, e.g., in December 2010 for FLAT_mM experiments (Fig.6a). [..[122]]Figure 6c gives several examples where the monthly budget of the prior estimate that are flat at the [..[123]]6- h scale [..[124]]determines if the corresponding corrections are positive or negative. This figure also illustrates the fact that the amplitude of the correction to the monthly estimates [..[125]]is not highly correlated with the corresponding prior uncertainty. Furthermore, the

10 fact that higher prior emission estimates are assigned higher prior uncertainties cannot explain the level of convergence of the sensitivity tests. In particular, it can not explain the robustness of the retrieved seasonal cycle of emissions when using flat priors. It neither explains the fact that the annual budget from INV_IER is closer to AIRPARIF 2010 than to AIRPARIF 2008. INV_MNH selected a significantly different set of gradients. However, it still constrains the inverted emissions towards the same levels of emissions as in the reference inversion (typically differences in monthly emissions are <5 %).

15   The improvement of the reference inversion compared to the initial inversion demonstrates the need for a narrow definition of the wind direction ranges, and more generally the need for a very careful selection of CO_2 data. This reveals the asset of following as much as possible the concept of assimilating cross-city CO_2 gradients to control the emissions at the whole city scale, and to filter out the poorly modelled influence of fluxes outside the Paris urban area. The assimilation of gradients cannot perfectly cancel this influence because firstly one cannot set up the inversion system to ensure that the selected

20 gradients correspond to the concentration variations of air masses that travel from the upwind to the downwind sites (at least due to uncertainties in the atmospheric transport) and secondly because the signal from fluxes outside the Paris varies during such a transport (due to atmospheric diffusion). However, results from Bréon et al. (2015) and from this study demonstrate that the assimilation of gradients succeeds in decreasing this signal. These studies also show that such a decrease strengthens the inversion results by limiting the problem of the uncertainties in the remote fluxes for

25 regional inversions (which is particularly critical in the Paris area as shown by Bréon et al., 2015) and the problem of the uncertainties in natural fluxes for urban CO_2 emission inversions. The positive insights from the evaluation of our results also strengthens the confidence in this relatively simple concept to estimate monthly budgets of the city emissions, even if it relies on the assimilation of a relatively small amount of data.

4.2   Problems to be solved

30 The different inversion tests still raise concerns for the inversion of the cities' monthly emission budgets. We expected that cross-city gradients would be weakly sensitive to the uncertainties in the distribution of the emissions within the Paris region

---

[122]removed: Tests with prior estimates

[123]removed: daily and

[124]removed: (Fig. 6b,c) give a lot of examples when the sign of the corrections

[125]removed: depends on the value of the flat prior and where the amplitude of this correction is not obviously driven by the

and the 6-h windows, which explains why we control, for a given 6-h window, a single scaling factor for the emissions of the entire urban area. The inversion results, however, are significantly affected by changes in the emission distribution. [..[126]]This does not necessarily question the control of a single scaling factor for the emissions of the whole urban area since reasonable results are obtained using the emission distribution from AIRPARIF but it reveals the need to rely on robust,

5   high resolution emission maps such as those produced by local agencies like AIRPARIF. However, many cities do not have such local inventories.

Bréon et al. (2015) have shown that the selection of afternoon data [..[127]]provides little constraint on night-time emissions. This is problematic since the diurnal cycle is highly uncertain in inventories. [..[128]]The differences between the results from FLAT_mD and FLAT_mH [..[129]]indicate that the poor representation of the diurnal cycle in FLAT_m[..[130]]H has a large

10 impact on the inverted monthly emissions. [..[131]]As highlighted above, the inversions, based on the diurnal cycle from AIRPARIF 2008, generally tend to decrease the prior estimates which can also be viewed as an impact from errors in this diurnal cycle. New approaches and techniques are needed to provide a direct constraint on the night-time emissions or to better extrapolate the information from daytime to night-time data to solve for this problem. The poor representation of the day-to-day variations [..[132]]in the flat prior of FLAT_mD does not seem to impact the results from this inversion, which are

15 close to that of FLAT_mM. Even though there is a large number of days, even sometimes weeks, during which no gradients are [..[133]]assimilated, the inversion does not strongly rely on the prior day-to-day variations within the months to correct the monthly mean emissions budgets. However, there is a critical lack of data, which is primarily due to the small number of sites available for this study, and thus to the relatively small wind sectors by which we select cross-city gradients. [..[134]][..[135]]This lack of data hinders the results of all inversions for specific months such as September, November, April, May and July,

20 when less than 30 1-hour mean gradients are assimilated. The month-to-month variability is thus often driven by the variability of the data availability. Results at the monthly scale are thus not systematically consistent with the different sensitivity tests. Monthly estimates can be weakened by missing or over-weighting high variations in the emissions over short time periods (e.g., due to a cold event). One can hope that this limitation could be overcome by an expansion of the observation network with stations all around the Paris urban area, which could ensure a continuous monitoring of the cross-city CO_2 gradients.

25 In December, the number of assimilated data is relatively high for both the reference inversion and INV_MNH. However, while the inversions increase the emissions compared to AIRPARIF 2008 during December when using ECMWF data and all

---

[126]removed: It

[127]removed: provide

[128]removed: However, the similarity between

[129]removed: results seems to

[130]removed: D does not much impact

[131]removed: The poor representation of the day-to-day variations in both flat priors is the main explanation for the positive bias in the annual estimates between these tests and the reference inversion. The dependence of the inversion to the prior

[132]removed: within the months can be explained by the

[133]removed: selected for the inversion . This

[134]removed: Therefore, one can hope that this limitation could be overcome by an expansion of the observation network with stations all around the Paris urban area. This could ensure a continuous monitoring of the cross-city

[135]removed: gradients. The

gradients (SW and NE gradients), this is not the case when assimilating subsets of the cross-city gradients only, or when using Meso-NH/TB. Consequently, there is no peak of inverted emission estimates in December. Neither is this a robust feature of the reference inversion. The absence of an emission peak in December is associated with the assimilation of NE gradients (i.e., due to the assimilation of NE gradients only, or, to the use of the Meso-NH/TEB meteorology which selects a larger fraction of NE gradients than its ECMWF counterpart).

