

We thank both reviewers for their comments on our manuscript. Thanks to them, this manuscript now conducts a discussion on the interesting topic of the spatial correlation in the observation errors, based on new tests described in Appendix A of the new manuscript. We provide below detailed responses to all of their comments and the version of the new manuscript where the changes are highlighted.

Response to the anonymous Referee #1

The authors do a good job of explaining a regional tool that can be used to evaluate observing systems. Much of this detail has already been covered in Broquet et al (2018), but they add some to the analysis presented there and extend to a more comprehensive state vector that includes the larger region, as well as exploring the impacts of satellite data precision and resolution on the inversions.

Everything as described is correct mathematically, and the results from the point of view of a linear least squares optimization are useful. I find the analysis related to the independence of various sources particularly interesting, as distinguishing from neighboring sources is critical for the mission of CO2M.

We thank the reviewer for these positive general statements.

As the authors highlight their exploration of the observational precision, it behooves me to point out that we have no reason to believe the assumption they make about independence of the errors in retrievals that are spatially near each other. In fact, work by Kulawik et al. (2020) and Worden et al. (2017) would suggest that the correlation length scales would be something more like 50km-100km for XCO₂. The assumption that errors scale by \sqrt{N} is particularly poor.

We made explicit in Section 2.1.2 that we purposely focus on the impact of the instrumental noise and we now clarify it in the introduction of our revised manuscript while mentioning the exploration of the topic of the correlations in the Appendix A:

In terms of errors in the XCO₂ data, the analysis focuses on random errors due to the instrumental noise that have no spatial correlations (even though the topic is explored in Appendix A).

We assume that errors in XCO₂ due to this instrumental noise bear no spatial correlations, following, e.g., Buchwitz et al. (2013). The analysis of current OCO-2 data at high spatial resolution demonstrates that the uncorrelated noise is a significant fraction of the total error on XCO₂ individual data (e.g. Reuter et al., 2019; Zheng et al. 2020).

Studying the correlated errors from the radiative transfer inverse modelling was out of the scope of our study.

I can appreciate that handling systematic errors in a classical uncertainty reduction framework is not straightforward,

Systematic errors can be accounted for in a classical uncertainty reduction framework, for instance based on a Monte Carlo approach (Broquet et al., 2018). However, assigning them remains difficult because:

- they are not described in the uncertainties calculated by existing retrieval schemes, in contrast to the retrieval noise caused by instrument noise. The publications cited above characterize them empirically based on the statistics of the retrieval small-scale variability (Worden et al., 2017) or on the statistics of the difference to reference retrievals that are themselves empirically related to WMO standards (Kulawik et al., 2019a,b).
- they depend on the specific measurement configurations and on the evolving skill of radiative transfer models and of empirical bias-correction systems, so that conclusions from existing missions may hardly apply to future ones.

but handling correlated observation errors should be doable.

Including correlations in the \mathbf{R} matrix of our inversion framework is definitely doable and we have conducted some tests with such correlations to support our discussion on this topic. In order to lighten the computations associated with the inversion of such a matrix when using a 900 km swath, we have actually tested it when considering 4 km resolution pixels rather than 2 km resolution ones.

A traditional way to model spatial correlations is to assume that they are isotropic and homogeneous in terms of spatial scale (as probably suggested by the reviewer when speaking about 50 to 100 km scale correlations). We have modeled such correlations using exponentially decaying functions $\exp(-d/D)$ of the distances d between two observation pixels, and tested their inclusion in \mathbf{R} with a 1-sigma uncertainty of 0.3 ppm for individual XCO₂ data and with $D=10, 50$ or 100 km (assuming the correlations apply to the total uncertainty in individual data i.e. not splitting \mathbf{R} into a random noise component and a component with spatial correlations). These tests are now described in Appendix A where the following Figure A3 is also provided. These results and this topic are now discussed in Section 4, with the following main messages:

- such isotropic correlations should make the errors much easier to filter by the inversion than actual spatial patterns of the systematic errors that can follow the same atmospheric dynamics as the signature of the targeted fluxes
- the situation is made even more optimistic here (like in most of studies with atmospheric inversion OSSEs) since our inversions are perfectly informed about the statistics of the observation errors (and in particular about their spatial scale)
- as expected, the results indicate that when introducing correlations with small spatial scales, the posterior uncertainties increase since these correlations yield larger budget of observation errors at the scale of the signatures of the targeted fluxes. Conversely, introducing correlations helps the inversion separate the observation error patterns from these signatures, and large correlation spatial scales lead to a decrease of the posterior uncertainties. The worst case corresponds to correlation scales that are function of the scale of the signatures of the targeted fluxes. It varies depending on whether we analyze results for point sources, cities or widespread emissions across regions. This is why (i) results are poorer with $D=10$ km than with no correlations, for all type of sources (ii) results become generally better with spatial correlation than without for point sources and cities respectively for $D \geq 10$ km and $D \geq 50$ km (iv) the

inversion of the largest of a given category of sources (in terms of amplitude) are generally more negatively or less positively impacted by the spatial correlations (iii) results are systematically worse with $D= 10$ to 100 km than with no correlations, for countryside emissions. Results for regional budgets mix all these behaviors in a complex way.

- These results and conclusions should be further investigated in future theoretical studies (to properly analyze the characterization of the spatial scale of the signature of the targeted fluxes) but this would fall out of the scope of this paper since (i) we want to keep the focus on the impact of random instrumental noise (ii) we believe that such a simple structure of the error correlations and their optimal learning by the inversion should be interpreted carefully to avoid misleading conclusions regarding the impact of actual systematic errors in the inversions.

Here is the corresponding text added in section 4:

Our new inversion framework allows accounting for a realistic simulation of the observation sampling and errors. Nevertheless, generating simulations of the systematic errors from the retrieval of XCO_2 data that are suitable for the purpose of our study would have been difficult. Systematic errors are not described in the uncertainties computed by existing retrieval schemes. Furthermore, they depend on specific measurement configurations and on the evolving skill of radiative transfer inverse models and of empirical bias-correction systems, so that their characterization based on diagnostics with existing missions may hardly apply to future ones. Simulating realistic patterns of cloud cover consistent with the meteorology for the different test cases investigated would have also been challenging. Finally, this study focuses on other parameters to allow exploring the sensitivity of the inversions to these parameters in depth.

In order to raise insights into the impact of errors with spatial patterns such as model and systematic errors, we have conducted experiments where spatial correlations are included in the observation error (in the \mathbf{R} matrix). We have tested isotropic and homogeneous spatial correlations exponentially decaying with distance, using various correlation lengths. The experiments and results are described in Appendix A since they are out of the scope of this study. The results indicate that including correlations in the observation errors tends to increase the budget of observation errors and thus to increase the posterior uncertainties in the flux estimates as long as the correlation length scale does not exceed that of the signature of the fluxes in the XCO_2 images. However, including correlations in the observation errors also tends to increase the ability to distinguish between the patterns of the observation error and of the signatures of the fluxes, and thus to decrease the posterior uncertainties, so that for large spatial correlation lengths, increasing the correlation length leads to a decrease of the posterior uncertainties. In our tests, the worst situation for the monitoring of the emissions in the study area corresponds to ~ 10 km correlation length scales. These results should be interpreted cautiously since the spatial patterns of the model and systematic errors are more complex than this traditional but simple modeling of spatial correlations and since in these tests, the inversion system is perfectly informed about the statistics of the observation error. Future studies will integrate more realistic simulations of observation sampling and errors from different concepts of spaceborne imagery, based on radiative transfer inverse modeling applied to realistic fields of surface and atmospheric conditions and instrumental specifications.

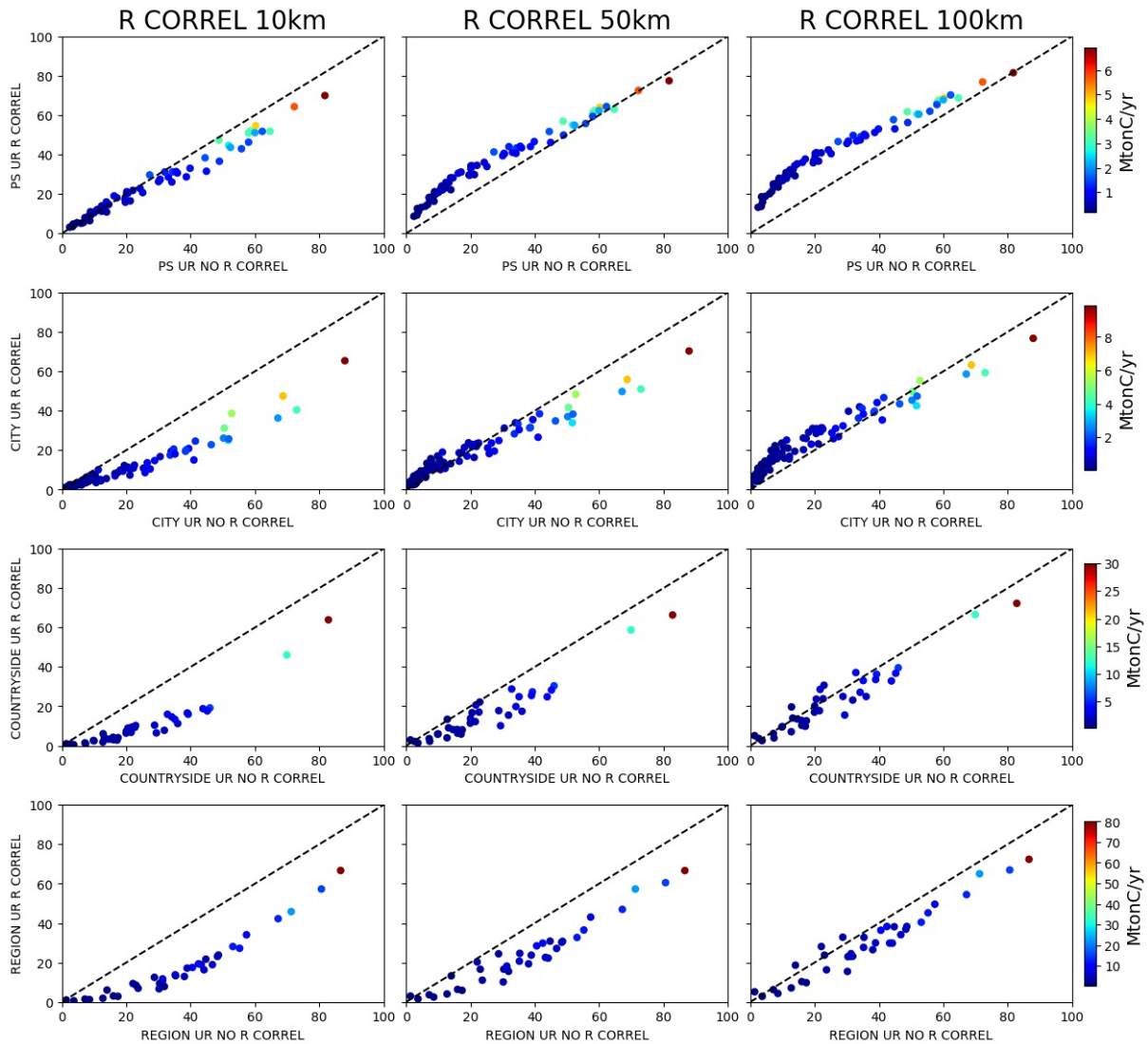


Figure A3: Uncertainty reductions (UR) for the emissions of the point sources (1st row), the urban areas (2nd row), the countryside areas (3rd row) and the regions (4th row) when considering spatial correlations (Y-axis) or no correlations (X-axis) between the observation errors. The 1st, 2nd and 3rd column correspond to correlations with a spatial scale of 10, 50 and 100 km respectively. The color of the dots corresponds to the annual budgets of the sources (color bar on the right of the figure). The dashed line is the 1:1 line.

This consideration is particularly important for small scale sources/sinks of the sort that the authors are claiming to constrain.

The dependence of the impact of such correlations on the type of sources is definitely an important aspect of the problem as highlighted above.

I think the paper is worthy of publication, but I do think that it will have more impact with this one extra factor considered.

As explained above, we have tried to consider this extra factor in the new version, while respecting the coherence of the study (by providing the details of the tests of sensitivity to the

spatial correlations in **R** in Appendix A) and acknowledging the unknown evolving nature of systematic errors.

I also recommend a bit more rigorous grammar and spelling check, as I noticed typographical errors and grammatical errors as I read the manuscript.

We have carefully proofread the text to significantly improve its grammar and spelling (see track changes in the revised manuscript).

I don't have specific comments, as the presentation is straightforward, and the figures are self-explanatory.

We thank the reviewer again for such a positive assessment.

Response to the Anonymous Referee #2

The authors have done a series of inversions in an OSSE framework to explore the feasibility of a wide swath CO₂-sensing satellite mission such as CO2M to quantify the fossil fuel (FF) CO₂ emissions from individual cities, clusters of cities, and regions. They have relied on uncertainty reduction as the primary metric for evaluating their inversions. The work they have done is mathematically correct, and the results follow expectations from prior experience with inversions, i.e., they “make sense”.

We thank the reviewer for his comments and for raising detailed discussions on our inversion set-up.

However, there are three critical shortcomings, owing to which I cannot recommend publication unless they make the suggested changes. Since making these changes is likely to require a fair amount of additional work, and may change their conclusions, I am classifying them as “major” changes.

As detailed below, the first “shortcoming” raised by the reviewer does not apply to the inversions that we have presented. Furthermore, we do not agree that suggestions of changes related to the “shortcoming #2 and #3” should be required. We disagree with many details of these suggestions, and we view the discussions on the corresponding topics as incentive for further investigations, while the set of original experiments is already significant, and while the manuscript is already long, analyzing many aspects of the inversion and satellite CO₂ data assimilation problems. However, since the first part of the third suggestion is also raised by the first reviewer, we provide some analysis in Appendix A and some discussions in the manuscript to address the corresponding topic.

First, in any CO₂-based effort to quantify FF CO₂ emissions, the biggest confounding factor is the nonFF variation of CO₂. This can be due to biosphere fluxes over continents, or, in a regional study, due to inflow/outflow at the boundary. This has been a well-known problem both in the context of estimating FF CO₂ from in situ and satellite CO₂ measurements (e.g., <https://doi.org/10.1002/2014GL059684> and <https://doi.org/10.1029/2019JD030528> respectively). Biases in the assumed NEE – which is very likely in any biosphere model – will lead to biases in derived FF CO₂, which makes any uncertainty reduction irrelevant. And yet,

in this set of OSSEs the authors skirt this very important issue, assuming the prior NEE to be unbiased.

This first “major” comment is misplaced since our control vector does include both NEE fluxes at hourly and regional scale and the boundary conditions so that uncertainties in these fluxes and conditions are clearly accounted for. We describe the prior NEE errors in terms of random time-correlated uncertainties rather than in terms of biases (to use the reviewer’s words), but:

- this is the usual statistical representation of uncertainties in NEE in inversions, in particular in those targeting NEE estimates (Rayner et al., 2019)
- we do not see how this distinction would change the conclusions of the paper

Furthermore, specific tests are conducted to analyze the impact of uncertainties in the NEE in section 3.3.2 which also discusses this “confounding factor” based on the correlations between the uncertainties in these fluxes and the anthropogenic emissions.

However, we take the opportunity of such a comment to revise section 2.4 to better clarify the set-up of uncertainties in the NEE and boundary conditions in the **B** matrix.

Here are the corresponding sentences added or modified in section 2.4. (see track changes in the revised manuscript):

- *Prior estimates of the boundary conditions for regional inversions are usually interpolated from large scale analysis or inversions. Such products can bear significant large-scale errors at the boundaries of Europe (Monteil et al., 2018). We reflect it by setting in **B** the standard deviation of the prior uncertainty in the scaling factor for the boundary conditions (see below).*
- *Finally, we use 1% for the STD of the prior uncertainty in the scaling factor associated to the boundary conditions (i.e. typically an uncertainty of ~4 ppm in the average boundary conditions). This value is quite pessimistic, but some tests in which this value was varied (not shown) demonstrate a very weak sensitivity of the results for the fluxes to this parameter.*
- *Despite the differences between the temporal variations of the hourly emissions from one control area to the other, or between natural and anthropogenic fluxes in \mathbf{H}_{distr} , these STD show a negligible variation of less than 1% and when considering the reference set-up for **B**, $\sigma_{hour} \sim 65\%$.*

While this does not make the results wrong, it makes them less than useful for evaluating the potential of a CO2M-like mission, which will surely have to contend with unknown biospheric CO2 fluxes.

Again, we explicitly account for uncertainties in the NEE in our inversions and we have a result and analysis section dedicated to the topic.

Second, in the context of satellite CO2 instruments, an additional complication is the data gap or poor data due to cloud cover, aerosol loading, and other factors. For example, for the currently flying OCO2 spectrometer, only a few percent of its footprints result in good quality XCO2 retrievals. For the domain the authors have chosen (NW Europe with lots of urban centers), both cloud cover and aerosol loading are important limiting factors. Yet, the authors explicitly ignore this complication, “The cloud cover and the corresponding gaps in the

spaceborne passive XCO₂ sampling are ignored” (L100). I understand that simulating realistic clouds and aerosols is difficult, but for an OSSE to be realistic, some attempt must be made. For example, the authors could have used the statistics of past cloud and aerosol data to introduce sampling gaps. Or, they could have taken the fraction of good quality retrievals among all footprints from an existing NIR XCO₂ instrument like OCO₂ and then downsampled their footprints. The results presented here without considering realistic sampling gaps are mathematically correct, but not very useful for evaluating the capability of any real CO₂ mission.