More generally, the assimilation of NE gradients seems to raise concerns while more satisfying results are obtained when using SW gradients. This applies also to the initial inversion, for which the NE direction corresponds to wider wind direction ranges. Thus, the problem cannot be related to a very specific source NE of Paris. When the wind blows from NE, the signature of emissions from remote, highly urbanised and industrialised areas (North-Eastern France, Benelux and Western Germany) should impact the CO_2 fields in the Paris area. On the opposite, the regions between the Atlantic Ocean and Paris are mostly rural. While the computation of gradients is an efficient way of limiting the signatures of the fluxes outside the Paris area on assimilated data, and while it effectively reduces these signatures to a small component in the simulated gradients, it does not ensure a total removal of such signatures in the measurements which may bear a more spatial heterogeneity than the modelling framework. The large-scale signature of the remote natural fluxes from SW may be more easily modelled or filtered out by the computation of gradients in the Paris area than the [..[136]]signature of emissions from NE. This could explain why the assimilation of NE gradients is more problematic than that of SW gradients. This could reveal another limitation in assimilating cross-city gradients. The high temporal variability of the ratio of assimilated NE gradients to SW gradients may be problematic for the monitoring of the month-to-month variability of the city emissions. Similarly, the small biogenic signal in the simulated gradients may be due to the use of an ecosystem model with moderate horizontal resolution. Measured gradients might be impacted by urban ecosystems that cannot be represented with this model. Due to the high density and compactness of the Paris urban area, we can assume that such urban ecosystems should have a low impact on the inversion of Paris emissions. This should be further investigated based on urban ecosystem modelling and monitoring (Nordbo et al., 2012).

The last major issue is the limited confidence in the posterior uncertainties computed by the inversion system. We purposely avoided analysing them in details in Sect. [..[137]]3. They provide qualitative insights on the behaviour of the inversion, i.e., posterior uncertainties remain close to the prior ones for night-time emissions, which are [..[138]]poorly constrained by using only afternoon CO_2 data (Bréon et al., 2015). The posterior uncertainties also vary as a function of the number of assimilated data. The different estimates from the sensitivity tests generally lie in the 68[..[139]]% confidence (1-σ) interval of the reference inversion. However, the posterior uncertainties look generally very low for specific months, despite the lack of confidence in the specific monthly estimates as discussed above, and despite the very limited number of assimilated data. During February and March, the posterior uncertainties from the reference inversion are lower than $0.69\,\mathrm{MtCO_2yr^{-1}}$. The large emission decrease

---

[136]removed: one

[137]removed:

[138]removed: not strongly

[139]removed:

of $1.32\,\mathrm{MtCO_2yr^{-1}}$ from February to March [..[140]]is surprising. The relative difference between the posterior and prior uncertainties when moving from the initial inversion to the more reliable reference inversion demonstrates how misleading the interpretation of theoretical uncertainties can be when [..[141]]several mathematical assumptions in the inversion are not met in practice. However, even though the configuration is [..[142]]far from perfect, the misfits between posterior estimates

5 and observations are still smaller than between prior estimates and observations. This gives a stronger confidence in posterior emission estimates than in the posterior uncertainties of these emissions. Sensitivity tests with the analysis of the posterior estimates only were conducted to give a better picture of the strength of the measurement constraint.

4.3 Perspectives

Despite these concerns, the results from this study are promising and several methodological improvements were found. The

10 [..[143]]inversion test of assimilating spatio-temporal gradients accounts for the time air parcels need to pass from the upwind site over the urban area to the downwind site. Such gradients should bear a smaller signal from fluxes outside the urban area than spatial gradients, which should help isolating this signal from the city emissions. The lack of data, however, prevented this inversion from significantly departing from its prior emission estimate. Such a strategy would be more appropriate if a larger amount of data was available, but it is impractical for our limited network: it exacerbates the loss of data from already

15 strict gradient selection criteria and degrades the overall emission retrieval compared to [..[144]]the reference inversion.

For the same reason, it would be inappropriate with our limited network to narrow the wind direction ranges to select gradients to less than $\pm 15^o$ of the transect between the downwind and upwind sites, even though, in principle, it would strengthen the decrease of the signature from fluxes outside the urban area.

The expansion of the network, [..[145]]in particular a full encirclement of the city with at least 8 sites (given that the wind

20 ranges for the selection of gradients between one upwind site and one downwind site cover 30 degrees in this study) should strengthen the results and could allow applying such new techniques that result in a stricter gradient selection. However, relying on such [..[146]]a measurement expansion may not be sufficient. Exploiting more information from the available dataset without violating or undermining our assumptions on the selection of cross-city [..[147]]gradients is a requirement to strengthen the observational constraint of the inversion. The Paris observation network has been set back since September 2014 in the

25 framework of the CARBOCOUNT-CITY and LE CO2 PARISIEN projects. Both projects aim to deploy more measurement sites than the CO2-MEGAPARIS project. However, relying on such a measurement network expansion may not be sufficient. New methods should be developed to exploit urban measurements [..[148]](Wu et al., 2016) which would allow to solve for the

---

[140]removed: seems

[141]removed: the configuration of the inversion is not perfect

[142]removed: not

[143]removed: Lagrangian inversion test

[144]removed: this

[145]removed: and

[146]removed: an increase of the measurement network

[147]removed: upwind-downwind

[148]removed: (Bréon et al., 2015) which also

27

spatial distribution of the emissions, which does not seem possible with the current monitoring network of peri-urban sites. This in turn [..[149]]could help assimilating data that do not necessarily bear the signature of the emissions from a large part of the city. Finally, developing methods to exploit [..[150]]morning, evening and night-time data would be necessary to constrain night-time fluxes. This is not necessary to improve the knowledge on the emissions based on atmospheric inversion, but this is

5   necessary to develop accurate tools for the operational monitoring and verification of the emissions based on this approach.

Even though it applies to the specific case of monitoring the CO_2 emissions from Paris, this study demonstrates the potential of an approach which can be adapted to a wide range of cities. [..[151]]The urban surrounding, spread, size, topography and meteorology of some cities increase the difficulty for catching cross-city [..[152]]gradients, and different strategies may be more adapted for such cases. The atmospheric inversion of the city emissions is still an emerging activity, but the present results

10   already raise some confidence in this concept, especially since many other resources (combining [..[153]]atmospheric CO_2 inversions with air quality monitoring, the development of new measurement types) could help overcoming the remaining challenges.

---

[149]removed: ,

[150]removed: late morningor evening measurements

[151]removed: It is currently tested in the city of Recife in Brazil.

[152]removed: downwind–upwind

[153]removed: atmospheric

[154]removed: is

[155]removed: Thales

References

AIRPARIF: Inventaire des émissions en Île-de-France. Résultats - année 2008, Tech. rep., AIRPARIF Surveillance de la Qualité de l'Air en Île-de-France, Paris, France, last access: 26 March 2015, 2012.

AIRPARIF: Bilan des émissions de polluants atmosphéríques et de gaz à effet de serre en Île-de-France pour l'année 2010 et historique 2000/2005. Méthodologies et résultats, Tech. rep., AIRPARIF Surveillance de la Qualité de l'Air en Île-de-France, Paris, France, last access: 26 March 2015, 2013.

Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramont, 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, Atm. Chem. Phys., 16, 6735–6756, doi:doi:10.5194/acp-16-6735-2016, 2016.

Boussetta, S., Balsamo, G., Beljaars, A., Panareda, A.-A., Calvet, J.-C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzaro, M., Papale, D., and van der Werf, G.: Natural land carbon dioxide exchanges in the ECMWF integrated forecasting system: Implementation and offline validation, J. Geophys. Res. Atmos, 118, 5923–5946, doi:10.1002/jgrd.50488, 2013.

Bréon, F.-M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Rémy, 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, Atm. Chem. Phys., 15, 1707–1724, doi:10.5194/acp-15-1707-2015, 2015.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T., and Ciais, P.: A European summertime CO_2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, J. Geophys. Res. Atmos., 116, doi:10.1029/2011JD016202, 2011.

Chevallier, F., Cias, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis, J. Geophys. Res, 115, doi:10.1029/2010JD013887, 2010.

CITEPA: Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France–Séries sectorielles et analyses étendues, Tech. rep., CITEPA Centre Interprofessionnel Technique d'Etudes de la Pollution Atmosphérique, Paris, France, last access: 9 September 2015, 2015.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis error statistics in observation space, Q. J. R. Meteor. Soc., 131, 3385–3396, doi:10.1256/qj.05.108, 2005.

Duren, R. M. and Miller, C. E.: COMMENTARY: Measuring the carbon emissions of megacities, Nat. Clim. Change, 2, 560–562, 2012.

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe – Part 1: mapping the atmospheric CO_2 signals, Atmos. Chem. and Phys., 7, 3461–3479, doi:10.5194/acp-7-3461-2007, http://www.atmos-chem-phys.net/7/3461/2007/, 2007.

Gibert, F., Schmidt, M., Cuesta, J., Ciais, P., Ramonet, M., Xueref, I., Larmanou, E., and Flamant, P. H.: Retrieval of average CO_2 fluxes by combining in situ CO_2 measurements and backscatter lidar information, J. Geophys. Res. Atmos., 112, doi:10.1029/2006JD008190, http://dx.doi.org/10.1029/2006JD008190, d10301, 2007.

Henne, S., Brunner, D., Oney, B., Leuenberger, M., Eugster, W., Bamberger, I., Meinhardt, F., Steinbacher, M., and Emmenegger, L.: Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling, Atmos. Chem. Phys., 16, 3683–3710, doi:10.5194/acp-16-3683-2016, http://www.atmos-chem-phys.net/16/3683/2016/, 2016.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: Climate performance and sensitivity to parametrized physics with emphasis on tropical convection, Clim. Dyn., 27, 787–813, doi:10.1007/s00382-006-0158-0, 2006.

Janssens-Maenhout, G., Dentener, F., Van Aardenne, J., Monni, S., Pagliari, V., Orlandini, L., Klimont, Z., Kurokawa, J., Akimoto, H., Ohara, T., Wankmueller, R., Battye, B., Grano, D., and Zuber, A.: EDGAR-HTAP: a Harmonized Gridded Air Pollution Emission Dataset Based on National Inventories, Tech. Rep. EUR report No. EUR 25299, European Commission Publications Office, JRC68434, Ispra, Italy, 2012.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703–4715, doi:10.1029/2000JD900581, 2001.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO_2 emissions: Minimum site location requirements for the Los Angeles megacity, J. Geophys. Res. Atmos., 118, 1577–1584, doi:10.1002/jgrd.50135, 2013.

Lac, C., Donnelly, R. P., Masson, V., Pal, S., Riette, S., Donier, S., Queguiner, S., Tanguy, G., Ammoura, L., and Xueref-Remy, I.: CO_2 dispersion modelling over Paris region within the CO_2-MEGAPARIS project, Atmos. Chem. Phys., 13, 4941–4961, doi:10.5194/acp-13-4941-2013, 2013.

Lafore, J. P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Héreil, P., Mascart, P., Masson, V., Pinty, J. P., Redelsperger, J. L., Richard, E., and Vilà-Guerau de Arellano, J.: The Meso-NH Atmospheric Simulation System. Part I: adiabatic formulation and control simulations, Ann. Geophys., 16, 90–106, doi:10.1007/s00585-997-0090-6, 1998.

Latoska, A.: Erstellung eines räumlich hoch aufgelösten Emissionsinventar von Luftschadstoffen am Beispiel von Frankreich im Jahr 2005, Master's thesis, Institut für Energiewirtschaft und Rationelle Energieanwendung, Universität Stuttgart, Stuttgart, Germany, 2009.

Lauvaux, T., Miles, N. L., Richards, S. J., Deng, A., Stauffer, D. R., Davis, K. J., Jacobson, G., Rella, C., Calonder, G.-P., and DeCola, P. L.: Urban Emissions of CO_2 from Davos, Switzerland: The First Real-Time Monitoring System Using an Atmospheric Inversion Technique, J. Appl. Meteor. Climatol., 52, 2654–2668, doi:10.1175/JAMC-D-13-038.1, 2013.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., 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), J. Geophys. Res. Atmos., 121, 5213–5236, doi:10.1002/2015JD024473, http://dx.doi.org/10.1002/2015JD024473, 2015JD024473, 2016.

Masson, V.: A physically-based scheme for the urban energy budget in atmospheric models, Bound.-Layer Meteorol., 1994, 357–397, 2000.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960, doi:10.5194/gmd-6-929-2013, 2013.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proc. Natl. Acad. Sci. U. S. A., 109, 8423–8428, doi:10.1073/pnas.1116645109, 2012.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6, 981–1028, doi:10.5194/gmd-6-981-2013, 2013.

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., and Robinson, P.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part III–Observation impact and observation sensitivity in the California Current System, Prog. Oceanog., 91, 74–94, 2011.

Nordbo, A., Järvi, L., Haapanala, S., Wood, C. R., and Vesala, T.: Fraction of natural area as main predictor of net CO_2 emissions from cities, Geophys. Res. Lett., 39, doi:10.1029/2012gl053087, 2012.

Rodgers, C. D.: Inverse methods for atmospheric sounding: Theory and practice, World Sci., London, 2000.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, Mon. Weather Rev., 139, 976–991, doi:10.1175/2010MWR3425.1, 2010.

Strong, C., Stwertka, C., Bowling, D. R., Stephens, B. B., and Ehleringer, J. R.: Urban carbon dioxide cycles within the Salt Lake Valley: A multiple-box model validated by observations, J. Geophys. Res., 116, doi:10.1029/2011JD015693, 2011.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Tans, P., Lehman, S., Davis, K. J., Miles, N. L., Richardson, S. J., Lauvaux, T., Cambaliza, M. O., Shepson, P., Gurney, K., Patarasuk, R., and Zondervan, A.: Towards quantification and source sector identification of fossil fuel CO_2 emissions from an urban area: Results from the INFLUX experiment, J. Geophys. Res. Atmos., 120, 292–312, doi:10.1002/2014JD022555, 2015.

Vogel, F. R., Thiruchittampalam, B., Theloke, J., Kretschmer, R., Gerbig, C., Hammer, S., and Levin, I.: Can we evaluate a fine-grained emission model using high-resolution atmospheric transport modelling and regional fossil fuel CO_2 observations?, Tellus B, 65, doi:10.3402/tellusb.v65i0.18681, 2010.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation network bring to the quantification of city CO_2 emissions?, Atm. Chem. Phys., 16, 7743–7771, doi:10.5194/acp-16-7743-2016, 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, Atmos. Chem. Phys. Disc., 2016, 1–51, doi:10.5194/acp-2016-218, http://www.atmos-chem-phys-discuss.net/acp-2016-218/, 2016.

Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, J. Geophys. Res-Atmos., 111, doi:10.1029/2005jd006003, 2006.

\AM@currentdocname .pdf

900

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf
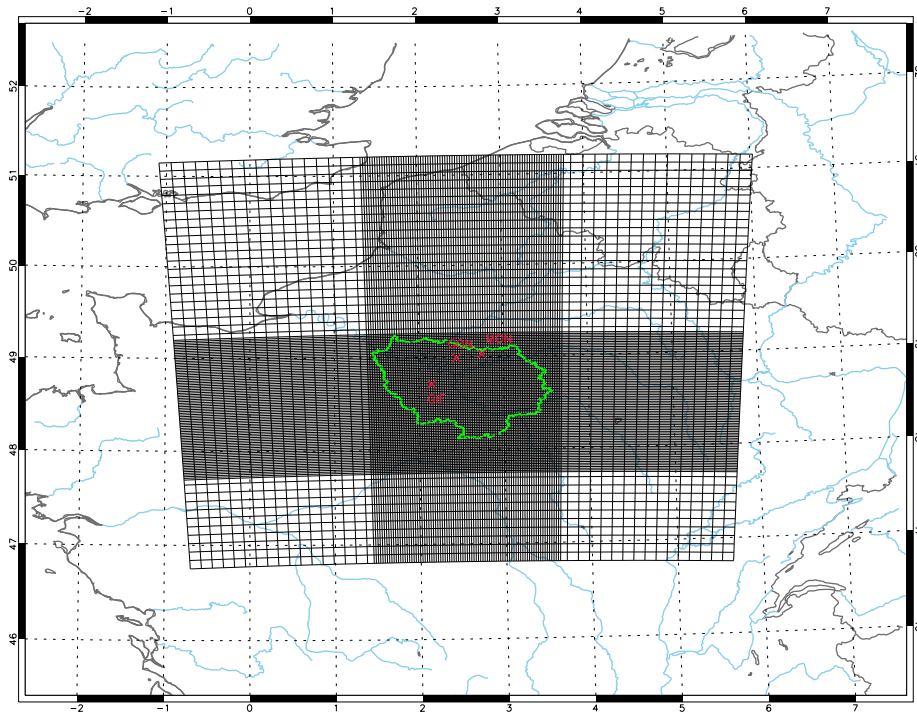
\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