Accounting for cloud cover would be important if we had to analyze results for a specific mission with a fixed spatial and temporal sampling, and over a long period of time, in order to highlight the frequency of its cloud free observations and potentially some asset of using a higher spatial resolution for the observations (as discussed in section 4). However, our study mainly focuses on the sensitivity of the inversion to other parameters. Adding the impact cloud cover, with its short-term chaotic component, in the experiments would not help raise such sensitivity curves and understand them. Furthermore, characterizing the sensitivity to cloud cover itself for individual images of plumes is not straightforward since it depends on how cloud patterns overlap with the plumes or not, especially if the set of cloud cover maps used for the study is limited. Of important note is that the impact of testing various cloud cover for the inversion of the emissions from a given city was analyzed in our previous study (Broquet et al., 2018).

From a technical point of view, the reviewer’s suggestion is challenging. The point is not just about having a realistic percentage of pixels removed from the image due to cloud cover, but also to have realistic pattern of gaps that are consistent with the meteorology. In particular, a uniform spread of cloudy pixels would necessarily impact the observation of a given plume in a partial way, while structured cloud coverage could either leave intact or completely hide a given plume. A way to tackle the problem is generally to simulate XCO₂ data based on real earth observation data. For example, Buchwitz et al. (2013) simulated the sampling by the proposed CarbonSat mission using MODIS data. This was relatively straightforward since the CarbonSat orbit was taken as that of Terra, since its swath was narrower and since its spatial resolution was coarser than that of MODIS data. However, here, we simulate theoretical observations every day during several months at various spatial resolutions. Building assumptions and processing MODIS data to achieve such a simulation would be a study in itself, and is clearly out of the scope of the present paper. We thus disagree with the reviewer’s suggestion to add an analysis of the impact of cloud cover on the results.

We now clarify this position in the manuscript (section 4):

Simulating realistic patterns of cloud cover consistent with the meteorology for the different test cases investigated would have also been challenging. Finally, this study focuses on other parameters to allow for exploring the sensitivity of the inversions to these parameters in depth.

Third, the authors use the posterior covariance matrix and uncertainty reduction as performance metrics for their inversions throughout the paper.

This is not not specific to this paper at all. It is a standard way to diagnose the skill of inversions (Rayner et al., 2019).

Given the importance of uncertainty calculation to the work (as opposed to reduction of biases in their priors),

There is no opposition: in the optimal estimation framework, the Kalman gain matrix that controls the random uncertainty reduction (together with the observation operator) also reduces biases in the prior. The mathematical demonstration is trivial.

I would like to see a more realistic specification of flux and data uncertainty. Currently they assume (i) uncorrelated retrieval errors (L193), which is unrealistic for the small footprint and dense sampling that they've given their satellite instrument,

We do not agree with this point since we explicitly focus on the random measurement errors from instrumental noise. Such correlations arise in the systematic errors that have been ignored in this study. As written to reviewer 1:

We made explicit in Section 2.1.2 that we purposely focus on the impact of the instrumental noise and we now clarify it in the introduction of our revised manuscript while mentioning the exploration of the topic of the correlations in the Appendix A:

In terms of errors in the XCO₂ data, the analysis focuses on random errors due to the instrumental noise that have no spatial correlations (even though the topic is explored in Appendix A).

We assume that errors in XCO₂ due to this instrumental noise bear no spatial correlations, following, e.g., Buchwitz et al. (2013). The analysis of current OCO-2 data at high spatial resolution demonstrates that the uncorrelated noise is a significant fraction of the total error on XCO₂ individual data (e.g. Reuter et al., 2019; Zheng et al. 2020).

Studying the correlated errors from the radiative transfer inverse modelling was out of the scope of our study.

Systematic errors can be accounted for in a classical uncertainty reduction framework, for instance based on a Monte Carlo approach (Broquet et al., 2018). However, assigning them remains difficult because:

- they are not described in the uncertainties calculated by existing retrieval schemes, in contrast to the retrieval noise caused by instrument noise. The publications cited above characterize them empirically based on the statistics of the retrieval small-scale variability (Worden et al., 2017) or on the statistics of the difference to reference retrievals that are themselves empirically related to WMO standards (Kulawik et al., 2019a,b).
- they depend on the specific measurement configurations and on the evolving skill of radiative transfer models and of empirical bias-correction systems, so that conclusions from existing missions may hardly apply to future ones.

However, including correlations in the \mathbf{R} matrix of our inversion framework is definitely doable and we have conducted some tests with such correlation to support our discussion on this topic. In order to lighten the computations associated with the inversion of such a matrix when using a 900 km swath, we have actually tested it when considering 4 km resolution pixels rather than 2 km resolution ones.

A traditional way to model spatial correlations is to assume that they are isotropic and homogeneous in terms of spatial scale (as probably suggested by the reviewer when speaking about 50 to 100 km scale correlations). We have modeled such correlations using exponentially decaying functions $\exp(-d/D)$ of the distances d between two observation pixels, and tested their inclusion in \mathbf{R} with a 1-sigma uncertainty of 0.3 ppm for individual XCO₂ data and with $D=10, 50$ or 100 km (assuming the correlations apply to the total uncertainty in individual data i.e. not splitting \mathbf{R} into a random noise component and a component with spatial correlations). These tests are now described in Appendix A where the following Figure A3 is also provided. These results and this topic are now discussed in Section 4, with the following main messages:

- such isotropic correlations should make the errors much easier to filter by the inversion than actual spatial patterns of the systematic errors that can follow the same atmospheric dynamics as the signature of the targeted fluxes
- the situation is made even more optimistic here (like in most of studies with atmospheric inversion OSSEs) since our inversions are perfectly informed about the statistics of the observation errors (and in particular about their spatial scale)
- as expected, the results indicate that when introducing correlations with small spatial scales, the posterior uncertainties increase since these correlations yield larger budget of observation errors at the scale of the signatures of the targeted fluxes. Conversely, when introducing correlations helps the inversion separate the observation error patterns from these signatures, and large correlation spatial scales lead to a decrease of the posterior uncertainties. The worst case corresponds to correlation scales that are function of the scale of the signatures of the targeted fluxes. It varies depending on whether we analyze results for point sources, cities or widespread emissions across regions. This is why (i) results are poorer with $D=10$ km than with no correlations, for all type of sources (ii) results become generally better with spatial correlation than without for point sources and cities respectively for $D \geq 10$ km and $D \geq 50$ km (iv) the inversion of the largest of a given category of sources (in terms of amplitude) are generally more negatively or less positively impacted by the spatial correlations (iii) results are systematically worse with $D= 10$ to 100 km than with no correlations, for countryside emissions. Results for regional budgets mix all these behaviors in a complex way.

These results and conclusions should be further investigated in future theoretical studies (to properly analyze the characterization of the spatial scale of the signature of the targeted fluxes) but this would fall out of the scope of this paper since (i) we want to keep the focus on the impact of random instrumental noise (ii) we believe that such a simple structure of the error correlations and their optimal learning by the inversion should be interpreted carefully to avoid misleading conclusions regarding the impact of actual systematic errors in the inversions.

Here is the corresponding text added in section 4:

Our new inversion framework allows accounting for a realistic simulation of the observation sampling and errors. Nevertheless, generating simulations of the systematic errors from the retrieval of XCO₂ data that are suitable for the purpose of our study would have been difficult. Systematic errors are not described in the uncertainties computed by existing retrieval schemes. Furthermore, they depend on specific measurement configurations and on the evolving skill of radiative transfer inverse models and of empirical bias-correction

systems, so that their characterization based on diagnostics with existing missions may hardly apply to future ones. Simulating realistic patterns of cloud cover consistent with the meteorology for the different test cases investigated would have also been challenging. Finally, this study focuses on other parameters to allow exploring the sensitivity of the inversions to these parameters in depth.

In order to raise insights into the impact of errors with spatial patterns such as model and systematic errors, we have conducted experiments where spatial correlations are included in the observation error (in the \mathbf{R} matrix). We have tested isotropic and homogeneous spatial correlations exponentially decaying with distance, using various correlation lengths. The experiments and results are described in Appendix A since they are out of the scope of this study. The results indicate that including correlations in the observation errors tends to increase the budget of observation errors and thus to increase the posterior uncertainties in the flux estimates as long as the correlation length scale does not exceed that of the signature of the fluxes in the XCO₂ images. However, including correlations in the observation errors also tends to increase the ability to distinguish between the patterns of the observation error and of the signatures of the fluxes, and thus to decrease the posterior uncertainties, so that for large spatial correlation lengths, increasing the correlation length leads to a decrease of the posterior uncertainties. In our tests, the worst situation for the monitoring of the emissions in the study area corresponds to ~10 km correlation length scales. These results should be interpreted cautiously since the spatial patterns of the model and systematic errors are more complex than this traditional but simple modeling of spatial correlations and since in these tests, the inversion system is perfectly informed about the statistics of the observation error. Future studies will integrate more realistic simulations of observation sampling and errors from different concepts of spaceborne imagery, based on radiative transfer inverse modeling applied to realistic fields of surface and atmospheric conditions and instrumental specifications

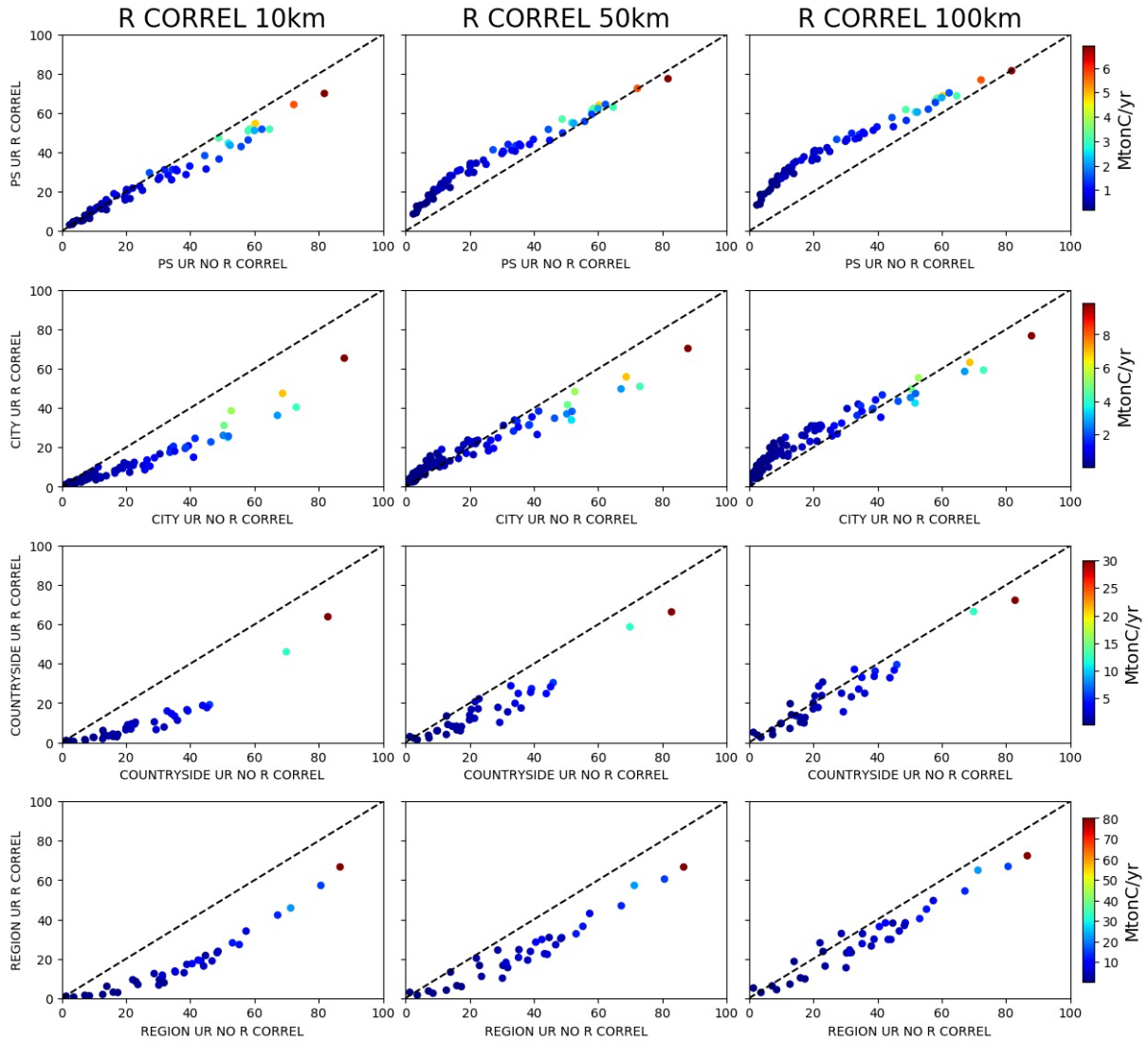


Figure A3 : Uncertainty reductions (UR) for the emissions of the point sources (1st row), the urban areas (2nd row), the countryside areas (3rd row) and the regions (4th row) when considering spatial correlations (Y-axis) or no correlations (X-axis) between the observation errors. The 1st, 2nd and 3rd column correspond to correlations with a spatial scale of 10, 50 and 100 km respectively. The color of the dots corresponds to the annual budgets of the sources (color bar on the right of the figure). The dashed line is the 1:1 line.

and (ii) no spatial correlation in their prior flux covariance B (L334), also unrealistic given the high spatial resolution of their fluxes.

The characterization of the correlations of the uncertainties in anthropogenic emissions of CO₂ from inventories (used as prior estimates) is an extremely complex topic, and, currently, there is certainly no consensus on how these correlations can be defined. In fact, one may argue for both positive and negative correlations. Positive spatial correlations arise from uncertainties in the emission factors applied at national scales in the inventories. Negative correlations arise from the spatial disaggregation of national inventories into gridded maps of emissions (Wang et al., 2018) use negative correlations between regions of the same country). Variations of activities from one city to another, from one plant to another... de-correlate

uncertainties. The use of local emission factors also de-correlates uncertainties. Finally, in a general way, the spatial correlations should be highly complex and have no reason to decrease with distance (as suggested by the reviewer's statement "given the high spatial resolution of their fluxes"), since e.g. uncertainties in emissions from two distant cities with large industrial activities should be more correlated than that between two neighbor cities with a completely different share of domestic, commercial, transport and industrial activities. State of the art estimates of the structures of uncertainties in the anthropogenic emissions (Super et al., 2020) strongly depend on the model used for the inventories and on the chosen assumptions regarding the major sources of uncertainties.

Given the current lack of characterization of correlations of the uncertainties in anthropogenic emissions of CO₂, ignoring them or assuming that they are null is a safe solution. We clarify this point in section 2.4 and in section 4:

The spatial correlations of the uncertainties in anthropogenic emission inventories is a complex topic and the current lack of characterization for such correlations led to such a conservative set-up (Wang et al. 2018; Wang et al., 2020; Super et al., 2020).

There is still a critical lack of knowledge and of characterization of the correlations in the uncertainties in the inventories (Wang et al. 2018, Wang et al., 2020). However, some extensive analyses are now conducted to fill this gap (Super et al., 2020).

NEE is controlled at regional scale only which implicitly translates into significant spatial correlations of the uncertainty in the bottom up estimates of the NEE.

I do not understand why they need to make either simplification, since in a batch inversion (as opposed to an iterative approximation like EnKF or 4DVAR) one can actually specify off-diagonals in both **B** and **R**. The simplifications (i) and (ii) make their posterior covariance, and the conclusions based thereon, not very relevant for real-world inversions of CO₂M-like satellite data.

This is not about simplification but about making realistic scenarios as detailed above and we strongly disagree with the last statement of the reviewer. Of note is that, in a general way, adding spatial correlations in **B** and **R** tends to simplify the inversion problem, to spoil the asset of solving for the fluxes and concentration at high resolution, and to increase the scores of uncertainty reduction. We think it would lead to over-optimistic results.

A general comment I would like to make about the three issues I raise above is that in inversions of real satellite data, modelers often have to make simplifying choices to make the problem tractable. E.g., ignoring off-diagonal elements in **R** is pretty common, although more recent inversions try to at least account for correlations in **R** by error inflation, aggregating, or data thinning (e.g., <https://doi.org/10.1029/2007GL030463>, <https://doi.org/10.5194/acp-13-8695-2013> and <https://doi.org/10.5194/acp-19-9797-2019>).

As detailed above, we can definitely include correlations in **R** and we did it in answer to both reviews. However, the corresponding set up and results are not relevant enough for our paper to be included within the result section so that they are kept for Appendix A.

Similarly, inversions with an iterative approximation like EnKF or 4DVAR often have an inexact posterior covariance due to computational limitations, while many inversions ignore biases in satellite retrievals because it's still an open problem. However, making too many simplifications in a single OSSE, as the authors have done here (unbiased priors and retrievals, no sampling gaps, uncorrelated retrieval errors, no spatial error correlation in fluxes) makes the results inapplicable to any real-world satellite instrument.

See our answers to all these general critics with which we obviously disagree.

I will also point out that several coauthors of this manuscript have previously authored papers stressing the importance of these complicating factors (e.g. <https://doi.org/10.1029/2007GL030463>, <https://doi.org/10.5194/amt-11-681-2018>, and <https://doi.org/10.1088/1748-9326/ab7835>) and published inversions with far more realistic assumptions, which makes the current set of simplifications all the more surprising.

We thank the reviewer for recalling these studies that highlighted various challenges which have often been disregarded in other studies. However, they should not be used as an obligation for us to solve all these long-standing challenges in this paper that is driven by its own scientific questions.

Chevallier (2007) studied the impact of hypothetical observation error correlations of 0.5 in neighboring observations supposedly caused by modelling errors (radiative transfer and atmospheric transport). Such errors are out of the scope of our study, as clearly stated in the submitted version, but, as we have explained, we have now included tests to touch this topic.

Broquet et al. (2018) explored the impact of errors in the emission spatial sampling and in the boundary conditions but discussed the shortcoming of testing these errors as biases due to the lack of ensemble simulations for these parameters. They studied the impact of cloud cover and systematic errors in the last step of their incremental analysis, after having diagnosed series of experiments ignoring it, in order to focus on other parameters first, which led to a major part of the conclusions. Our introduction and discussion / conclusion highlight the amount of new topics explored here in cloud free conditions without systematic errors.