\AM@currentdocname .pdf

Figure 1. Map of the Northern France modelling domain. The monitoring sites are depicted as red crosses, while the green line denotes the boundaries of the Île-de-France region. Black lines show the model grid: 2 km×2 km spatial resolution in the centre, 2 km×10 km and 10 km×10 km spatial resolution for the surroundings, respectively.
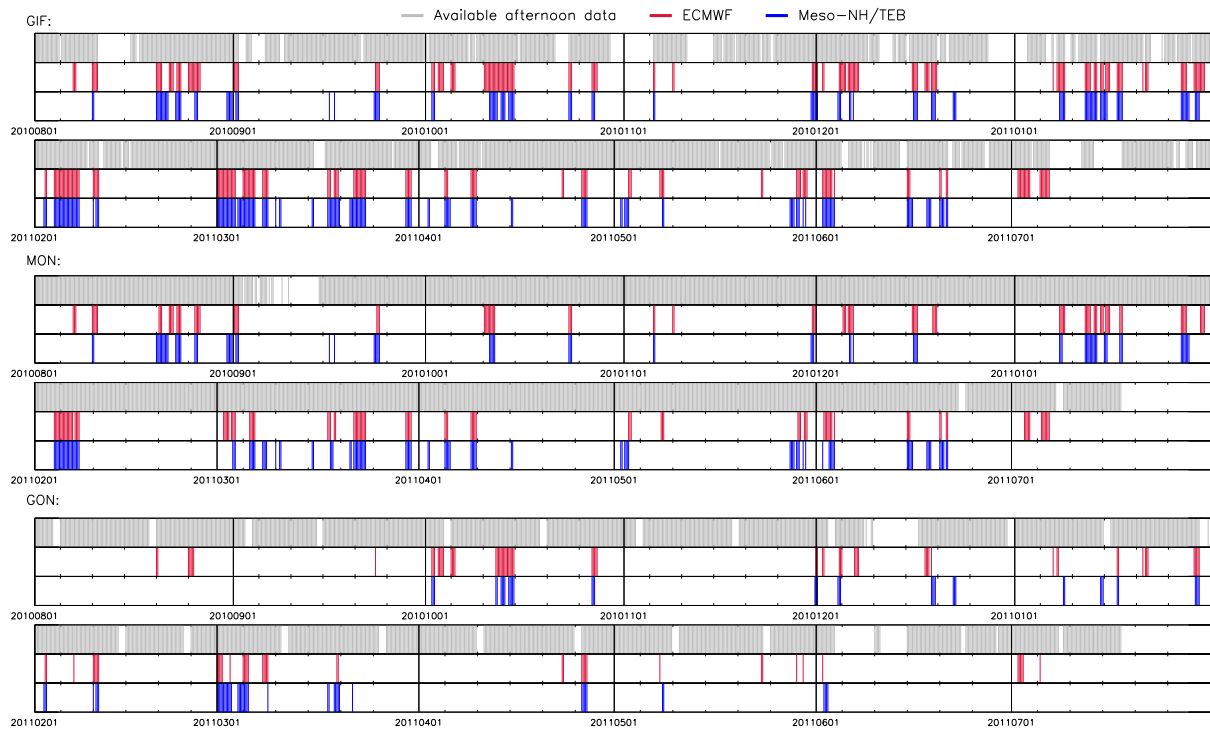
Figure 2. Afternoon (12-16 h) CO_2 data availability during the CO2-MEGAPARIS project (August 2010–July 2011) for the different monitoring sites used in this study. Available hourly observed data are displayed as grey vertical lines. Red and blue: observations that are actually assimilated when using the reference (stringent) gradient selection criteria. Red: Selection when using ECMWF wind fields for the wind estimation. Blue: Selection when using Meso-NH/TEB wind fields for the wind estimation.
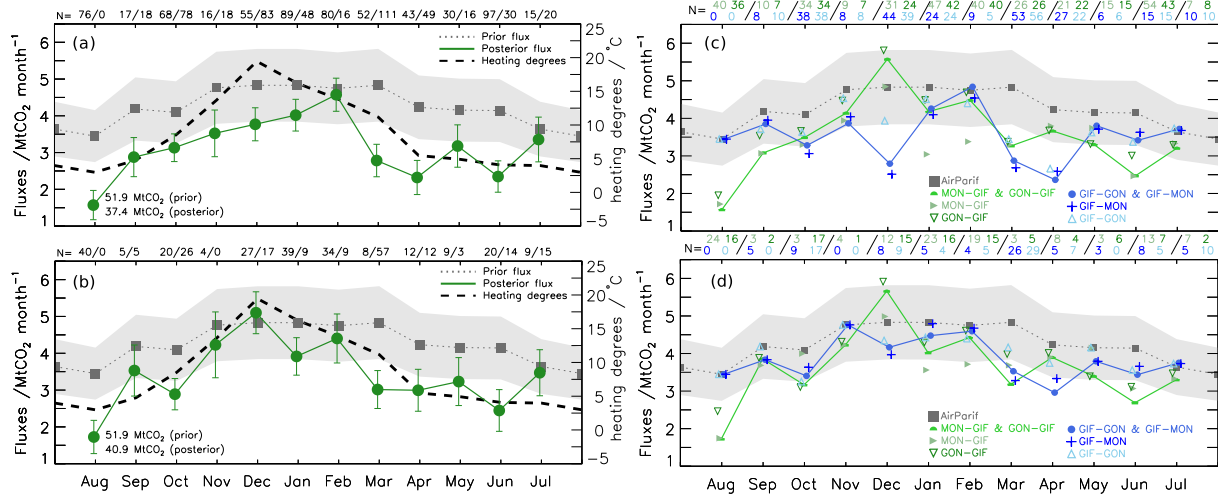
Figure 3. Monthly fossil fuel CO_2 emissions prior and posterior estimates from inversions in $MtCO_2$. Left column: Prior estimates (line) $\pm$ the standard deviation of uncertainties (shade) are displayed in grey while posterior estimates (line) $\pm$ the standard deviation (bars) are displayed in green. (**a**): Results using the initial inversion configuration. (**b**): Results using the reference inversion configuration. Monthly mean heating degrees for the centre of Paris, obtained from ECMWF's operational analysis, are displayed in black. Numbers at the top are those of the CO_2 mole fraction gradients assimilated for SW- or NE winds, respectively. Prior and posterior annual emission estimates are displayed in the left bottom corner. Right column: Results using the initial (**c**) and reference (**d**) configuration of the inversions but assimilate only subsets of selected gradients (see Sect. 2.2.) Colour-coded numbers at the top are those of the assimilated gradient by each subset. Prior estimates (line) $\pm$ the standard deviation of uncertainties (shade) are displayed in grey. Symbols are slightly shifted to prevent overlap.
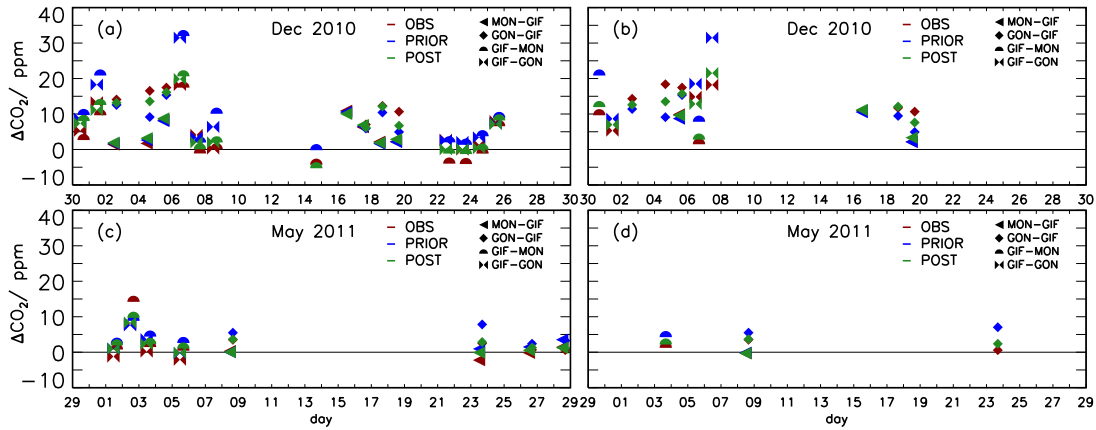
Figure 4. Time series of mean afternoon (12–16 h) CO_2 mole fraction differences (ΔCO_2) between downwind and upwind sites for December 2010 (upper panel) and May 2011 (lower panel). Measured CO_2 mole fraction differences are shown in red; prior and posterior CO_2 mole fraction differences in blue and green, respectively. (**a**) and (**c**): Selection of gradients relies on wind direction ranges of the initial inversion configuration. (**b**) and (**d**): Wind direction ranges are limited to a narrow SW/NE wind corridor across the city following the reference inversion configuration (see Sect. 2.2).
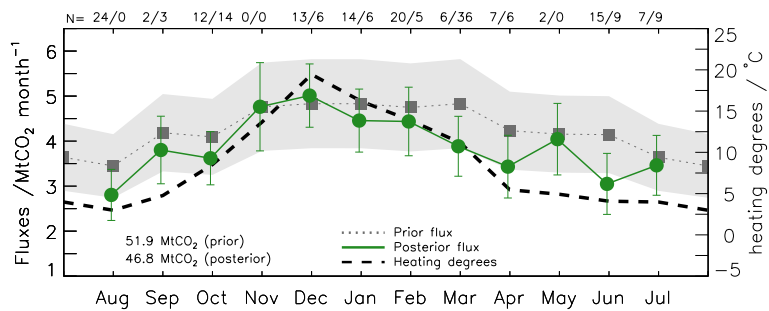

Figure 5. Results for the Lagrangian experiment (see Sect. 2.2 and Tab. 1). As for Fig. 3c, but a 2-hour time lag between downwind and upwind measurements is introduced.
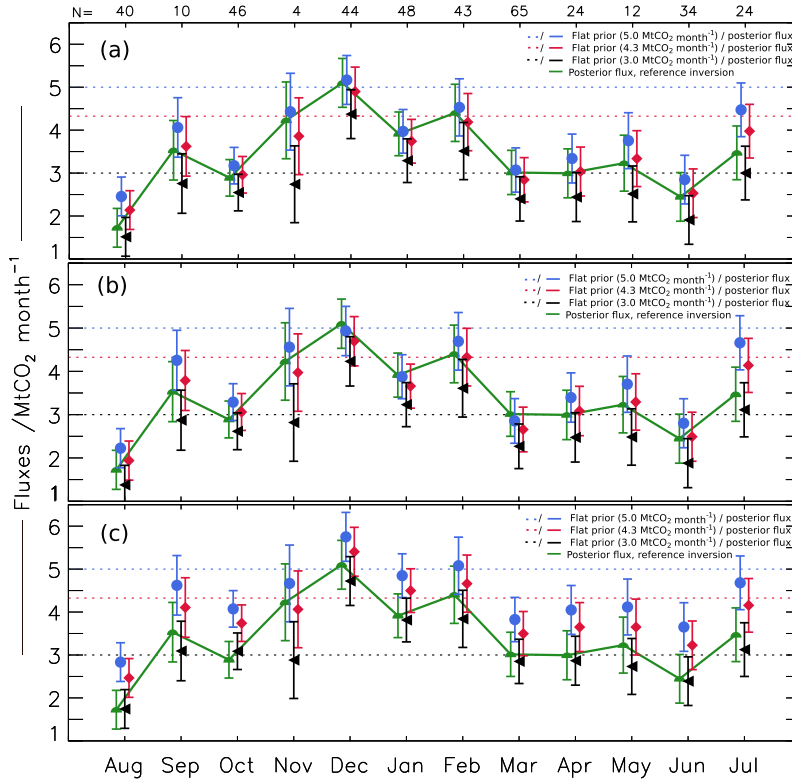
Figure 6. Sensitivity of monthly fossil fuel CO_2 emissions upon $\mathbf{x}_b$. Monthly fossil fuel CO_2 estimates $\pm$ the standard deviation of their uncertainties are shown for inversions that use $3\,\mathrm{MtCO_2 month^{-1}}$ (black), $4.3\,\mathrm{MtCO_2 month^{-1}}$ (red), and $5\,\mathrm{MtCO_2 month^{-1}}$ (blue) monthly prior emissions. **(a)** Priors are flat at monthly scale. **(b)** Priors are flat at daily scale. **(c)** Priors are flat at 6-h scale (see Sect. 2.6 for details). Fluxes obtained by the reference inversion are displayed in green. Numbers at the top denote the number of assimilated CO_2 mole fraction gradients. Symbols are slightly displaced to prevent overlap.
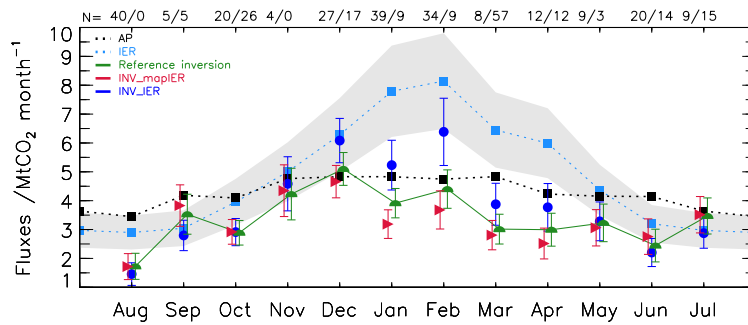
Figure 7. Sensitivity of monthly fossil fuel CO_2 emissions upon $\mathbf{H}^{\mathrm{map}}$. Red: Monthly fossil fuel CO_2 emissions estimates $\pm$ the standard deviation of their uncertainties obtained from the reference inversion (green), INV_mapIER (red), and INV_IER (blue), respectively. Monthly fossil fuel CO_2 emissions prior estimates