The problem of the uncertainties in biofuel emissions and of human respiration from cities has been raised by Ciais et al. (2020) as a problem for the spaceborne observation but not tackled by any inversion system yet. Actually, this publication shows that this problem will have to be addressed by complementary data or source of information since one can hardly adapt an atmospheric inversion approach that assimilates satellite XCO₂ images only. We now discuss this additional source of uncertainty in our results (section 4):

A last significant simplification of the general problem of the inversion of the anthropogenic emissions based on XCO₂ data has been stressed by Ciais et al. (2020). Anthropogenic emissions of CO₂ bear a major share of emissions from biofuel combustion which can hardly be separated spatially from the fossil fuel combustion component. Furthermore, the emissions of CO₂ by human respiration represent a significant portion of the total CO₂ emitted from cities. The XCO₂ data and the atmospheric inversion approaches can hardly be used to distinguish between these different components if it cannot rely on complementary data. This factor was ignored here, as well as in most of the studies dedicated to the inversion of anthropogenic CO₂ emissions at city scale to regional scale.

Beside these major issues, here are a few minor points that need correcting or clarifying:

1. L156: Delete “and vertical”. The vertical resolution comes later.

Done

2. L161: A high spatial resolution (~2x2 km²) implies higher temporal resolution as well. If the driving winds are 3-hourly, what provides high frequency variation in the CHIMERE winds?

Before any simulation, a preprocessing stage interpolates the 3-hourly winds from ECMWF at 1h-resolution and at the spatial resolution of CHIMERE to generate the wind forcing in input to the model. CHIMERE itself interpolates the hourly forcing at each of its physical time step (Menut et al., 2013). To clarify this point, we add a sentence in section 2.1.1: *These three-hourly fields are interpolated at the spatial and temporal resolution of CHIMERE.*

3. L175: The gradients in column CO₂ due to the top 30% of the atmosphere would be small, agreed, but how large are they? Signals in column CO₂ are deceptively small, so terms that seem to be negligible are not always negligible.

The analysis of the CO₂ concentrations at 11:00 produced by the response functions associated to the emissions between 5:00 and 6:00 of several cities and point sources shows that these CO₂ concentrations at the upper level of the model are indeed negligible with respect to the ones within the boundary layer (with a factor < 1e-6). This supports our approximation to prescribe the CO₂ concentrations in the upper layers of the atmosphere. To clarify this point, we add in section 2.1.1 the sentence:

This is supported by the lack of signal in our simulations of the atmospheric signatures of the surface fluxes in the upper layer of the model.

4. L180: Switch 92.8° and 705 km.

Done

5. L227: Since the quantity directly estimated is the FF CO₂ emission between 5 and 11 local time, to estimate the total emissions one would need an accurate diurnal cycle. What is the uncertainty in the diurnal cycle of FF CO₂ emissions?

The structures of temporal correlations of the uncertainties in the estimates of emissions from inventories is a very complex topic (as well as that of their spatial correlations). The uncertainty in the diurnal cycle of the emissions has hardly been characterized and will strongly depend on the inventories that are considered. Some insights regarding these uncertainties can be found in Super et al. (2020).

However, this part of the paper was not addressing the topic of the estimate of the total emissions (over the day or over the year). Note that we have mentioned this topic at the end of section 4 (*Exploiting further capabilities of the inversion framework: potential of complementary observation systems and results at larger temporal scales*).

6. Section 2.2.2 and elsewhere: The word “controlled” keeps confusing me. Do you mean “estimated”, as in part of the control vector? Or do you mean “controlled” as in kept in control, static, not changed? I suspect you mean the former, but “controlled” in English can also signify “not allowed to change”. I’d suggest using the word “estimated” or “optimized” if you mean the former.

Throughout the text, we use the term “control – controlled (by the inversion)” in a unique, clear and explicit way that is equivalent to “correct / adjust / update - corrected / adjusted /

updated”. Rayner et al. (2019) acknowledges this traditional denomination even though they promote the term « target »: « We term the set of these quantities the “target variables” of the problem. They are also called unknowns, parameters or control variables. » (they do not speak about « estimated » or « optimized » variables). We prefer « control » than « target » since the inversion often has to control some variables which are not a scientific target to explicitly account for and tackle uncertainties from these degrees of freedom. In this OSSEs framework where the posterior estimate of the control parameters is not derived, the term “estimated” and “optimized” would not be ideal, and we do not think that “estimated vector” or “estimated region” sounds very well. Finally, the term optimized can be misleading in statistical inversion, especially when focusing on the posterior error covariance matrices i.e. on the statistical distribution of the posterior uncertainty around its optimal estimate rather than on this optimal estimate.

7. L280: Is this an over-determined problem? Then that’s not very common in flux inversions, and is likely due to the unrealistic correlations in **R** and **B** (one of my major concerns).

1) We are speaking here about the size of control and observation vectors in a statistical inversion problem, not about the question whether the problem is over or under constrained. Statistical inversion problems where the number of observations is larger than the number of control variables, especially when studying spaceborne imagery of XCO₂, is quite common (e.g. Reuter et al., 2014; Kemp et al., 2014; Pillai et al., 2016).

2) Furthermore, adding correlations in **B** will implicitly reduce the size of the space of uncertainty in the fluxes, not the opposite.

3) Finally, mathematically speaking, changing correlations in **B** and **R** will not change the size of the control and observation vectors.

Regarding the major concerns on **R** and **B**, see the discussions above.

8. L290: Is random noise added to **y**, consistent with **R**?

At this stage of the paper, the text has not said that we will generate pseudo observation **y**. It just provides the theoretical basis for the following. The following clarifies the fact that we do not generate such pseudo-data.

9. L304: Typo, change XO₂ à XCO₂

Done

10. L348: With the assumptions detailed here at the grid scale, what is the uncertainty on the (say) annual total or seasonal total NEE and FF? Aggregate numbers are easier to make sense of than grid-scale specifications.

We do not want to enter into discussions about the uncertainties in annual fluxes and about their link with uncertainties at the 6-hour scale in this part of the paper. This requires discussing the temporal correlations of the prior uncertainties in inventories of the anthropogenic emissions at the daily to annual scale. Again, this is a highly complex topic (Wang et al., 2020; Super et al. 2020), and the general topic is discussed in the last part of

section 4 (*Exploiting further capabilities of the inversion framework: potential of complementary observation systems and results at larger temporal scales*).

11. L399: “Figures 2i” likely means all the subplots of Figure 2. In that case, just say “Figures 2”, no need to add the “i”.

Done

12. L437: Speed is one aspect of the wind, direction is the other. Since wind direction determines how well plumes present themselves to a satellite that is going one way, uncertainty in wind direction must be considered as well as speed uncertainty. Was that done here?

This section does not discuss uncertainties in the wind but the sensitivity to wind variations. This specific part of the text demonstrates the influence of wind speed on the results. The impact of changes in the wind direction is significant for narrow swaths and discussed later.

13. L477: Remove “uncertainty”, **B** is just the prior covariance matrix.

No, absolutely not. **B** is the prior uncertainty covariance matrix i.e. **B** is the covariance matrix of the uncertainty in the prior, not the covariance matrix of the prior.

14. L519: Again, I’d like to see the uncertainties on aggregated fluxes, such as annual totals.

We conduct 6-hour inversions. There is no need in this study to make assumptions on temporal correlations of the prior uncertainty on larger timescales to propose some corresponding annual uncertainties.

15. L560: In Figure 7, why do larger emissions have smaller uncertainty reductions?

The Figure 7 shows that the general tendency is to get larger uncertainty reductions for sources with larger emissions. However, other factors driving the variations in uncertainty reductions may soften this conclusion for sources that have similar emissions: e.g., if the plumes from sources are driven by highly different wind speeds. The influence of the wind on the inversion results is analyzed in sections 3.2.1 and 3.3.1.. And, as discussed in sections 3.2.2 and 3.3.1., the UR for a given source is also sensitive to the level of uncertainty in NEE around, and to the potential loss of part of its plume at the edge of the satellite swath. We clarify this point by adding the sentence in section 4: *Beside these sensitivities to the source amplitudes and to the uncertainties in NEE, the variations in UR are also driven by the wind (sections 3.2.1 and 3.3.1) and by the potential loss of part of the atmospheric signatures of sources at the edges of the satellite swath (section 3.2.2).*

16. L582: “... and thus by the variability of these fluxes, during the month of May”. This only matters because the uncertainty on the NEE is larger in May, right? Because this metric/score does not care about the actual prior NEE.

Yes, it does, since we define relative values for the prior uncertainties in the NEE rather than absolute values for the prior uncertainties in the NEE. So if the NEE varies, the resulting absolute value of the corresponding prior uncertainty will vary too.

We clarify this point in section 3.3.1 with the sentence:

Using constant prior relative uncertainties in the natural fluxes (as for the anthropogenic emissions) yields large absolute uncertainties in May and low absolute uncertainties in March.

17. L794: “Efforts have been made to limit the amplitude of such errors in the concept of the new CO2M mission. Our new inversion framework allows accounting for a realistic simulation of the observation sampling and errors.” I disagree with this statement in the context of this paper, especially the part about a realistic simulation, because of the three major points raised above.

Again, see our answers to these points. We do not think that this sentence has to be modified.

References

Broquet, G., Bréon, F.M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., Chevallier, F., Wu, L., and Ciais, P.: The potential of satellite spectro-imagery for monitoring CO₂ emissions from large cities, *Atmos. Meas. Tech.*, 11, 681–708, <https://doi.org/10.5194/amt-11-681-2018>, 2018.

Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V., Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO₂ and CH₄ retrieval errors by error parameterization, *Atmos. Meas. Tech.*, 6, 3477–3500, <https://doi.org/10.5194/amt-6-3477-2013>, 2013.

Chevallier, F.: Impact of correlated observation errors on inverted CO₂ surface fluxes from OCO measurements. *Geophysical Research Letters*, 34(24), 2007.

Ciais, P., Wang, Y., Andrew, R., Bréon, F. M., Chevallier, F., Broquet, G., ... & Zheng, B.: Biofuel burning and human respiration bias on satellite estimates of fossil fuel CO₂ emissions. *Environmental Research Letters*, 15(7), 074036, 2020.

Kemp, S., Scholze, M., Ziehn, T., & Kaminski, T.: Limiting the parameter space in the Carbon Cycle Data Assimilation System (CCDAS). *Geoscientific Model Development*, 7(4), 1609-1619, 2014.

Kulawik, S. S., O'Dell, C., Nelson, R. R., and Taylor, T. E.: Validation of OCO-2 error analysis using simulated retrievals. *Atmospheric Measurement Techniques*, 12(10), 2019a.

Kulawik, S. S., Crowell, S., Baker, D., Liu, J., McKain, K., Sweeney, C., Biraud, S. C., Wofsy, S., O'Dell, C. W., Wennberg, P. O., Wunch, D., Roehl, C. M., Deutscher, N. M., Kiel, M., Griffith, D. W. T., Velasco, V. A., Notholt, J., Warneke, T., Petri, C., De Mazière, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollock, D. F., Morino, I., Uchino, O., Hase, F., Feist, D. G., Roche, S., Strong, K., Kivi, R., Iraci, L., Shiomi, K., Dubey, M. K., Sepulveda, E., Rodriguez, O. E. G., Té, Y., Jeseck, P., Heikkinen, P., Dlugokencky, E. J., Gunson, M. R., Eldering, A., Crisp, D., Fisher, B., and Osterman, G. B.: Characterization of OCO-2 and ACOS-GOSAT biases and errors for CO₂ flux estimates, *Atmos. Meas. Tech.*, in review, 2019b.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., ... & Mailler, S.: CHIMERE 2013: a model for regional atmospheric composition modelling. *Geoscientific model development*, 6 (4), 981-1028, 2013.

Pillai, D., Buchwitz, M., Gerbig, C., Koch, T., Reuter, M., Bovensmann, H., Marshall, J., and Burrows, J. P.: Tracking city CO₂ emissions from space using a high-resolution inverse modeling approach: a case study for Berlin, Germany, *Atmos. Chem. Phys.*, 16, 9591–9610, <https://doi.org/10.5194/acp-16-9591-2016>, 2016.

Rayner, P. J., Michalak, A. M., & Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry. *Atmospheric Chemistry & Physics*, 19(22), 2019.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., ... & Butz, A.: Satellite-inferred European carbon sink larger than expected. *Atmospheric Chemistry and Physics*, 14(24), 13739-13753, 2014.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., ... & Burrows, J. P.: Towards monitoring localized CO₂ emissions from space: co-located regional CO₂ and NO₂ enhancements observed by the OCO-2 and S5P satellites. *Atmospheric Chemistry and Physics*, 19(14), 9371-9383, 2019.

Super, I., Dellaert, S. N., Visschedijk, A. J., & Denier van der Gon, H. A.: Uncertainty analysis of a European high-resolution emission inventory of CO₂ and CO to support inverse modelling and network design. *Atmospheric Chemistry and Physics*, 20(3), 1795-1816, 2020.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., ... & Tao, S.: Potential of European 14 CO₂ observation network to estimate the fossil fuel CO₂ emissions via atmospheric inversions. *Atmospheric Chemistry and Physics*, 18(6), 4229-4250, 2018.

Wang, Y., Broquet, G., Bréon, F.-M., Lespinas, F., Buchwitz, M., Reuter, M., Meijer, Y., Loeschner, A., Janssens-Maenhout, G., Zheng, B., and Ciais, P.: PMIF v1.0: an inversion system to estimate the potential of satellite observations to monitor fossil fuel CO₂ emissions over the globe, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2019-326>, in review, 2020.

Worden, J. R., Doran, G., Kulawik, S., Eldering, A., Crisp, D., Frankenberg, C., ... & Bowman, K.: Evaluation and attribution of OCO-2 XCO₂ uncertainties. *Atmospheric Measurement Techniques*, 10(7), 2759-2771, 2017.

Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., & Zhao, Y.: Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2. *Atmospheric Chemistry and Physics*, 20(14), 8501-8510, 2020.

A local to national-scale inverse modeling system to assess the potential of spaceborne CO₂ measurements for the monitoring of anthropogenic emissions

Diego Santaren¹, Grégoire Broquet¹, François-Marie Bréon¹, Frédéric Chevallier¹, Denis Siméoni², [Bo Zheng¹](#), Philippe Ciais¹

¹Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Gif-sur-Yvette, France

²Thales Alenia Space, La Bocca, France

Correspondence to: Diego Santaren (diegosantaren@gmail.com)

Abstract. This work presents a flux inversion system for assessing the potential of new satellite imagery measurements of atmospheric CO₂ to monitor anthropogenic emissions at scales ranging from local intense point sources to regional and national scales. While the modeling framework keeps the complexity of previous studies focused on individual and large cities, this system encompasses a wide range of sources to extend the scope of the analysis. This atmospheric inversion system uses a zoomed configuration of the regional transport model CHIMERE which covers most of Western Europe with a 2-km resolution grid over [Northern France](#), [Western Germany](#) and Benelux. For each day of March and May 2016, over the 6 hours before a given satellite overpass, the inversion controls separately the hourly budgets of anthropogenic emissions in this area from ~300 cities, power plants and regions. The inversion also controls hourly regional budgets of the natural fluxes. This enables the analysis of results at the local to regional scales for a wide range of sources in terms of emission budgets and spatial extent while accounting for the uncertainties associated [with-to](#) natural fluxes and the overlapping of plumes from different sources. The potential of satellite data to monitor CO₂ fluxes is quantified [by-with](#) posterior uncertainties or uncertainty reductions (URs) from prior inventory-based statistical knowledge.

A first analysis focuses on the hourly to 6-hour budgets of the emissions of the Paris urban area, and on the sensitivity of the results to different characteristics of the images of vertically integrated CO₂ (XCO₂) corresponding to the spaceborne instrument: the pixel spatial resolution, the precision of the XCO₂ retrievals per pixel, and the swath width. This sensitivity analysis provides a correspondence between these parameters and thresholds on the targeted precisions [on-of](#) emission estimates. However, the results indicate a large sensitivity to the wind speed and to the *prior* flux uncertainties. The analysis is then extended to the large ensemble of point sources, cities and regions in the study domain, with a focus on the inversion system ability to monitor separately neighbor sources whose atmospheric signatures overlap and are also mixed with those produced by natural fluxes. Results highlight the strong dependence of uncertainty reductions to the emission budgets, to the wind speed and [on](#) whether the focus is on point or area sources. With the system hypothesis that the atmospheric transport is perfectly known, the results indicate that the atmospheric signal overlap is not a critical issue. For the emissions within the 6-hours before a satellite overpass, [URs](#) of more than 50% can only be achieved for power plants and cities whose annual

emissions are more than $\sim 2 \text{ MtC}\cdot\text{yr}^{-1}$. For ~~more~~-regional budgets encompassing more diffuse emissions, this threshold increases up to $\sim 10 \text{ MtC}\cdot\text{yr}^{-1}$. The results suggest therefore an imbalance of the monitoring capabilities towards high and dense sources.

35 1. Introduction

Comprehensive information about anthropogenic CO_2 emissions integrated at the scale of power plants, cities, regions and countries up to the globe would allow decision makers to track the effectiveness of emission reduction policies in the context of the Paris Agreement on Climate and other voluntary emission reduction efforts. By observing the CO_2 plumes downwind of large cities and industrial plants, and atmospheric signals at a few to several hundred km scales, future high-resolution
40 spectro-imagery of the column-average CO_2 dry air mole fraction (XCO_2) from space may help addressing this need (Ciais et al., 2015; Pillai et al., 2016; Pinty et al., 2017; Schwandner et al., 2017; Broquet et al., 2018). The Copernicus Anthropogenic Carbon Dioxide Monitoring mission (CO2M, (Pinty et al., 2017)) is a prominent example of such a strategy. The CO2M concept relies on a constellation of sun-synchronous satellites with XCO_2 spectral-imagers to be deployed from 2025 by the European Commission and the European Space Agency (ESA). It will be based on passive radiance
45 measurements in the Short-Wave InfraRed (SWIR), a part of the spectrum that is sensitive to CO_2 and CH_4 concentrations throughout the troposphere including the boundary layer, like almost all the space missions that have been dedicated to ~~GreenHouseGreenhouse~~ Gas (GHG) monitoring so far (Crisp ~~et al.~~, 2018).

Much remains to be understood and to be developed in order to ensure that such a constellation informs about emissions with enough detail to be relevant for policy makers. In this context, Observing System Simulation Experiments (OSSEs) of
50 atmospheric inversions with synthetic images of XCO_2 data have ~~been-supporte~~ding the design of the space missions that will monitor the anthropogenic emissions (Buchwitz et al., 2013b; Pillai et al., 2016; Broquet et al., 2018). So far, they have mainly focused on plume inversions for some large plants and cities. However, Wang et al. (2019) estimated that cities and plants emitting more than $10 \text{ MtC}\cdot\text{yr}^{-1}$ like Berlin (in the study by Pillai et al. (2016)) and Paris (in the study by Broquet et al. (2018)) represent less than $\sim 7\%$ of the global CO_2 emissions. Furthermore, the studied cases are generally quite isolated
55 from other large CO_2 sources, facilitating the distinction of their plumes in the XCO_2 images, while plumes from neighbor sources could overlap and hamper the attribution to ~~each-underlying-emission~~the targeted source~~city or plant~~. Finally, the signature of emissions in spaceborne imagery does not consist only of clear plumes from cities, industrial clusters and point sources. Despite the large atmospheric signature of the natural fluxes, atmospheric inversions may have the potential to exploit other spatial variations in XCO_2 fields to quantify regional to national budgets of more diffuse sources or of all types
60 of sources, even when the overlapping of several plumes prevent from quantifying the emissions from individual cities and point sources. Therefore, there is a need to extend the OSSEs to a representative range of sources with various emission budgets and spreads, and various distances to other major sources, and to a larger range of spatial scales.

We have developed a high-resolution inversion system for the monitoring of CO₂ anthropogenic emissions at spatial scales ranging from local intense point sources like industrial sites to regional and national scales. Furthermore, this system accounts for the uncertainty in the natural fluxes. Our current simulation domain covers most of Western Europe with an extensive ensemble of cities, plants and diffuse CO₂ emissions. We use an analytical inversion methodology, which is the most adapted approach to efficiently test an ~~important-large~~ number of observation scenarios (section 2.1.2.), a high-resolution atmospheric transport model (section 2.1.1) and a high spatial resolution distribution of the emissions derived from different inventory products developed by the Institut für Energiewirtschaft und Rationelle energieanwendung (IER) of the University of Stuttgart (section 2.2).

The analytical inversion system follows the traditional Bayesian formalism of the atmospheric inversion. Of direct relevance here, it derives uncertainty statistics of its “posterior” emission estimates for the controlled sources (plants, cities, countryside areas or whole regions) from: i) the assumed uncertainties in the budgets derived from “prior” emission inventories (built on statistics of the fossil fuel consumption, activity data and emission factors), ii) the spaceborne XCO₂ observation sampling and precision, and iii) its atmospheric transport model. The improvement of the knowledge on the emissions enabled by the satellite imagery is quantified here in terms of “uncertainty reduction”, i.e. of the relative difference between the prior and posterior uncertainties.

The inversion system solves for hourly budgets of the emissions from the different types of sources over the 6 hours before the satellite observation of the corresponding area. ~~Indeed~~, Broquet et al. (2018) showed ~~indeed~~ that, due to atmospheric diffusion, the atmospheric signatures of emissions from a mega-city like Paris ~~which-that~~ are detectable in satellite XCO₂ images; ~~(made with current measurement capabilities)-in-satellite-XCO₂-images~~; correspond approximately to the city emissions occurring within less than 6 hours before the satellite overpass. This duration should be even shorter for the range of sources analyzed in our study since most of them have lower emissions than Paris. The analysis of the results will primarily focus on the 6h-budgets of the emissions before the satellite observation. However, controlling the hourly budgets allows evaluating the capability to solve for the temporal profiles of the emissions. It also allows accounting for some level of independence ~~of-thebetween~~ uncertainties in the emissions from different hours, which limits the ability to cross and extrapolate information throughout the 6-hour windows. This point is critical for cities whose detectable atmospheric signatures are representative of emissions on durations shorter than 6-hours, and thus for which there is no direct constraint from the satellite observation on the first hours of such 6-hour windows.

The OSSEs presented in this study use a rather simple simulation of the XCO₂ observation sampling and errors from a single helio-synchronous satellite over the area of interest. The aim is indeed to provide a general understanding of the performance of the inversion system and of its potential to monitor anthropogenic emissions with spaceborne XCO₂ imagery rather than to evaluate a precise mission configuration with precise orbital parameters and instrumental specifications. In terms of errors in the XCO₂ data, the analysis focuses on random errors due to the instrumental noise that have no spatial correlations (even though the topic is explored in Appendix A). Nevertheless, a large range of values for the precision (assumed to be homogeneous in the satellite field of view), horizontal resolution and swath of ~~a-the~~ spaceborne instrument are tested to

assess the impact of these parameters on the inversion results, which can potentially support the design of future missions (section 3.2.).

100 Furthermore, in order to get a wide range of atmospheric transport conditions (in particular ~~for-of average~~ wind speeds, (Broquet et al., 2018)) and natural flux conditions (Pillai et al., 2016), inversions are performed for each day of March and May 2016. We work with full images of the plumes from the targeted sources each day, by flying a satellite any day in ~~its~~ ~~their~~ vicinity with a large swath. Partial images of the plumes will be analyzed when studying the sensitivity to the swath width (over Paris) only. The cloud cover ~~and large aerosol loads~~, and the corresponding gaps in the spaceborne passive XCO₂ sampling are ignored. In any case, the satellite crosses the area of interest at 11:00 (local time used hereafter) similar 105 to what is currently recommended for the CO2M mission (Pinty et al., 2017), so that the inversion controls the hourly emissions of the sources between 5:00-11:00.

The inversion system and the corresponding transport model extend from ~~S~~southern France to ~~northern~~North Germany and from ~~w~~Western UK to ~~e~~Eastern Germany (Fig. 1). However, the grid of the transport ~~model~~ is zoomed and the analysis focuses on a 2 km-resolution sub-domain covering the ~~n~~North of France (in particular Paris), ~~s~~South ~~e~~East England (in 110 particular London), ~~w~~West Germany, Belgium, Luxembourg and the Netherlands.

The first part of the analysis concerns the monitoring of the emissions of Paris and its suburb, which represent the most populated and densest urban area of Europe. Broquet et al. (2018) chose this megacity as a study case because its emissions are high (~11-14 MtC·yr⁻¹ for 2013 according to the AIRPARIF inventory (Stauffer et al., 2016; AIRPARIF 2013)), concentrated and relatively distant from other major sources. ~~Moreover, T~~the topography of the region is ~~moreover~~relatively 115 flat and the average wind speed is moderate: 7 m·s⁻¹ ~~on average~~ at 100 m above the ground level (Broquet et al., 2018). The XCO₂ plume generated by the Paris emissions has a relatively simple structure that often emerges well from the background.

~~Therefore, T~~the monitoring of the emissions of Paris constitutes ~~thus~~a very favorable case with respect to other cities in Europe. Broquet et al. (2018) performed some analysis of the sensitivities of the inversion results ~~as-a-function-of~~to the wind speed and ~~of-to~~ the XCO₂ spaceborne spectro-imagery average precision and horizontal resolution. However, they tested a 120 limited number of values for these observation parameters, and in particular few high precisions (< 2 ppm) and a single high spatial resolution ~~value~~ (< 4x4 km²) ~~value~~, while the refinement of the specification of new missions requires understanding

of the sensitivity to choices of precision at the 0.1 ppm scale and of resolution at the 1 km² scale. ~~Indeed~~, these choices ~~hav~~ing large impacts on the design of the instrument and therefore on its cost (Pinty et al., 2017). Furthermore, Broquet et al. (2018) performed all their OSSEs with an unique hypothesis on the prior uncertainties in the 1-hour to 6-hour budgets of 125 the emissions from Paris, while acknowledging that ~~this-the characterization-specification of these prior uncertainties~~ could have a significant impact on the results and that the uncertainties in the inventories at such a temporal scale are difficult to assess. Therefore, this study performs a deeper investigation of the sensitivity to the observation precision and spatial resolution, to the wind speed and to the characterization of the uncertainties in the prior estimate of the emissions (section 3.2.).

130 The second part of this study considers the full ensemble of ~~emission~~-sources, from point sources to regions, in the 2-km resolution sub-domain. This sub-domain encompasses the Netherlands, Belgium and ~~w~~Western Germany which are characterized by densely populated areas distributed over a network of medium-sized towns and by a large number of strong point sources, with ~~for example~~, some power plants in ~~w~~Western Germany emitting ~~for example~~ more than 5 MtC·yr⁻¹. The ability of the inversion system to disentangle the plumes ~~of from~~ neighbor sources is therefore well challenged in these areas.

135 This paper is organized as follows: section 2 details the theoretical and practical framework of both the inversion system and ~~of the~~ OSSEs conducted in this study. Section 3 analyses the results relative to the monitoring of the emissions of Paris and in particular the corresponding sensitivity ~~analyses-tests~~ (section 3.2). This section also diagnoses the potential of the inversion system to monitor the anthropogenic emissions at the point source, city and regional scales in the area where flux and concentrations are simulated at 2-km resolution (section 3.3). Section 4 addresses the robustness and extent of the conclusions that can be derived from this study and proposes some perspectives regarding the analytical inversion system and the monitoring of anthropogenic emissions based on satellite data.

2. Inverse modeling system and OSSEs

In the following sections, we describe the different structural elements of the analytical inversion system: the gridded inventories used to define the point or area sources to be controlled, and to map their emissions (section 2.2), the simulation of the atmospheric CO₂ and XCO₂ signatures of the controlled sources using the atmospheric transport model CHIMERE and the matrix computation of the posterior uncertainties in the emissions (sections 2.1.1 and 2.3). We also describe the observations and parameters chosen for the OSSEs in this study: the XCO₂ observation sampling and errors (section 2.1.2) and the prior uncertainties in the emissions, natural fluxes and boundary conditions (section 2.4). New OSSEs with the analytical inversion system can easily be conducted with other options for these ~~These~~ observations and parameters ~~can be easily modified to conduct new OSSEs with the analytical inversion system~~. However, for the sake of clarity, the descriptions of i) the components of the inversion system and of ii) the options for the OSSEs are intertwined.

2.1. XCO₂: transport model simulations and pseudo-data

2.1.1. Simulations of CO₂ and XCO₂ with the CHIMERE model

To compute the 4D CO₂ signatures of surface CO₂ fluxes in the study domain and for March and May 2016, and of the domain CO₂ boundary conditions, we use the regional atmospheric transport model CHIMERE (Menuet et al., 2013). This Eulerian mesoscale model was designed to simulate pollution (Pison et al., 2007) but has also been used for CO₂ atmospheric inversions, and in particular for city-scale inversions of the emissions from Paris (Bréon et al., 2015; Staufer et al., 2016; Broquet et al., 2018). It has shown high skill in simulating the daily and synoptic variability of the atmospheric CO₂ concentrations at European CO₂ continuous measurements sites (Patra et al., 2008).

160 The domain of our CHIMERE configuration covers most of Western Europe (Fig. 1) between latitudes $\sim 42^\circ\text{N}$
 (nNorthern Spain) and $\sim 56^\circ\text{N}$ (nNorthern Germany) and between longitudes 6°W (eEastern Ireland) and $\sim 17^\circ\text{E}$ (eEastern
 Germany). The horizontal ~~and vertical~~ resolutions of the zoomed grid of this configuration ranges from 2 to 50 km; the $2\text{ km} \times 2\text{ km}$
 resolution sub-domain being appropriate to simulate the atmospheric signature of a dense network of sources in
 165 ~~nNorthern~~ France, Belgium, the Netherlands, Luxembourg and ~~wWestern~~ Germany. The zoom and ~~the~~ extent of the
 CHIMERE grid together link the simulation of CO_2 at local scales in this area of interest with the transport of CO_2 at the
 European scale while mitigating the computational cost. The model has 29 sigma vertical layers that extend from the surface
 to 300 hPa. Model concentration outputs are averaged at the hourly scale. The meteorological forcing is from the $9\text{ km} \times 9$
 km- and 3-hour- resolution analysis of the European Center for Medium-Range Weather Forecasts (ECMWF). These three-
hourly fields are interpolated at the spatial and temporal resolution of CHIMERE. The CO_2 concentrations used to impose
 170 the conditions at the initial time and at the lateral and top boundaries of the CHIMERE domain are from the analysis of the
 Copernicus Atmosphere Monitoring System (CAMS, (Inness et al., 2019)) at $\sim 16\text{ km}$ resolution. The products used to
 impose surface CO_2 fluxes in the model are detailed below in section 2.2.

X CO_2 observations and the corresponding signatures of fluxes are simulated from the CO_2 3D fields from CHIMERE at
 175 11:00. For the sake of simplicity in the OSSEs conducted here, since ~~we use only~~-synthetic data ~~are used only~~ and ~~since~~-a
 rather simple modeling of the spaceborne observation ~~is used~~, the computation of X CO_2 assumes that the vertical weighting
 function of the CO_2 column-averaging (kernel) is vertically uniform. For a given model pixel at latitude lat and longitude
 lon , X CO_2 is thus computed from the vertical average of the CO_2 mole fractions simulated by the model:

$$\text{XCO}_2(lat, lon) = \frac{\int_{P_{top}}^{P_{surf}} \text{CO}_2(lat, lon, P) dP + \overline{\text{CO}_2}(P_{top}) P_{top}}{P_{surf}(lat, lon)}, \quad (1)$$

180 where P designates the atmospheric pressure, P_{surf} the atmospheric surface pressure and P_{top} (300hPa) the pressure ~~ceiling at~~
~~the top boundary~~ of the model. For pressures lower than P_{top} , we assume that the CO_2 concentrations equal the horizontally-
 average ~~of the~~ top-level mixing ratios in CHIMERE ($\overline{\text{CO}_2}(P_{top})$). Indeed, we do not expect significant spatial gradients of
 CO_2 over the simulation domain in the upper atmosphere. This is supported by the lack of signal in our simulations of the
atmospheric signatures of the surface fluxes in the upper layer of the model.

2.1.2. X CO_2 pseudo-data sampling and error

185 As detailed in section 2.3 below, the OSSE framework of the inversions requires the location and time of the individual
 X CO_2 data, and the associated error statistics, but not the explicit values of the synthetic observations themselves. In this
 study, we consider pseudo satellites with a Low Earth Orbit (LEO) whose altitude and inclination parameters are similar to
 the ones of the A-Train (705 km and 98.2° 98.2° and 705 km respectively, (Parkinson et al., 2006)). The satellite observations
 are assumed to ~~happen-occur~~ at 11:00 in the morning. Successive tracks ~~for-of~~ a single satellite on this orbit are distant by
 190 ~~about~~ ~ 25 degrees. However, we do ~~neither~~ study the potential of a specific satellite nor that of a constellation of such

LEO satellites depending on ~~of~~ their number. ~~Furthermore, F~~his study focuses on ~~statistics-of-the-imagery~~results at the scale of 6-hours. Therefore, it ~~will~~ considers single satellite tracks for any day that do not correspond to a specific positioning of a satellite on the chosen orbit: when studying the emissions of the Paris urban area, ~~it follows~~we use the track ~~which-that crosses-is nearly centered on~~ this city every day and various swaths are considered to study the sensitivity of the results to this parameter (section 3.2.2). For the study of ~~results for the multiple-ensemble of~~ sources contained within the 2-km sub-domain (section 3.3), ~~it follows~~we use a track centered over Belgium ~~every day~~ and ~~we use~~ a 900-km swath to ensure a full coverage of the plumes from these sources (the sensitivity to a realistic range of swath widths is not investigated in this second set of analysis).

Our OSSEs assess the impact on the inversion results of the measurement noise ~~from~~of the satellite instrument only, ignoring the errors associated to the radiative transfer inverse modeling for the retrieval of the XCO₂ data from the radiance measurements (Buchwitz et al., 2013a; Broquet et al., 2018), and in particular any “systematic error”. The errors on the XCO₂ data at the spatial resolution of the measurements are thus assumed to be Gaussian, unbiased and uncorrelated in space or time. The distribution of the standard deviation (STD) for these errors is also assumed to be uniform and these errors are therefore summarized by a single value of STD (denoted the data precision hereafter).

A large number of scenarios are tested for the observation specifications: the precision on the individual XCO₂ data varies between 0.3 and 2 ppm, and the spatial resolution of the ground pixels can take the following values: 2 km×2 km (longitude × latitude), 2 km×3 km, 3 km×3 km, 3 km×4 km and 4 km×4 km. The reference is a precision of 0.6 ppm and a spatial resolution of 2×2 km². These values are similar to the characteristics of the ~~simulated-simulation of~~ CO2M data used in the study of Wang et al. (2020). When studying the sensitivity of the results over Paris, to the swath of the instrument, ~~the swath is varied~~s from 100 km to 600 km, with a reference value of 300 km.

2.2. CO₂ fluxes

2.2.1. Maps and time-series of anthropogenic emissions and natural fluxes

High resolution maps of anthropogenic emissions are needed to define appropriate point and areas sources to be controlled by the inversion. High resolution maps of anthropogenic and biogenic fluxes are also needed to distribute the controlled local to regional budgets of these fluxes ~~over-on~~ the spatial grid of the CHIMERE model. Finally, such maps are needed to provide insights ~~into~~ the typical budgets of fluxes at the control resolution, and thus to quantify the prior uncertainty in these budgets with a suitable order of magnitude.