by AIRPARIF are depicted in black while IER's monthly estimates $\pm$ the standard deviation of uncertainties are depicted in sky blue and grey, respectively. Note the different scale of the ordinate compared to Fig. 3, 5 and 6.
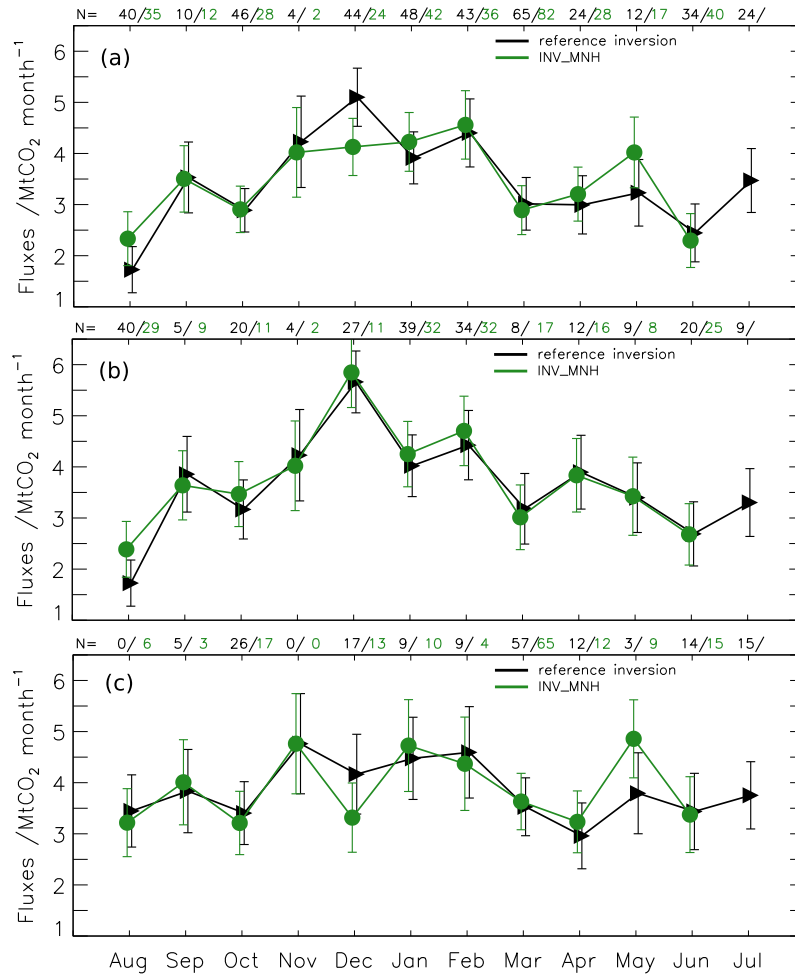
Figure 8. Sensitivity of monthly fossil fuel CO_2 budgets upon meteorological data. Displayed are the estimates $\pm$ the standard deviation of their uncertainties obtained from the reference inversion (green) and INV_MNH (black), respectively. Numbers at the top denote color-coded the number of assimilated gradients. **(a)** Assimilation of both SW and NE gradients. **(b)** Assimilation of SW gradients. **(c)** Assimilation of NE gradients.
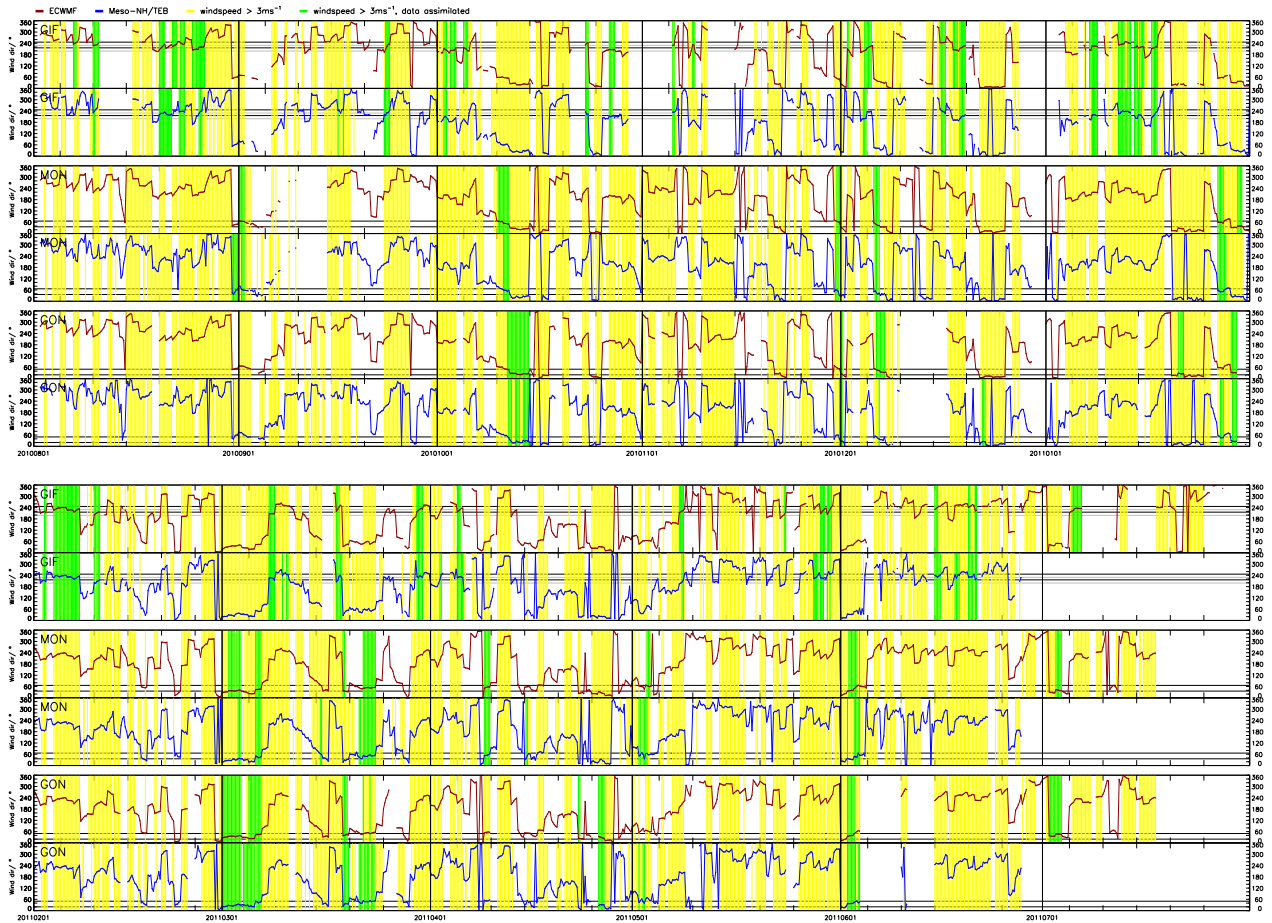
Figure A1. Time series of mean wind directions during afternoon (12-16 h) at the different monitoring sites used in this study. Solid horizontal lines denote the range of wind directions used by the reference (stringent) gradient selection (see Sect. 2.2). Red: Wind directions as simulated by ECMWF. Blue: Wind directions as simulated by Meso-NH/TEB. Yellow vertical lines indicate wind speed $>3 \, \mathrm{ms}^{-1}$. Green vertical lines: Data are actually assimilated when using the reference (stringent) gradient selection criteria.