The anthropogenic CO₂ emissions are extracted from several datasets compiled by IER (Pregger et al., 2007; Thiruchittampalam et al., 2012). These datasets provide maps of the annual budgets per sectors of anthropogenic activities over different domains, and at different spatial resolutions. We have merged and re-gridded them to derive a map of the annual budgets ~~of emissions~~ over the entire grid of the CHIMERE configuration. The emissions corresponding to France and Germany are extracted from the respective IER national maps for 2005 at a 1-minute resolution, while emissions in Belgium,

Luxembourg and the Netherlands are derived from an IER 1-km product covering Northern Europe for 2005; The IER 5-km resolution map covering the whole Europe for 2008 is used for the emissions over the rest of the domain. The gridded spread sources in these IER maps are interpolated on the CHIMERE grid but the large point sources are relocated as point sources in individual CHIMERE grid cells. We then derive the hourly emission maps from the yearly-annual emission maps by applying the convolution of IER typical temporal profiles specific to each country and sector. These profiles include seasonal, daily and diurnal variations of emissions for large sectors such as traffic, power demand, domestic heating or air conditioning (Pregger et al. 2007).

The IER maps for France, Germany, Northern Europe and the whole Europe correspond to annual budgets for years (2005 and 2008) that can be different from one area to the other, and which are different from the year chosen for the atmospheric transport and for the natural fluxes (2016). This could raise some inconsistencies if assimilating real data in the inversion. However, this study is based on OSSEs with some strong simplifications regarding the observation system since the overarching target is a general understanding of the behavior and potential of the inversion. This requires the use of a high resolution and realistic distribution of the emissions in space and time, but not a precise estimate of their amplitudes for a given year.

The land surface natural fluxes are derived from 8-km resolution simulations made with the Vegetation Photosynthesis and Respiration Model (VPRM) model for the year 2016. This prognostic model delivers hourly values of Net Ecosystem Exchange (NEE) by assimilating satellite and meteorological data (Mahadevan et al., 2008). The values of these biogenic fluxes NEE values are interpolated over the CHIMERE area at the hourly time scale. Natural ocean fluxes are ignored.

2.2.2. Controlled areas

The resulting hourly maps of anthropogenic CO₂ emissions for spread sources and large point sources are decomposed spatially to define the areas for which hourly emission budgets are controlled by the inversion: large point sources, cities, remaining parts of regions where from which point sources and cities have been extracted (covering diffuse emissions only), and full regions when point sources and cities are not controlled separately but altogether with the diffuse emissions. Hourly budgets of the natural fluxes are controlled for full regions only, the regions used for the control of anthropogenic and biogenic fluxes being identical.

The definition of the regions is done considering the whole domain. It corresponds to administrative regions of France, Belgium, the Netherlands, Luxembourg and Germany, and to three additional large "regions": the United Kingdom, Switzerland and the rest of the domain. This subdivision results in 67 regions (Fig.1). These 67 regions correspond to the spatial resolution of the natural fluxes in the inversion.

Point sources and cities are controlled individually in the 2-km resolution part of the CHIMERE grid only (section 2.1.1., Fig. 1). In the 39 regions entirely comprised within this sub-domain, we individually control the 84 point sources (e.g., factories, power plants...) whose annual emissions are larger than 0.2 MtonC.yr⁻¹. The maps of the remaining emissions (excluding these point sources) in each of these 39 regions are then processed to extract large urban areas to be controlled

independently, ensuring at least one controlled urban area per region, and that no controlled urban area overlaps two different regions. An algorithm of pattern recognition has been designed for such an extraction, with the idea that urban areas correspond to clusters of adjacent high emitting pixels (also followed by (Wang et al., (2019)). After having applied a Gaussian filter to smooth the spatial distribution of the emissions, the large urban areas are defined by a label-connecting algorithm (Stockman et al., 2001; Shapiro et al., 2000) which defines-identifies the clusters of adjacent points whose emissions are above a predefined threshold. As the density and extension of cities considerably-vary considerably amongst the different regions, the parameters of the pattern recognition algorithm, i.e. the standard deviation of the Gaussian filtering and the emission threshold, are different for each region to ensure that each region contains at least one controlled urban area (Fig. 2). As a result, we identify 152 control urban areas within the 2-km resolution sub-domain. They are characterized by a wide range of budgets and spatial spread of their emissions, the annual budgets ranging between ~ 0.07 MtonC.yr⁻¹ and ~ 9.9 MtonC.yr⁻¹ (with a mean and a standard deviation of ~ 0.8 MtonC.yr⁻¹ and ~ 1.5 MtonC.yr⁻¹ respectively), and areas ranging from ~ 8 km² to ~ 2400 km² (with a median value of ~ 240 km²).

The remaining emissions, after having extracted the large point sources and urban areas in the 39 regions, are considered to be diffuse and called hereafter the “countryside” emissions. The inversion controls their budgets in each region. The analysis of the results at the regional scale for these 39 regions will consider either the countryside emissions only (i.e. focusing on the individual control variables), or their aggregation with the emissions from the point sources and cities within the same region (i.e. considering the full geographical extent of the regions). The inversion also directly controls the total budget of the emissions for the 28 regions that are not fully comprised in the 2-km sub-domain (most of these 28 regions do not actually-overlap this sub-domain at all). Overall, the control of countryside or total regional emissions adds 67 controlled areas (corresponding to the 67 regions) for the anthropogenic emissions so that the inversion controls the hourly budgets of anthropogenic emissions for 303 areas (84 point sources, 152 urban areas and 67 countryside or regional areas) and the hourly budgets of natural fluxes for 67 areas.

2.3. Analytical flux inversion

2.3.1 Theoretical framework

The inversion system follows a traditional analytical inversion approach based on the Bayesian formalism and assuming that error statistics follow Gaussian errors-statisticsdistributions (Tarantola et al., 1987; Broquet et al., 2018). The system controls factors that scale the hourly budgets of the different control areas for the anthropogenic and biogenic fluxes defined in section 2.2.23. It also controls a single scaling factor applied to the CO₂ field used to impose the initial, lateral and top CO₂ boundary CO₂-conditions (BC)-of the model, since such boundary conditions generally bear important large-scale uncertainties that can impact the estimates of sources within the domain (Broquet et al., 2018). In the OSSEs of this study, more-specifically-for-each-day-in-March-or-May-2016, the inversion periods for each day in March or May 2016 cover the 6 hours (5:00-11:00) before 11:00, when the satellite observations are supposed to be made. The number of control variables

(2221 = 370 controlled areas × 6 time slots + 1 control variable for the **BC**boundary conditions), is sufficiently small for analytical calculations to solve the inverse problem analytically. However, building the matrix **H** which that encompasses the atmospheric transport operator **H** and that is described below (section 2.3.2) requires a large computational burden.

For a given inversion period, we define the control vector **x** as the set of controlled scaling factors for the hourly emission flux budgets and the boundary conditions. The prior uncertainty in **x** is assumed to be follow a Gaussian distribution and to be unbiased. It and is thus characterized by the uncertainty covariance matrix **B**.

In this study, the observation vector **y** is defined by the XCO₂ concentrations in the transport model horizontal grid cells sampled by the observations. The simulation of **y** based on a given estimate of **x** is given by the linear observation operator **H**: **x** → **y**=**Hx**, which chains three operators. The first operator **H**_{distr} distributes the controlled hourly budgets of emissions in space within the controlled areas, and provides the spatial and temporal mapping of the boundary conditions whose scaling factor is controlled by the inversion. The second operator **H**_{transp} is the atmospheric transport from the emissions and the boundary conditions to the full CO₂ and XCO₂ fields. Finally, the third operator **H**_{sample} performs the XCO₂ sampling at the location of the XCO₂ data (section 2.1.2). Differences between **Hx** and observed values for **y** arise due, on the one hand, to uncertainties in **x**, and, on the other hand, to the combination of errors in the observation operator and in the observation data that are altogether called altogether “observation errors”. The errors from the observation operator are strongly associated with the atmospheric transport model errors (Houweling et al., 2010; Chevallier et al., 2010), but also with the discretization and spatial resolution of the transport and inversion problems, which raise representation and aggregation errors (Kaminski et al., 2001; Bocquet et al., 2011). Assuming that they are follow Gaussian and unbiased distributions like the prior uncertainties, these observation errors are fully characterized by the observation error covariance matrix **R**. The **H**, **B** and **R** matrices must be explicitly estimated in the analytical inversion framework (section 2.3.2 and 2.4).

The Bayesian theory (Tarantola et al., 1987) states that the statistics of the knowledge on **x** knowing i) the prior estimate of **x**, ii) the observed values for **y** and iii) **H** as a link between the **x** and **y** spaces, follow a Gaussian and unbiased distribution.

The uncertainty in such a posterior estimate is thus fully characterized by the posterior uncertainty covariance matrix **A** given by:

$$\mathbf{A} = [\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}]^{-1} \quad (2)$$

The analysis of **A** and its comparison to **B**, aggregated or not over different spatial and temporal scales, are the critical diagnostics in this study to assess the potential of inversions assimilating XCO₂ images. The score of uncertainty reduction for a given flux budget is a common indicator for evaluating the performance of an observation system. It is defined as the relative difference between the STD of the prior (σ_{prior}) and posterior (σ_{post}) uncertainties in this flux budget ($UR = 100 \times (1 - \frac{\sigma_{post}}{\sigma_{prior}})$). If the assimilation of satellite observations perfectly constrains a given flux budget, the corresponding UR equals 100%. If this assimilation does not provide any information on the flux budget, UR equals 0%.

2.3.2. Building the observation operator matrix \mathbf{H}

320 The analytical inversion system is essentially built on the explicit computation of $\mathbf{H}=\mathbf{H}_{\text{distr}}\mathbf{H}_{\text{transp}}\mathbf{H}_{\text{sample}}$. The different columns of \mathbf{H} correspond to the signatures (or “response functions”) in the observation space of the different control variables, *i.e.* of the different hourly emissions for each control area, and of the boundary conditions. They are computed by applying the sequence of operators $\mathbf{H}_{\text{distr}}$, $\mathbf{H}_{\text{transp}}$ and then $\mathbf{H}_{\text{sample}}$ to each control variable set to 1, keeping the others null (Broquet et al., 2018). $\mathbf{H}_{\text{distr}}$ is defined based on the flux maps detailed in sections ~~2.2.2.1. and 2.2.2.~~ $\mathbf{H}_{\text{transp}}$ corresponds to the CHIMERE model and to the vertical integration of CO_2 into XCO_2 presented in section 2.1.1, while $\mathbf{H}_{\text{sample}}$ corresponds to the sampling, on the transport model grid, of the simulated XCO_2 values according to the spatial distribution of the pseudo-observations (section 2.1.2). A generalized \mathbf{H} is actually stored for the analytical inversion system to anticipate any option for $\mathbf{H}_{\text{sample}}$, by recording the full CO_2 and XCO_2 fields from the application of $\mathbf{H}_{\text{distr}}\mathbf{H}_{\text{transp}}$ to each control variable, *i.e.* the full CO_2 and XCO_2 signatures of each control variable.

330 2.4. Practical implementation of the OSSEs

While, in principle, \mathbf{R} should characterize both the ~~errors in~~ XCO_2 data ~~errors~~ and the errors from the observation operator \mathbf{H} , this study focuses ~~only~~ on the impact of the observation sampling and errors only. It ignores the errors from the observation operator. Moreover, the observation errors are restricted to the measurement noise which is uncorrelated in space and time as detailed in section 2.1.2. The different \mathbf{R} matrices used for the OSSEs (depending on the observation sampling and noise) are thus all diagonal. The errors on the individual pseudo-observations are described by an uniform precision (σ_{XCO_2}) depending on the chosen satellite configuration (section 2.1.2). However, the observation vector \mathbf{y} is defined by the transport model grid rather than by the precise location and coverage of the data. ~~Therefore,~~ ~~the~~ diagonal elements of \mathbf{R} follow ~~this~~ the aggregation of n_{obs} pseudo observations with uncorrelated errors (where n_{obs} is potentially greater than 1) within each model grid cell corresponding to an element of \mathbf{y} , so that the resulting STD of the errors for this element is given

340 by $\frac{\sigma_{\text{XCO}_2}}{\sqrt{n_{\text{obs}}}}$.

Prior estimates of anthropogenic emissions and biogenic fluxes are generally provided by inventories and ecosystem model simulations such as those used here in $\mathbf{H}_{\text{distr}}$ to distribute the fluxes at high resolution. \mathbf{B} should characterize uncertainties in such products and is thus ~~be~~ set with values corresponding to typical relative uncertainties in the budgets from the maps detailed in section 2.2.1. Prior estimates of the boundary conditions for regional inversions are usually interpolated from large scale analysis or inversions. Such products can bear significant large-scale errors at the boundaries of Europe (Monteil et al., 2019). We reflect it by setting in \mathbf{B} the standard deviation of the prior uncertainty in the scaling factor for the boundary conditions (see below). When constructing the \mathbf{B} matrices in all our OSSEs, we assume that there is no correlation between the prior uncertainties associated to different controlled emission areas or between these uncertainties and the one associated to the boundary conditions. The spatial correlations of the uncertainties in anthropogenic emission inventories is a complex topic and the current lack of characterization for such correlations led to such a conservative set-up (Wang et al. 2018; Super

345

350

et al., 2020). However, we model the temporal correlations between prior uncertainties in scaling factors associated to different hourly natural or anthropogenic flux budgets of the same controlled emission area by using an exponentially decaying function with a correlation time scale τ (like, for instance, in (Bréon et al., 2015)):

$$\rho_{i,j} = e^{-\frac{|j-i|}{\tau}} \quad (3)$$

355 where j and i are the indices of two corresponding hours. The STD of the prior uncertainties in the scaling factors for the different hourly budgets of the same controlled area are fixed to an identical value σ_{hour} . Finally, we assume that the STD of the prior uncertainties in scaling factors for the 6-hour budgets of the natural fluxes or of the anthropogenic emissions from a controlled area during 5:00-11:00 (to be applied to the budgets from the IER and VPRM products presented in section 2.2.1 and used to build $\mathbf{H}_{\text{distr}}$) is fixed to a value σ_{Budget} that is the same for all control areas: typically 50% or 100% of the 6-hour budgets. The STD of the prior uncertainties in the scaling factors for the hourly budgets of the controlled areas σ_{hour} are then derived based on these different assumptions. The reference parameters for \mathbf{B} are fixed to $\tau = 3$ hours and $\sigma_{\text{Budget}} = 50\%$. Despite the differences between the temporal variations of the hourly emissions from one control area to the other, or between natural and anthropogenic fluxes in $\mathbf{H}_{\text{distr}}$, these STD show a negligible variation of less than 1% and when considering the reference set-up for \mathbf{B} , $\sigma_{\text{hour}} \sim 65\%$. The sensitivity of the inversion to the values of σ_{Budget} and τ is assessed in section 3.2.3. Finally, we use 1% for the STD of the prior uncertainty in the scaling factor associated to the boundary conditions (i.e. typically an uncertainty of ~ 4 ppm in the average boundary conditions). This value is quite pessimistic, but some tests in which this value was varied (not shown) demonstrate a very weak sensitivity of the results for the fluxes to this parameter.

3. Results

3.1. High-resolution simulations of XCO₂

375 Previous sections documented how, for each 6-hour period, the inversion system exploits the simulated XCO₂ fields at 11:00 to constrain each hourly budget of the anthropogenic or natural fluxes of the controlled areas between 5:00 and 11:00. The CHIMERE full XCO₂ simulations between 5:00 and 11:00 with the anthropogenic emissions, natural fluxes and/or domain boundary conditions detailed in sections 2.1 and 2.2 are used in this section to compare the overall signatures of these components and of the controlled areas, and to discuss their overlapping. Figure 2b shows the XCO₂ signatures of all the anthropogenic emissions in the domain excepted those from the 84 point sources controlled individually by the inversion in the 2-km resolution subdomain (that are illustrated in Fig. 2a). Figure 2c integrates the XCO₂ produced by the 84 point sources and shows the signature of all the anthropogenic emissions. Finally, Figure 2d displays the superposition of the XCO₂ signatures of all the anthropogenic emissions, of the natural fluxes and of the boundary conditions. For all these figures, XCO₂ values are taken at 11:00 and are provided by simulations between 5:00 and 11:00 on May 23rd which is a day of strong northwest-southeast wind ($\sim 10 \text{ m}\cdot\text{s}^{-1}$ over Paris at 700 m above ground level).

Formatted: Superscript

The strong plumes from the megacities of Paris and London are easily distinguished when considering the signature of anthropogenic emissions in Fig. 2b and 2c, with their amplitude exceeding 0.3 ppm at 100 km downwind of these cities, and with sharp gradients of XCO₂ at their edges. The relative narrowness, extended length and small intensity of the plumes shown in the Figs. 2b and 2c are explained by the magnitude of the wind speed on May the 23rd ($\sim 9 \text{ m s}^{-1}$ over Paris). The characteristics of those plumes vary considerably with respect to the wind speed and the inversion results are strongly impacted by this parameter (section 3.2.1 and Broquet et al. (2018)).

Figures 2b and 2c also show that the overlapping of plumes from urban areas ~~of in~~ Belgium and ~~in~~ the Netherlands produces XCO₂ patterns whose amplitudes are comparable to that of the plumes from Paris and London. ~~However, Due~~ to the urban density of those countries, the level of distinction ~~of between~~ the individual signatures of the different cities is ~~however~~ weak. If we exclude Paris, ~~n~~Northwestern France has a much less dense urban fabric with scattered cities of small ~~extension extents~~. ~~Thus, This~~ sparse distribution allows the relatively weak plumes ~~of from~~ cities ~~emissions~~ to be visible, whereas the more diffuse XCO₂ signatures of the ~~anthropogenic countryside~~ emissions ~~of the countryside~~ do not form any distinguishable patterns (Fig. 2b).

The comparison between Figs. 2b and 2c highlights the plumes from some of the large 84 point sources within the 2-km resolution subdomain (section 2.2.2~~3~~). The amplitude of these plumes can locally reach that of Paris but such an increase above the background occurs on a much smaller extent: ~~for instance~~, the one of the power plant close to Dunkerque ($\sim 51^\circ$, $\sim 2.3^\circ$) in the ~~n~~Northern French coast reaches 0.4 ppm but its width does not extend to more than 5 km ~~for instance~~ (Fig. 2c). The capacity of our high resolution transport model to simulate narrow XCO₂ plumes from point sources or urban areas distinct from ~~that of~~ neighbor or surrounding sources is revealed by the example of several point sources in Belgium as well as that of the oil refinery of Grandpuits (48.59° , 2.94°) whose plume stands out of the large plume from the Paris urban area (Fig. 2c). The 2-km resolution zoom of the model grid allows distinguishing those features which would ~~otherwise~~ be blurred in a coarser resolution transport model.

When including the XCO₂ produced by the natural fluxes and the boundary conditions, identifying the features produced by the anthropogenic emissions is more difficult (Fig. 2d). The atmospheric signatures of Paris, of London and of the high-emitting power plants are hardly differentiated from patterns produced by the boundary conditions and natural fluxes even though they are still visible. The ~~isolated plumes of low amplitudes from~~ scattered cities with small ~~extensions extents~~ and low emission budgets ~~which produce isolated plumes of low amplitudes~~ can hardly be seen. The boundary conditions and the natural fluxes tend indeed to produce signatures whose amplitude is often larger than ~~or at least comparable to that of the signal from~~ the anthropogenic emissions, ~~with which they interfere, or at least comparable, which interfere with and~~ This blurs ~~thise~~ signals of the ~~anthropogenic~~ emissions, especially ~~with those of the~~ when the emissions are diffuse ~~anthropogenic emissions~~. Boundary conditions and natural fluxes are however much more ~~homogeneously~~ distributed ~~homogeneously~~ than the anthropogenic emissions, that are localized over a small fraction of the surface. As a consequence, the boundary conditions and the natural fluxes produce smooth XCO₂ fields (Fig. 2d) while the anthropogenic emissions ~~from sources inside the domain~~ produce heterogeneous fields with ~~smaller and~~ finer structures and sharper gradients (Fig. 2c). Therefore,

the separation between the two types of fluxes could rely on the differences in terms of spatial scales of their ~~spatial~~ ~~atmospheric~~ signatures, or on a precise knowledge of the atmospheric transport patterns.

This first qualitative overview of the atmospheric signatures could imply that the ability to quantify the budgets of emissions ~~from~~ ~~for~~ the two megacities, ~~from~~ ~~for~~ most of the 84 large point sources, and for large regions in the ~~n~~North ~~e~~East should be much larger than for the individual urban areas in most of the domain or for the countryside emissions. However, this diagnostic ~~test~~ relies on a qualitative assessment of Figures 2i. In Section 3.3, we will quantitatively analyze the ~~implications~~ ~~for~~ ~~the~~ inversion results ~~of the differences in terms of amplitude or spatial structure between the XCO₂ signatures as a function of the type of sources.~~

3.2. Potential of the satellite images to monitor the anthropogenic emissions of a megacity: sensitivity studies

This section assesses the performance of our inversion assimilating XCO₂ images to monitor the anthropogenic emissions from the Paris area, as a function of the wind conditions ~~for speed and direction~~, of the XCO₂ observation precision, resolution and swath, and of the configuration of the prior error covariance matrix **B**. Results are relative to the inversion control area that covers most of the Paris urban area (Fig. 2.a.) ~~and whose contours have been delimited by the pattern recognition algorithm described in section 2.2.3.~~ The analysis is based on 62 6-hour~~s~~ inversion tests with satellite images ~~nearly~~ centered on this area for each day of March and May 2016. With the reference 300-km wide swath, such images cover the plumes from the Paris urban area entirely ~~for~~ ~~in~~ most wind conditions (~~{~~Broquet et al. 2018~~}~~). In the following, the wind speed is characterized by its averaged value at 700 m above ~~the~~ ground level over the inversion control area corresponding to Paris and over the period corresponding to the ~~chosen~~ diagnostic ~~test~~: over 5:00-11:00 when analyzing the uncertainty in the budgets of the emissions corresponding to the full 6h-period of inversion, or over the time interval [hh,11:00] when analyzing the uncertainty in the hourly budget of the emissions between the hours hh and ~~++:00~~hh+1.

3.2.1. Impact of the wind speed

A first set of inversions is conducted with the reference values for the precision, resolution and swath of the satellite observation and for the parameters of the prior uncertainty (~~respectively 0.6 ppm, 2×2 km², 300 km, 50% and 3 hours~~). These inversions are applied to 12 different days in March 2016 which present a range of average wind speeds from 2 to 14 m·s⁻¹ (Fig. 3). We investigate results in winter, when the amplitudes of the biogenic fluxes are low, to mitigate the influence of these fluxes, and of their variability, on the URs ~~fore~~ the Paris emissions. Note that the time profiles modeling the variability of the anthropogenic emissions ignore day-to-day variations (except between week-end and working days) which almost removes the influence of the variability of these emissions when studying ~~only results the month of in~~ March ~~only~~. Results are presented in terms of prior and posterior uncertainties in both the 1-hour and 6-hour budgets of emissions from the Paris urban area. The directions of the wind are predominantly meridional so that the selection of the swath has no impact. The main analysis and conclusions in this section are similar to that of Broquet et al. (2018), ~~and so~~ we ~~thus~~ present them briefly.

Results in Fig. 3 illustrate the fact that larger wind speeds lead to smaller uncertainty reductions for the 6h-emission budgets: On March 3 ~~when~~ the average wind speed is $13.5 \text{ m}\cdot\text{s}^{-1}$; ~~and~~ the UR ~~for~~ ~~with~~ the reference values for the precision, resolution and swath is 74% while on March 10 ~~the~~ (average wind speed $\sim 1.8 \text{ m}\cdot\text{s}^{-1}$ ~~and~~) the UR is ~~of~~ 97% (Fig. 3a). Stronger winds ~~decreases~~ the UR because, due to an increased atmospheric dilution, the amplitude of the city plume is smaller which decreases the signal-to-noise ratio for the inversion. However, when considering the UR for hourly emissions, this rule may not apply for wind speeds lower than $6 \text{ m}\cdot\text{s}^{-1}$. For this range of wind speeds, the posterior uncertainty in ~~individual~~ 9:00-10:00 and 10:00-11:00 emissions increases with decreasing wind speed (Fig. 3b). The inversion system shows difficulties in distinguishing the atmospheric signatures produced by consecutive hourly emissions because these signatures have a significant overlap when the wind speed is low. This explanation is confirmed by the negative correlations found between the uncertainties in consecutive hourly emissions since the magnitude of these negative correlations increases when the wind speed decreases (Fig. 3c). Important negative correlations explain also that for low wind conditions, 6h-emission budgets are better constrained even though hourly emissions can be poorly constrained. The overestimation of some hourly emissions is compensated by the underestimation of other hourly emissions.

The uncertainty reductions for 1h- and 6h-budgets are ~~important~~ ~~high~~ for a large range of wind values: ~~for the range considered in this study, in all the tests here,~~ the URs for the 6h-budgets are above 74% and the UR for the 1h-budgets after 7:00 are above 62% (Fig. 3a). Concerning the 1h-budgets of the 5:00 to 7:00 emissions, the corresponding URs significantly decrease for wind speeds above $10 \text{ m}\cdot\text{s}^{-1}$. ~~In particular,~~ ~~the~~ UR for the 5:00 to 6:00 emission drops below 20% ~~above this value for the wind speed.~~ This behavior is consistent with the fact that the signatures of emissions occurring well before the satellite overpass have been much more diffused through atmospheric transport at the observation time than that of later emissions.

3.2.2. Impact of the precision, resolution and swath of the satellite images

Figures 4a-d show the ~~uncertainty reductions~~ URs for the 6 h-emission budgets of March 3 (with a strong wind), March 10 (with a low wind), May 23 (with a strong wind) and May 27 (with a low wind). This figure ~~is associated to the first analysis made in this section and~~ corresponds to a second set of inversions performed with the range of options for satellite data precisions and resolutions presented in section 2.1.2 but with the ~~observation~~ swath ~~of the satellite track,~~ the relative prior uncertainty in the 6 h-budgets ~~of fluxes,~~ and the correlation length scale for the prior uncertainties in hourly fluxes fixed to the reference values (respectively 300 km, 50% and 3 hours). The sensitivity of the UR to the measurement precision and resolution increases with stronger winds. For example, on May 27 (~~under~~ "weak" wind) and 23 (~~under~~ "strong" wind), the UR increases by 16% and 48% respectively between inversions with 2 ppm precision data and inversions with 0.3 ppm precision data (at the resolution of $2 \text{ km} \times 2 \text{ km}$). For those two days, the UR decreases by 6% and 20% respectively between inversions with $2 \text{ km} \times 2 \text{ km}$ resolution data and inversions with $4 \text{ km} \times 4 \text{ km}$ resolution data (~~for~~ ~~with~~ a precision of 0.6 ppm). The comparison between results on March 3 and 10 confirms such a high sensitivity ~~for~~ ~~under~~ stronger winds. It is related to the fact that the slope of the convergence of the UR towards 100% with better precision and finer resolution is

smaller with low wind speeds, which generate higher UR than high wind speeds. For similar reasons, the sensitivity to the precision decreases at finer resolution, and the sensitivity to the resolution decreases with better precision (Figs. 4a-d).

The comparison between results obtained when doubling the random measurement error of the individual observations and when multiplying by four the value of their spatial resolution provides insights ~~into~~ on the exploitation of the fine scale patterns of the XCO₂ image by the inversion. Indeed, both changes result in doubling the resulting error at coarse resolution, but doubling the random measurement error at fine resolution conserves the capability to exploit information at this fine resolution unlike coarsening the spatial resolution of the image. Figs. 4a-d show that scores of UR with 2 km × 2 km resolution and 2 ppm precision data are extremely close to that with 4 km × 4 km resolution and 1 ppm precision data. UR_s with 2 km × 2 km resolution and 1.2, 1, 0.8 or 0.6 ppm precision data are also similar to UR_s with 4 km × 4 km resolution and respectively 0.6, 0.5, 0.4 or 0.3 ppm precision data. This indicates that the inversions here do not really take advantage of the information on the fine scale patterns of the plume from Paris.

A third set of inversions is conducted to study the sensitivity of the results to ~~varying~~ the width of the satellite swath while keeping all other observation and inversion parameters to reference values. This sensitivity is modulated by the wind conditions: the speed and direction of the wind control the spread and position of the plume and thus the value of the swath which fully covers the extent over which the amplitude of the plume is significant for the inversion, i.e., the value of the swath above which the results do not change any more (Figs. 4e-h). This threshold value of the swath is lower for smaller wind speed. For wind directions across the satellite track, the UR_s for the 6-h-emissions of the Paris area are no longer sensitive to the increase of the swath above a value of 100 km and 400 km for wind speeds ~~superior~~ lower than ~~to~~ 8 m·s⁻¹ and ~~to~~ 9 m·s⁻¹ respectively. The sensitivity to the swath is null (except if considering very low values for the swath of the order of the width of the plume from Paris) for wind directions along the satellite track as we, since ~~considering~~ satellite tracks centered on Paris.

3.2.3. Impact of the definition of the prior uncertainties ~~on~~ the CO₂ fluxes

The prior uncertainty covariance matrix **B** has a strong influence on the scores of posterior uncertainties when its “amplitude” is comparable to, or much larger than, the one of the **H^TRH** matrix (see Eq. 2), i.e. once the prior uncertainties are comparable or much larger than the projection of the observation errors in the control space. The relative prior uncertainty~~ies~~ in the 6 h-emission budgets (σ_{Budget} , section 2.4.), which characterizes the diagonal of **B**, is one of the critical drivers of the relative weight given by the inversion to the prior information and to the observations.

In a fourth set of inversions, we thus analyze the sensitivity of the inversion results for 6-hour~~s~~ emission budgets to σ_{Budget} , with values for this parameter ranging between 0 and 100%. This set of inversions uses the reference values of the observation parameters and for the temporal autocorrelation of the prior uncertainties. Figs. 5a-b shows the corresponding results on March 3 (with-under strong wind) and on March 10 (with-under low wind), to highlight the dependence of this sensitivity to the wind speed. The curves of UR as a function of σ_{Budget} have an inflection point for values around 50%. For low values of σ_{Budget} , the UR is sensitive to this parameter, the posterior uncertainty balancing the prior uncertainty and the

515 projection of the observation error. For large values, the UR converges asymptotically towards 100%, and the posterior uncertainties are dominated by the projection of the observation error (i.e. the posterior estimate of the emission essentially relies on the top-down information from the observations). The observational constraint on the inversion is larger on March 10 than on March 3 since the wind is much lower on the former. As a consequence, the qualitative threshold of σ_{Budget} above which the URs are not much sensitive to this quantity is smaller on March 10 than on March 3: 30% and 50% respectively.

520 These results over Paris suggests an empirical choice of a reference value for $\sigma_{\text{Budget}} \geq 50\%$, in the absence of any factual knowledge about σ_{Budget} . With 50% as a reference value, we focus our analysis of the posterior uncertainties on such the projection of the information from the observations and we nearly neglect the prior information, while keeping an assumption regarding the prior uncertainties that could seem consistent or even optimistic compared to series of assessment of the errors in inventories for cities at daily scale (Wang et al., 2020). However, for other cities, point or area sources with smaller amplitudes, the observational constraint is lower. The relative weight between the projection of the observations and the prior information is then more balanced than for Paris, and the prior uncertainty still has a significant impact on the posterior uncertainties when using $\sigma_{\text{Budget}}=50\%$. In order to study the pure projection of the observation errors, results using $\sigma_{\text{Budget}}=100\%$ will thus be analyzed along with that using $\sigma_{\text{Budget}}=50\%$ in section 3.3.

530 The other important parameter defining the \mathbf{B} matrix in this study is τ (section 2.4.). By construction, the increase of the corresponding auto-correlations for in the prior uncertainties at the hourly scale in \mathbf{B} does not modify the prior uncertainties of in the 6 h emission budgets. However, it could help the inversion crossing the information on different hourly budgets to better constrain the overall budget of emissions. A fifth set of inversions with the reference values for the observation parameters and for σ_{Budget} is conducted to test the sensitivity to τ , with values for this parameter from 0 to 6 hours (0 h indicating that there is no temporal correlation in \mathbf{B} , and 3 h being the reference value), on March 10 and 3. The analysis shows that, actually, the increase of τ hardly impacts the results for the 6 h-budgets (not shown) but significantly changes the results for the hourly budgets (Figs. 5c-d). The auto-correlation brings information about the temporal distribution of the emissions, constraining how the 6h-emission budgets are distributed at the hourly scale. This impact is more significant when the XCO₂ signatures of the hourly emissions overlap, i.e. for hourly emissions between 5:00-7:00 when the wind speed is high and for almost all the hourly emissions when the wind speed is low. But However, this better knowledge about the temporal variations from auto-correlations does not appear to improve the knowledge on the 6 h-budgets.

3.3. Potential of satellite images to monitor anthropogenic emissions at the regional, city and local scales.

545 This section synthesizes the inversion results from at the local scale (for power plants, industrial facilities) to the regional scales over most of the model 2-km resolution model subdomain, using a sixth set of inversions assimilating images that cover this subdomain entirely (satellite observations with a 900-km swath centered on Belgium, (Fig.1). This set of inversions covers all the days of March and May 2016 in order to analyze the impact of the wind speed and of the natural fluxes on the results. The prior relative uncertainty in the 6-hour budgets of the emissions are alternatively set to

$\sigma_{\text{Budget}}=50\%$ and 100% . These inversions use the reference parameters for the observation precision and resolution and for the temporal auto-correlation of the prior uncertainties in hourly emissions (0.6 ppm, $2 \text{ km} \times 2 \text{ km}$ and 3 hours respectively).

Results over most of the 2-km resolution ~~model~~-subdomain using different observation spatial resolutions and precisions will briefly be discussed in section 4.

3.3.1. Overview of the inversion performance

Figure 6 gives a geographical overview of the scores of UR in the 2-km resolution sub-domain. The largest scores of UR for 6-hour budgets are obtained for the mega-cities of Paris and London with ~~a~~-mean values over the two months considered ~~>80% over the two months considered~~. Mean UR can also be $>60\%$ for several cities of Belgium and the Netherlands and for a large number of point sources (power plants and large industrial facilities) within the dense industrial area of western Germany, although these sources are close to each other or to other significant point and area sources.

In a general way, the scores of UR increase with the magnitude of the emissions (Fig. 7). This increase is more important when considering lower emission values due to the asymptotic convergence of the UR towards 100% for high emission values (with a point of inflection for emitting-sources of ~~around~~ $\sim 2 \text{ MtC}\cdot\text{yr}^{-1}$ in the curves of Fig. 7). The increase of the UR as a function of the budgets of emissions is different if considering point or area sources. As expected, the largest URs are obtained for narrower sources like point sources (Fig. 7a) and ~~the~~-cities (Fig. 7b) which generate plumes with smaller extents but larger amplitudes than diffuse countryside emissions. When using $\sigma_{\text{Budget}} = 100\%$, the mean URs are larger than 50% for all point sources and cities with an emission rate larger than $2 \text{ MtC}\cdot\text{yr}^{-1}$, ~~but while for to~~ achieving the same 50% UR, an emission rate of at least $4 \text{ MtC}\cdot\text{yr}^{-1}$ is needed for regional countryside emissions (Fig. 7c). The gap is even larger when using $\sigma_{\text{Budget}} = 50\%$, with mean URs that are systematically larger than 50% for annual emission budgets of point sources and cities larger than $2 \text{ MtC}\cdot\text{yr}^{-1}$, but for annual emission budgets of regional countryside emissions larger than $7 \text{ MtC}\cdot\text{yr}^{-1}$.