Figure A2. Time series of mean wind speed during afternoon (12-16 h) at the different monitoring sites used in this study. Solid horizontal lines denote a $3\,\mathrm{ms}^{-1}$ wind speed threshold. Red: Wind speed as simulated by ECMWF. Blue: Wind speed as simulated by Meso-NH/TEB. Yellow vertical lines indicate wind directions are within the range of wind direction used by the reference (stringent) gradient selection. Green vertical lines: Data are actually assimilated when using the reference (stringent) gradient selection criteria.

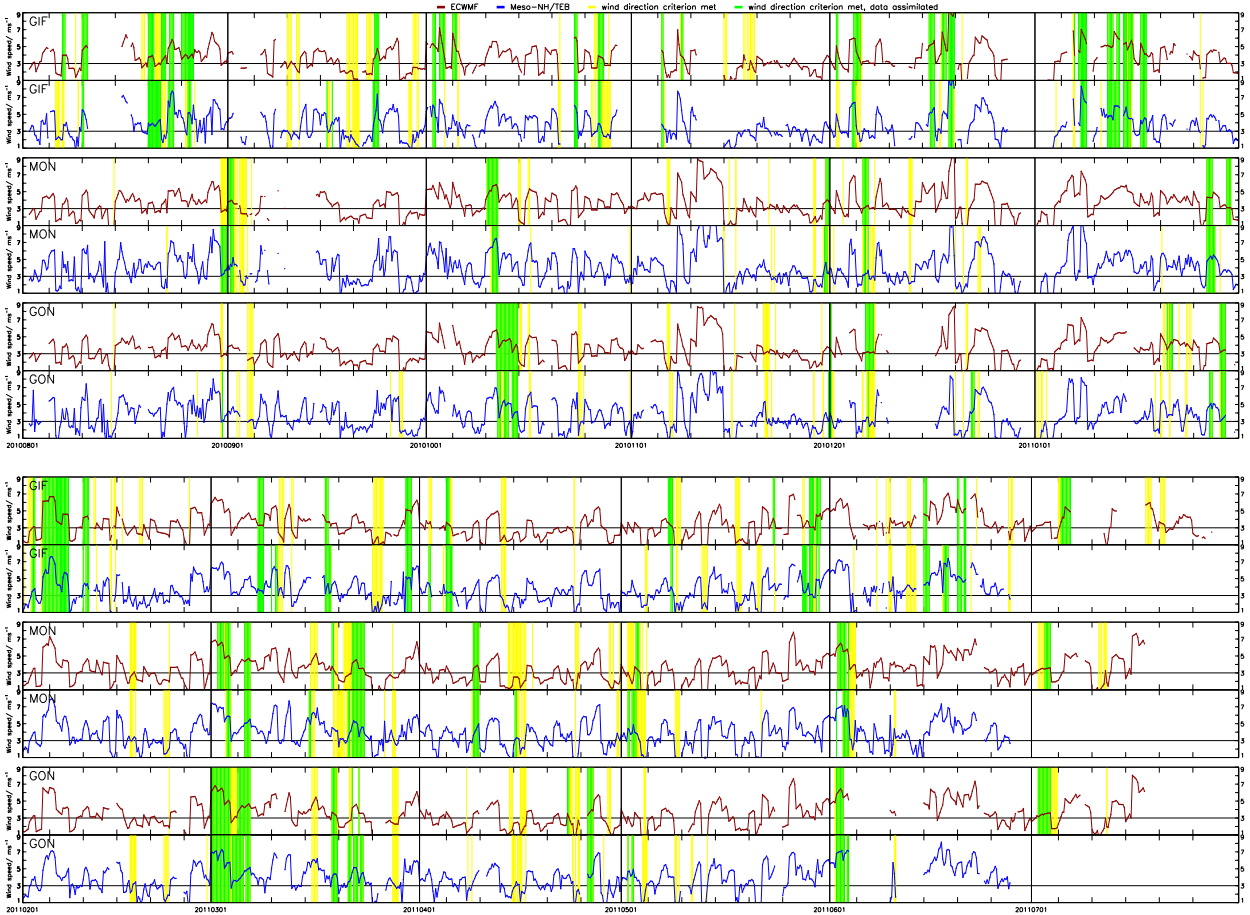Table 1. Summary of the different inversion configuration and Île-de-France (IdF) annual fossil fuel CO_2 emissions from different inventories and inversion results. Priors that are flat at the monthly, daily and 6-hourly scale are denoted M, D and H, respectively (see section 2.6 for the details). Posterior estimates are derived from inversions using the operator and prior estimate indicated in the corresponding line of the table.

Inversion	Acronym	H			$\mathbf{y}^{\mathrm{f}}$	$\mathbf{x}^{\mathrm{b}}$	IdF annual fossil fuel CO_2 emissions in $\mathrm{MtCO_2}$	
		$\mathbf{H}^{\mathrm{samp}}$	$\mathbf{H}^{\mathrm{trans}}$	$\mathbf{H}^{\mathrm{map}}$			Prior	Post
Initial	*ini*	$\mathbf{H}^{\mathrm{samp}}_{ini-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{AP}$	$\mathbf{y}^{\mathrm{f}}_{ini-ECM}$	AP08	51.9	37.4
Reference	*ref*	$\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{AP}$	$\mathbf{y}^{\mathrm{f}}_{ref-ECM}$	AP08	51.9	40.9
Sensitivity	FLAT_4.3H					H	51.9	47.1
Tests	FLAT_4.3D					D	51.9	41.1
	FLAT_4.3M					M	51.9	41.4
	FLAT_3.0H	$\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{AP}$	$\mathbf{y}^{\mathrm{f}}_{ref-ECM}$	H	36.0	37.1
	FLAT_3.0D					D	36.0	33.0
	FLAT_3.0M					M	36.0	33.0
	FLAT_5.0H					H	60.0	52.2
	FLAT_5.0D					D	60.0	45.3
	FLAT_5.0M					M	60.0	45.3
	INV_mapIER	$\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{IER}$	$\mathbf{y}^{\mathrm{f}}_{ref-ECM}$	AP08	51.9	39.0
	INV_IER	$\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{IER}$	$\mathbf{y}^{\mathrm{f}}_{ref-ECM}$	IER	60.1	45.5
	INV_MNH	$\mathbf{H}^{\mathrm{samp}}_{ref-MNH}$	$\mathbf{H}^{\mathrm{trans}}_{MNH}$	$\mathbf{H}^{\mathrm{map}}_{AP}$	$\mathbf{y}^{\mathrm{f}}_{ref-MNH}$	AP08	51.9	[1]
[..[156]]Time lag	*lag*	$\mathbf{H}^{\mathrm{samp}}_{ref-ECM}$	$\mathbf{H}^{\mathrm{trans}}_{ECM}$	$\mathbf{H}^{\mathrm{map}}_{IER}$	$\mathbf{y}^{\mathrm{f}}_{ref-ECM}$	AP08	51.9	46.8
Emissions for the year 2010 as given by AIRPARIF (2013)								41.8

[1] Meso-NH/TEB data are available up to June 2011 only

Table 2. Annual and seasonal bias, standard deviation (STD), root mean square error (RMSE) and coefficient of determination (r^2) of prior model-data misfit and posterior model-data misfit for the initial inversion (experiment ini) and the reference inversion (experiment ref), respectively. All values, except for r^2, are given in ppm.

	Bias				STD				RMSE				r^2			
	ini		ref		ini		ref		ini		ref		ini		ref	
	prior	post	prior	post	prior	post	prior	post	prior	post	prior	post	prior	post	prior	prior
Annual	2.50	0.33	3.04	0.36	3.60	2.21	3.77	2.20	4.38	2.23	4.84	2.22	0.53	0.80	0.53	0.81
JJA	2.20	0.23	2.70	0.39	2.31	1.59	2.54	1.62	3.18	1.60	3.70	1.66	0.13	0.45	0.03	0.34
SON	2.41	0.28	3.73	0.38	3.49	1.98	2.95	2.05	4.23	2.00	4.74	2.07	0.35	0.75	0.27	0.61
DJF	2.35	0.48	2.79	0.44	4.21	2.51	4.55	2.29	4.82	2.55	5.33	2.32	0.61	0.84	0.55	0.85
MAM	3.01	0.26	3.29	0.20	3.65	2.38	4.01	2.63	4.72	2.39	5.17	2.62	0.22	0.56	0.07	0.50