When aggregating the results for point sources, cities and countryside emissions at the regional scale, the relative prior uncertainty becomes significantly smaller than the values used for individual sources since we assume that there is no correlation between their uncertainties: the mean prior uncertainty for the regions is then of $\sim 33\%$ when assuming a 50% prior error on the 6 h-budgets of point sources, cities and countryside areas which make these regions. And, the emission threshold above which the URs for the regional budgets are larger than 50% becomes $10 \text{ MtC}\cdot\text{yr}^{-1}$ and $7 \text{ MtC}\cdot\text{yr}^{-1}$ when using $\sigma_{\text{Budget}} = 50\%$ and 100% , respectively (Fig. 7d). These thresholds are larger than the ones corresponding to individual point sources and cities as given above, but the overall performance of the inversion system at the regional scale is better with respect to that of the point sources and cities when analyzing the relative posterior uncertainties: for $\sigma_{\text{Budget}} = 50\%$, the mean value is ~~of~~ 22% for the total regional budgets while it is ~~of~~ $\sim 40\%$ for the point sources and cities budgets (Fig. A1).

The results for the different types of sources are shown for four regions of Belgium in Fig. 8. This figure provides an illustration of the general results seen in Fig. 7. It shows that the URs for emissions ~~from of~~ the largest urban areas (emitting more than $2 \text{ MtC}\cdot\text{yr}^{-1}$) are as high as that for the overall emissions of their respective region although the budgets of emissions from these urban areas are much smaller than that of their regions. As suggested above, smaller prior uncertainties

580 | in the regional budgets lead to similar URs for cities and regions budgets even though the relative posterior uncertainties in
regional budgets are much smaller (Fig. 8b). When comparing point sources and cities which are characterized by the same
prior uncertainty, the relative magnitudes of the URs are determined by the relative magnitudes of the emissions: URs are
thus much higher for the largest urban areas than for point sources and cities that emit much less CO₂. But, the comparison
585 | illustrates that, ~~even though despite lower they have lower emission budgets in these regions~~ amplitudes, cities are better
constrained than countryside areas. This is in agreement with the enhanced capacity of the inversion system to monitor
city~~ies~~ emissions with respect to more diffuse ~~countryside~~-emissions. This figure also qualitatively illustrates the ability of
the inversion system to separate neighbor emission sources: the point source and city of Liège (left blue bar for the region of
Liège in the figure) contained within the region of Liège are characterized by significant URs even though the point source is
590 | within the city of Liege and its plume ~~is~~ completely overlapped ~~with-by~~ the plume from the rest of the city. We will analyze
more systematically and quantitatively the capacity of the ~~framework-inversion~~ to disentangle the ~~information-signals~~
produced by neighbor sources in section 3.3.3.

The URs for the 6 h-emission budgets show an important variability over the 62 inversion days as illustrated in Fig. 7. When
using $\sigma_{\text{Budget}} = 50\%$, the standard deviations of the day-to-day variations of the URs for the point sources, cities and
595 | countryside areas, are on average, equal to ~12%, ~8.3% and 12.2%, respectively. These values are important with respect to
the temporal mean of the values of UR (26%, 16% and 27% when averaging across all the point sources, cities and
countryside areas respectively). These variations are associated to variations in the wind speed at the daily scale as ~~it was~~
evidenced for the Paris case in section 3.2.1. However, when considering results for the months of March and May together,
they are also ~~influeened-driven~~ by the time profiles of the anthropogenic emissions that are characterized by a strong
600 | decrease of emissions between March and May due to the reduction of residential heating. Moreover, the UR variability is
also determined by ~~that of the uncertainties in~~ the natural fluxes which are also very different from March to May. ~~The~~
~~natural fluxes; havewith~~ large negative amplitudes in May when they are dominated by the primary production and smaller
positive amplitudes in March when they are mostly restricted to the heterotrophic respiration. ~~Using constant prior relative~~
~~uncertainties in the natural fluxes (as for the anthropogenic emissions) yields large absolute uncertainties in May and low~~
605 | ~~absolute uncertainties in March.~~ Furthermore, as the primary production related to photosynthetic processes is mostly driven
by the radiative forcing and then by the daily variation of the cloud cover, natural fluxes ~~and their prior uncertainties~~ are also
characterized by a strong day-to-day variability during the month of May whereas in March, they are not because of a weak
day-to-day variability of heterotrophic respiration. Cross sensitivity studies comparing the influence of the above drivers (not
shown), indicate the predominant influence of the daily variability of the wind speed on the variability of UR for the
610 | anthropogenic emissions estimates, for most sources. This conclusion should however be nuanced for some regions and
countryside areas where the scores of UR for the anthropogenic emission estimates is ~~highly~~ impacted by the inversion of the
natural fluxes, and thus by the variability of these fluxes, during the month of May (see section 3.3.2. below).

3.3.2. Impact of the uncertainties in the biogenic fluxes

615 The analysis of XCO₂ patterns produced by the different CO₂ fluxes (section 3.1) suggests that ~~the~~ large signatures of the biogenic fluxes in May could impact the monitoring of the anthropogenic emissions. In order to weigh the impact of the uncertainties in biogenic fluxes, we conduct experiments where these uncertainties are ignored. In these experiments, the mean UR_s for the budgets of the regional and countryside anthropogenic emissions in May ~~are~~ equal to ~31% and ~41% respectively (using $\sigma_{\text{Budget}}=50\%$). When accounting for uncertainties in biogenic fluxes, ~~the~~ mean URs decrease down to ~21% and ~31% respectively (Figs. 9c-d). This reveals some difficulty of the inversion system to separate countryside emissions from biogenic fluxes which is also illustrated by the important anti-correlations (-40% on average) between the corresponding ~~posterior~~ uncertainties. During May, the smaller amplitudes and rather diffuse nature of countryside emissions with respect to natural fluxes (~~data~~-not shown), and the overlapping of their atmospheric signatures (Fig. 2) explain why the inversion system only has a limited ability to distinguish the countryside emissions. Oppositely, during March, the smaller amplitudes of natural fluxes compared to countryside emissions (~~data~~-not shown) explains why the inversion can better filter the signature of countryside emissions from that of natural fluxes.

620 Contrarily to the UR_s for countryside emissions and regional budgets, the UR_s for the point sources and cities are hardly impacted by the uncertainties in biogenic fluxes during the month of May, even when the emission budgets of these sources are smaller than 1 MtC·yr⁻¹ (Figs. 9a and b), and even though these budgets are quantitatively lower than the absolute value of the regional budgets of biogenic fluxes. Consistently, the ~~posterior uncertainty in the inverted 6-hour budgets of the~~ emissions of cities or point sources is weakly correlated with that in the 6-hour budgets of biogenic fluxes in their respective region (-2% and -5% on average for the cities and point sources respectively). Therefore, the visual inspection of Fig. 2 may wrongly suggest that the plumes of the smallest point sources and cities controlled individually can hardly be separated from the signature of biogenic fluxes. The differences in terms of spatial scales of the atmospheric signatures appear to be the main driver of the skill of the inversion ~~for to separate~~ anthropogenic sources from biogenic fluxes.

635 3.3.3. Separation of the different anthropogenic emission sources

In order to estimate the ability of the inversion system to ~~separately~~ monitor anthropogenic sources whose atmospheric signals overlap ~~separately~~, we focus on pairs of sources contained within ~~a the~~ same region and we assess whether the sum of the variances (Var) associated to the inverted emissions for each source is comparable with the absolute value of their covariance (Cov), which is nearly systematically negative as a result of the uncertainty in the spatial attribution of emitted CO₂ to individual sources in the inversion. This criterion means, if the covariance is negative, that the variance associated to the ensemble of the two sources is much smaller than the sum of the variances associated to each source, ~~given since the mathematical expression governing that~~ the variance of the sum of two random variables X_a and X_b ~~is given by:~~ $\text{Var}(X_a+X_b)=\text{Var}(X_a)+\text{Var}(X_b)+2*\text{cov}(X_a,X_b)$. In terms of inversion, this case describes the situation when a pair of

Formatted: Font: Not Italic

Formatted: Font: Not Italic

645 sources is much better constrained than each of its individual sources ($\text{Var}(X_a+X_b) \ll \text{Var}(X_a) + \text{Var}(X_b)$), *i.e.* when the inversion system does not entirely manage to disentangle overlapping signals and to constrain independently each source.

650 Figure 10 indicates that in general, the system has an overall-good ability to ~~independently~~ constrain independently the different sources within a given region: point sources, cities, countryside areas, natural fluxes. In this figure, indeed, the number of pairs of sources that are characterized by an important negative covariance term ($2 \cdot \text{cov}(X_a, X_b)$) compared to the sum of the individual variances ($\text{V}(X_a) + \text{V}(X_b)$) is much lower than the number of pairs of sources that are characterized by a relatively small covariance term. Only 20 pairs of sources out of a total of 890 pairs show a covariance term that is larger than 25% of the sum of the variances (squares below the blue line in Fig. 10a). Most pairs for which the inversion system may be unable to distinguish individual sources in a completely independent way consist in a point source and its surrounding urban or countryside area. Only one case consists in two point sources that are located close to each other near Karlsruhe in Germany and, no cases consist in pairs of urban/urban or urban/countryside areas. The 20 cases of less separable pairs of sources identified above could be associated to situations for which the inversion system could have difficulties to independently monitor two sources but this conclusion should be nuanced when all terms (sum of the variances and covariances) are small: for example, four pairs of sources characterized by important negative correlations for example show an UR for each individual sources which is larger than 50% (squares in the top-right quadrant of the Fig. 10b). For these cases, the problem of distinction between the two sources applies to a moderate residual posterior uncertainty, and the inversion still gets a relatively precise estimate of the emissions.

660 Moreover, a low covariance between ~~pair of~~ sources can often be explained by the lack of constraint on their total budget or on one of the two sources, rather than by a good separation between the two sources. ~~For Examples of:~~ pairs of sources with an important UR for their total emissions and a small covariance while only one individual source in each pair is well constrained- ~~are can be seen in~~ represented by the top-left and bottom-right ~~of pairs of sources~~ in Fig. 10b, (where the UR for the total is larger than 75% while one of the sources shows an UR lower than 20%). Pairs of sources characterized by relatively low values of covariance but small values of UR for both the total and the two sources are identified in the top-right quadrant of Fig. 10a and the downbottom-left quadrant of Fig. 10b. These pairs are characterized by low individual sources and total emissions whose amplitudes are below the thresholds required by the inversion system to produce reliable estimates (see section 3.3.1.).

670 4. Discussion and conclusion

We have presented a new and comprehensive high-resolution atmospheric inversion system to assess the potential of the satellite imagery of XCO₂ for the monitoring of anthropogenic emissions from local to national scales. This system has been designed to deal with a wide range of different sources in terms of emission amplitude, distribution and spatial scales and to account for the overlapping between the signatures from different anthropogenic sources and natural sources and sinks. To cover the local and regional scales while mitigating computational costs, this inversion system is based on an atmospheric

transport model with a zoomed grid whose horizontal resolution ranges from 2 km in the most resolved domain to 50 km at its edges. The area of highest horizontal resolution (2 km) encompasses regions where the hourly emission budgets of a large ensemble of cities and point sources are inverted separately for their hourly emission budget, i.e. that of Northern France, Belgium, the Netherlands, western Germany, Luxemburg and London. Urban areas are delimited using a clustering algorithm and large industrial plants are selected based on an emission threshold. In total, the system controls the hourly budgets of 303 different anthropogenic sources (plants, cities or regions), the hourly budgets of biogenic fluxes for 67 regional budgets of biogenic fluxes regions and the a scaling factor applied to the 6-hour average initial and lateral and top boundary conditions for each 6-hour inversion window.

Formatted: Font: Italic

685 *A performance simulation tool*

The analytical formulation ~~allows easily testing~~makes it easy to test different scenarios ~~for of~~ observation systems that monitor anthropogenic emissions but, in this study, the OSSEs with the inversion framework have been performed in order to investigate the potential of the satellite imagery of XCO₂ from helio-synchronous orbits only, depending on instrumental configurations or other factors such as the wind or the amplitude of the emissions. Such a spaceborne imagery similar to CO2M may become the critical component of operational atmospheric inversion systems for the monitoring of CO₂ anthropogenic emissions. These operational inversion systems will likely have to jointly assimilate data from many of the existing satellite missions, from ground-based networks and from the planned spectro-imagers. Moreover, in addition to CO₂ concentrations, co-emitted pollutants and radiocarbon should ~~as well~~ be assimilated as well. Ultimately, OSSEs would have to integrate all these components of the observation system (Ciais et al., 2015; Pinty et al., 2017). However, we need OSSEs considering the spaceborne imagery of XCO₂ as a stand-alone system to determine the instrumental parameters ensuring that this imagery can bring emission estimates with sufficient coverage and accuracy so that it can serve as a backbone for ~~such~~ the ~~emission~~ operational monitoring of emissions. Furthermore, in a context where there is a lack of ground based and spaceborne networks that are suitable for the monitoring of CO₂ anthropogenic emissions, these OSSEs can help better understanding the needs in terms of complementary observation components.

The analytical inversion system built for this study allows testing an important set of observational parameters and situations. Once the atmospheric transport functions associated to the different ~~emissions~~ sources have been computed, the derivation of the posterior uncertainties is fast and results can be delivered for a wide and detailed range of specifications on the spatial resolution, precision (focusing here on the errors in XCO₂ from the instrumental noise) and swath of the satellite XCO₂ images. Extending the study of Broquet et al. (2018), the sensitivities of the inversion results to these three parameters have been tested on the example of the inversion of the emissions of Paris (section 3.2.2.). We varied these parameters without accounting for current limitations in space technologies (or from cost issues) which imposeing potential trade-offs between their configurations. The impact of the wind speed on the determination of the 6-hours and hourly budgets of the emissions

is analyzed by assimilating images over different days of March and May. Finally, the flexibility of the inversion system also allowed testing the impact of the way ~~a priori errors/prior uncertainties~~ on the fluxes are prescribed.

The inversion framework produces curves of sensitivity that give several qualitative and quantitative insights into the optimal configurations of the satellite imagery: ~~If we consider the monitoring of the Paris emissions, the impact of the swath on the results is limited for swaths larger than 400 km (section 3.2.2 and Fig. 4). As the megacity of Paris produces plumes whose intensity and extension are amongst the highest with respect to anthropogenic emissions sources in Western Europe, this supports the use of a swath close to 400 km in order to remove the influence of this instrument parameter on the ability to catch the full extent of plumes from targeted sources in the center of the satellite field of view.~~

The study of the UR for the Paris emissions indicates (i) that a $4 \text{ km} \times 4 \text{ km}$ spatial resolution is sufficient, provided that the precision at this spatial resolution is very high, and (ii) that the inversion hardly makes use of patterns at finer spatial resolution. However, as discussed in section 3.2, that might come from the relatively large extent of the plume from Paris, and from ~~their~~ distances ~~from-between this plume and~~ other major plumes. We have thus extended the analysis of the sensitivity to the observation spatial resolution and precision to all the days of March and May and to the ensemble of controlled regional, city and local sources (as for the analysis in section 3.3) ~~using-testing~~ a reduced set of resolutions and precisions ~~to be tested~~. The results are displayed in Fig. 11. This experiment confirms that results for point sources and narrow cities are more sensitive to the availability of the information at the reference resolution of 2 km than that for Paris.

The UR for point sources and cities emitting less than $2 \text{ MtC}\cdot\text{yr}^{-1}$ is larger with a spatial resolution of $2 \text{ km} \times 2 \text{ km}$ and a precision of 1.2 ppm than with a spatial resolution of $4 \text{ km} \times 4 \text{ km}$ and a precision of 0.6 ppm, (i.e. with the same precision at 4 km resolution but without information about the patterns at scales finer than 4 km). However, the differences are not really significant, and Fig. 11 tends to confirm that a spatial resolution of $4 \text{ km} \times 4 \text{ km}$ would be fine if achieving a very good precision: typically, 0.3 ppm if willing to get the results from the reference configuration of the observation with 2 km resolution and 0.6 ppm precision. With a $2 \text{ km} \times 2 \text{ km}$ resolution, the mean URs for all type of emissions would increase on average by $10.1 \pm 1.5\%$ if the precision increases by a twofold factor (from $\sigma = 1.2 \text{ ppm}$ to 0.6 ppm and from 0.6 ppm to 0.3 ppm). While this represent a dramatic increase of the UR when changing the precision from 1.2 ppm to 0.6 ppm, the relative impact is smaller from 0.6 to 0.3 ppm and the reference value of 0.6 ppm precision at 2 km resolution appear to be a balanced option. Finally, the results for the Paris case point out that, at the reference precision of 0.6 ppm, the performance of the inversion system would be nearly equivalent at the resolutions of $2 \text{ km} \times 2 \text{ km}$ and $2 \text{ km} \times 3 \text{ km}$ (Fig. 4). The spatial resolution could thus be relaxed to $2 \text{ km} \times 3 \text{ km}$ compared to the reference resolution without degrading much the precision of the inversions.

If we consider the monitoring of the Paris emissions, the impact of the swath on the results is limited for swaths larger than 400 km (section 3.2.2 and Fig. 4). As the megacity of Paris produces plumes whose intensity and ~~extension-extension~~ are amongst the highest with respect to anthropogenic emissions sources in Western Europe, this supports the use of a swath close to 400 km ~~in order to remove the influence of this instrument parameter on the ability to catch~~ ~~to ensure that~~ the full extent of plumes from targeted sources ~~nearin~~ the center of the satellite field of view ~~is caught~~. However, a narrower swath

would be sufficient for the large majority of cities which emit less than Paris, and thus have shorter plumes. This would impact the monitoring of the large megacities emitting more than 10 MtC/yr like Paris, which represent ~~around~~~~more than~~ 9% of the emissions from cities and power plants over the globe (Wang et al., 2020), but ~~section 3.2.2 shows that~~ this impact is limited for Paris in most meteorological conditions as long as ~~it~~ ~~the swath~~ is larger than 250 km (90% of the cases for the 62 days of inversion, not shown). A general conclusion from these results is that the swath is a less critical parameter than the pixel precision and resolution.

Overall, these results support the reference configurations for the pixel resolution and precision and for the swath width that broadly correspond to the current ones for the CO2M mission, even though they suggest a relaxation of the spatial resolution to 2 km x 3 km and an extension of the swath up to 400 km. However, since this study does not account for technical and cost trade-offs between these parameters, it cannot provide a full assessment for their combination. -Furthermore, by flying a satellite any day over the areas of interest in the sensitivity tests, we restrained the role of the swath to covering a more or less important portion of the plumes, with a lack of sensitivity when the swath exceeds the plume length or when the wind blows along the satellite track. In practice, a critical role of larger swaths (like that of ~~using~~ more satellites in a constellation) is to increase the number of situations for which the plume from a given city can be seen, which is ignored in this study. Finally, the spatial resolution could also play a role when accounting for cloud cover since it could help ~~to~~ ~~increase~~ ~~ing~~ the spatial coverage. A more realistic simulation of the observation sampling by specific missions or constellations, with one or several satellites following real tracks and including the impact of cloud coverage ~~and large aerosol loads~~, and some realistic technical constraints between the parameters, is needed to fully assess the right balance between higher precision, finer spatial resolution and larger swaths.

Ability to monitor anthropogenic emissions at the regional, city and local scales

At first, a qualitative overview of the atmospheric signatures could imply that the ability of the inversion system to quantify the budgets of emissions from most of the 84 large point sources and from the two megacities of Paris and London should be much larger than for smaller ~~individual~~-cities, and that countryside emissions should hardly be constrained by the inversion (section 3.1). ~~However,~~ ~~the~~ quantitative ~~analysis of the results from the OSSEs analysis~~-described in section 3.3. ~~however~~ shows that the capacity of detection of the inversion system is equivalent for point sources and cities: uncertainties ~~on-in~~ both type of sources ~~emitting more than 2 MtC·yr⁻¹~~ are reduced on average by more than 50% ~~above an emission threshold of 2 MtC·yr⁻¹~~ when the resolution and precision of the satellite data are equal to 2 km × 2 km and to 0.6 ppm respectively, and when the prior uncertainties ~~on-in~~ the fluxes have been set to 50% of the 6 h-budgets (section 3.3.1). With the same parameters, the threshold ~~on the level of emissions to get such an UR~~ is much higher (~10 MtC·yr⁻¹) for ~~the anthropogenic emissions of~~ the countryside areas and ~~for~~ the whole regions. ~~However~~ ~~but~~, for these ~~types~~ of sources, the overall ~~level of~~ performance of the inversion system is still comparable to that for point sources and cities: for the 67 regions considered in this study, the mean URs for the total emissions and for the countryside emissions are of 37% and 27.4% respectively

(section 3.3.1.). The ~~relatively~~ lower performance of the inversion system to monitor the countryside and regional anthropogenic emissions could be partly related to the impact of the biogenic fluxes on the determination of these types of emission sources contrary to the point sources and cities (section 3.3.2). ~~Beside these sensitivities to the emissions source amplitudes and to the level of uncertainties in biogenic fluxes, the variations in uncertainty reductions~~ URs are also driven by the wind (sections 3.2.1 and 3.3.1) and by the potential loss of part of the atmospheric signatures of sources at the edges of the satellite swath (section 3.2.2).

We have also ~~addressed~~ ~~investigated~~ the capacity of the inversion system to deal with the overlapping of XCO₂ plumes produced by nearby sources (section 3.3.3). As satellite images give snapshots of the XCO₂ distribution at a given time (11:00) and as atmospheric transport mixes the contribution of the different sources over the 6 hours before the satellite overpass, it is important to assess the ability of the inversion system to disentangle the information coming from different sources at different times. With the hypothesis adopted in this study, most point sources and cities can be ~~independently~~ monitored ~~independently~~ by the inversion system. This ability could be partly related to the high-resolution modeling that allows describing and catching the fine scale patterns of the XCO₂ signatures of the point sources even if the inversion is weakly sensitive to patterns at scales <4 km (Fig. 11).

Beyond the capacity to monitor separately the emissions of point sources, urban and countryside areas, the additional benefit of ~~separately~~ controlling ~~separately~~ these emissions within a given region is the mitigation of the ~~so-called~~ aggregation errors (~~Kaminski et al., 2001~~) when inverting the total budget of a region. In analytical inversions, these errors arise when a control of the emissions at a too coarse resolution limits too much the ability to fit the actual spatial distribution of the emissions and of the concentrations. These errors are evidenced for our study by inverting regional budgets without considering any internal separation into cities, point sources or countryside areas in the control vector: ~~with such a configuration~~, the results are ~~indeed~~ significantly and wrongly more optimistic than the ones obtained with the inversions that ~~consider~~ ~~use~~ this subdivision (Fig. A2).

800 *An optimistic framework*

This study aims at assessing the projection of ~~only~~ the sampling and observation random noise into the emission estimates. It combines optimistic assumptions that prevent from assuming that the level of posterior uncertainties achieved in the OSSEs here should correspond to that of the inversions with real data. The ability of the inversion system to separately monitor point sources and cities (section 3.3.3) and to quantify regional budgets of diffuse emissions is partly due to the underlying assumption made in this study that the atmospheric transport is perfectly known. We ignored the transport modeling errors and implicitly supposed that the position and extent of the ~~city or point sources~~ XCO₂ plumes ~~from cities or point sources~~ are perfectly simulated. Accounting for the transport modeling errors by assuming that they can be described as a random noise uncorrelated in space and time, as is usually done in atmospheric inversions, would not fundamentally change the results of this study. The impact of such an error would be equivalent to decreasing the precision of the observations. However, as a

consequence of errors in the wind speed and direction, or in the ~~structure of the CO₂~~ transport model ~~itself~~, the simulation of narrow and localized plumes can poorly match the actual ones and strongly affect the inversion of the emissions from point sources and cities. The actual signature of diffuse emissions could also be different in terms of shape, and extent from the actual one. An inversion procedure that would simultaneously control both transport parameters and CO₂ emissions within a coupled meteorological-CO₂ transport model (Kang et al., 2011) may partially address these issues although it would greatly increase the complexity of the inversion, in particular by introducing potentially large non-linearities in the observation operator. New methods based on imagery processing (Corpetti et al., 2009), plume detection (Kuhlmann et al., 2019), Gaussian plume modeling (Nassar et al., 2017) or direct computation of fluxes through detected plumes may also help overcoming realistic transport uncertainties to invert the emissions corresponding to point sources and cities (Varon et al., 2018). However, these methods may hardly deal with diffuse emissions whose signatures have low amplitudes, and should be difficult to detect in XCO₂ images. The recovering of countryside emissions and of regional budgets of the emissions may thus be optimistic in this study.

Moreover, the configuration of our inversion system uses many of the traditional assumptions of ~~atmospheric inversions systems~~ among which the rather simple characterization of ~~some the~~ sources of uncertainties ~~by with~~ Gaussian distributions, which may underestimate their impact ~~in the inversions~~. Furthermore, some sources of observation errors, such as the ~~so-called~~ systematic errors which bear spatial correlations, are purposely ignored in the present study. Such errors exacerbate the problem of the identification of the signatures of point source, city to diffuse emissions in addition to have a larger error budget than random noise ~~on at~~ the spatial scales of such signatures. There is a clear need to assess the impact of such correlated errors ~~with a similar system than the one presented in this study~~. Lastly, the extent of the observation sampling is made rather optimistic by ignoring cloud cover ~~and the loss of data due to large aerosol loads~~. In a general way, the results of this study could be seen as optimistic and as an upper limit of the skill of the inversions using satellite images only but, also as good indicators of the sensitivity of the uncertainty reduction to various parameters and drivers.

Broquet et al. (2018) provided insights into the impact of cloud cover and systematic errors in the XCO₂ images. They used realistic simulations of satellite samplings and errors made for the CarbonSat mission by (Buchwitz et al. 2013a). Results have indicated that, when accounting for cloud cover, the satellite data could efficiently constrain emissions from Paris only ~20 days per year, and that the impact of the systematic errors anticipated for that mission is such that the system would hardly be able to reduce errors in the emission estimates if ~~such these~~ errors are not filtered or controlled for. ~~However,~~ ~~e~~fforts have been made to limit the amplitude of ~~such systematic~~ errors in the concept of the new CO2M mission.

Our new inversion framework allows accounting for a realistic simulation of the observation sampling and errors. ~~Nevertheless, generating simulations of the systematic errors from the retrieval of XCO₂ data that are suitable for the purpose of our study would have been difficult. Systematic errors are not described in the uncertainties computed by existing retrieval schemes. Furthermore, they depend on specific measurement configurations and on the evolving skill of radiative transfer inverse models and of empirical bias-correction systems, so that their characterization based on diagnostics with existing missions may hardly apply to future ones. Simulating realistic patterns of cloud cover consistent with the~~

845 meteorology for the different test cases investigated would have also been challenging. Finally, this study focuses on other
parameters to allow exploring the sensitivity of the inversions to these parameters in depth.
In order to raise insights into the impact of errors with spatial patterns such as model and systematic errors, we have
conducted experiments where spatial correlations are included in the observation error (in the **R** matrix). We have tested
850 isotropic and homogeneous spatial correlations exponentially decaying with distance, using various correlation lengths. The
experiments and results are described in Appendix A since they are out of the scope of this study. The results indicate that
including correlations in the observation errors tends to increase the budget of observation errors and thus to increase the
posterior uncertainties in the flux estimates as long as the correlation length scale does not exceed that of the signature of the
fluxes in the XCO₂ images. However, including correlations in the observation errors also tends to increase the ability to
855 distinguish between the patterns of the observation error and of the signatures of the fluxes, and thus to decrease the
posterior uncertainties, so that for large spatial correlation lengths, increasing the correlation length leads to a decrease of the
posterior uncertainties. In our tests, the worst situation for the monitoring of the emissions in the study area corresponds to
~10 km correlation length scales. These results should be interpreted cautiously since the spatial patterns of the model and
systematic errors are more complex than this traditional but simple modeling of spatial correlations and since in these tests,
860 the inversion system is perfectly informed about the statistics of the observation error. In particular, future studies will
integrate more realistic simulations of observation sampling and errors from different concepts of spaceborne imagery, based
on radiative transfer inverse modeling applied to realistic fields of surface and atmospheric conditions and instrumental
specifications.

865 A last significant simplification of the general problem of the inversion of the anthropogenic emissions based on XCO₂ data
has been stressed by Ciais et al. (2020). Anthropogenic emissions of CO₂ bear a major share of emissions from biofuel
combustion which can hardly be separated spatially from the fossil fuel combustion component. Furthermore, the emissions
of CO₂ by human respiration represent a significant portion of the total CO₂ emitted from cities. The XCO₂ data and the
atmospheric inversion approaches can hardly be used to distinguish between these different components if it cannot rely on
complementary data. This factor was ignored here, as well as in most of the studies dedicated to the inversion of
870 anthropogenic CO₂ emissions at city to regional scales.

Exploiting further capabilities of the inversion framework: potential of complementary observation systems and results at larger temporal scales

875 Our analysis is restricted to a window from 5:00 to 11:00 corresponding to the period which might be constrained by the
satellite observation from a heliosynchronous satellite with 11:00 local overpass time. This provides little information about
the capacity of the inversion system to monitor daily to monthly budgets. To address this need and improve the results of the
inversions, the modularity of our data assimilation framework could integrate multiple streams of data in order to increase
the temporal and spatial coverage of the information (Moore et al., 2018; O'Brien et al., 2016).

880 In future studies, our inversion system could also integrate more realistic hypotheses concerning the description of the prior
uncertainties in the emissions. Uncertainties in inventories are indeed difficult to characterize and this study assumed, rather
arbitrarily, that the relative ~~error for~~ uncertainty in the 6h budgets was of 50% or 100% (section 2.4). We also assumed that
the prior uncertainties had low temporal correlations, i.e. correlation over timescales of less than 6h, and we neglected the
spatial correlations between the budgets for the different point sources, cities and countryside areas. Actual anthropogenic
emissions and inventories have complex cycles at daily and weekly scales together with a large temporal and spatial
885 variability. There is still a critical lack of knowledge and of characterization of the correlations in the uncertainties in the
inventories (Wang et al. 2018, Wang et al., 2020). However, some extensive analyses are now conducted to fill this gap
(Super et al., 2020). ~~w~~We can expect some stronger spatial and temporal connections than assumed in this study, which
would increase the transfer of information from the atmospheric observations, *i.e.* both the uncertainty reduction for the
sources and time windows covered by the observations, and the uncertainty reduction for other sources, time windows, and
890 for large spatial and temporal scales.

Appendix A: testing the impact of spatial correlations in the observation error

When addressing the impact of errors from the atmospheric modeling and observation, our study focuses on the random
error from the satellite instrumental noise which is assumed not to bear spatial correlations (following, e.g., Buchwitz et al.,
2013a). It does not study the impact of two major sources of uncertainties in the inversions: the transport modeling errors
895 and the systematic errors in the XCO₂ retrievals from the inverse radiative transfer modeling (Hobbs et al., 2017). These
errors raise spatial correlations in the overall observation errors (Worden et al., 2017; Broquet et al., 2018).
Beside the need to limit the scope of the study to a reasonable extent, one reason for not studying the impact of these errors
here is that they depend on the specific modeling and measurement configurations and on the evolving skill of transport and
radiative transfer models and of empirical bias-correction systems (see, e.g., the dramatic improvement of the agreement
900 between NASA's Orbiting Carbon Observatory-2 retrievals and reference ground-based retrievals between version 7 and
version 8 of NASA's algorithm in Fig. 18 of O'Dell et al., 2018). Therefore, conclusions from existing missions may hardly
apply to future ones. Another reason is that there are characterized empirically and that we still lack of robust and theoretical
ways to describe them. However, in order to feed the discussion and perspectives regarding such errors, we provide here a
first exploration of the impact of spatial correlations in the observation errors.
A traditional way to model spatial correlations between observation errors is to assume that they decrease with the distance,
and that they are isotropic and homogeneous in terms of spatial scale (Chevallier, 2007). Therefore, we model such
correlations using exponentially decaying functions ($\exp(-d/D)$) of the distances d between two observation pixels, the
parameter D defining the correlation length scale. Their inclusion in \mathbf{R} (sections 2.3.1) is tested by applying such correlations
905 to the total observation errors, i.e., not defining the total observation errors has a combination of spatially correlated and
910 uncorrelated errors, but as a single component with spatial correlations. This choice eases the analysis of the impact of

spatial correlations here, which aims at raising first general insights on the behavior of the inversion when including them rather than at providing a quantification of such an impact.

In a new set of tests with our inverse modeling framework, we use the set-up which covers the sources over most of the 2-km resolution model subdomain based on 900 km wide XCO₂ images (section 3.3). We set **R** with a 1-sigma uncertainty of 0.3 ppm for all XCO₂ data, and with D=10, 50 or 100 km. In order to lighten the computations associated with the inversion of a matrix **R** that is no longer diagonal (Eq. 2), we perform the inversions by considering observations with 4 km resolution pixels rather than 2 km resolution ones. The choice of a data precision of 0.3 ppm allows thus to get results for a configuration that is close to the reference one (with data resolution and precision of 2 x 2 km² and 0.6 ppm respectively; see section 3.2.2 and 4 regarding the similarity of the results from the inversions done with a given instrumental precision and resolution and the inversions done with an instrumental precision twice smaller and a resolution four times larger).

When introducing correlations with small spatial scales (D=10km), the uncertainty reductions for all type of emission sources are lower than for the inversions performed without correlations within the **R** matrix (Fig. A3). When using D=50km, this is still true for most of the cities and for countryside and regional emission areas. However, the uncertainty reductions get larger than when ignoring spatial correlations for all point sources (except 2) and the cities that emit less than ~25MtC/year. Finally, when using D=100km, the uncertainty reductions get larger than when ignoring spatial correlations for all point sources, all cities emitting less than 40MtC/year, and a significant number of countryside and regional emission areas. (i) This behavior, (ii) the analysis of images of the plume from Paris with the different types of observation errors tested here (Fig. A4), (iii) the fact the cities with the largest emission rate are generally also those with the largest spatial extent, and (iv) the fact the sources with larger emission rates have an atmospheric signature which can be distinguished on larger spatial extent, lead us to the following interpretation of the impact of the spatial correlations in **R**:

- when introducing correlations with small spatial scales in **R**, the posterior uncertainties in flux estimates increase since these correlations yield larger budget of observation errors at the scale of the signatures of the targeted fluxes. This impact tends to saturate when the spatial correlation scale in **R** reaches and gets larger than the scale of the signatures of the targeted fluxes.
- conversely, the increase of the correlation scales helps the inversion separate the observation error patterns from these signatures.
- These two opposed effects lead to a worst case in terms of posterior uncertainties in the emission estimates that correspond to correlation scales in **R** that are function of the scale of the flux atmospheric signatures. In particular, it varies depending on whether we analyze results for point sources, cities or widespread emissions across regions.

These insights into the impact of spatial correlations in **R** call for further investigations. In particular we should test more complex patterns that could better correspond to actual model and systematic retrieval errors (e.g. following surface and atmospheric structures). We should also challenge the potential of the inversion to separate such structured observation errors from flux atmospheric signatures in conditions where this inversion is not perfectly informed about the error spatial correlations.

945 **Acknowledgements**

This work was supported by the Chaire Industrielle Trace ANR-17-CHIN-0004-01 cofunded by the ANR French national research agency, THALES ALENIA SPACE, SUEZ-Environnement, and TOTAL-Raffinage Chimie. The authors are grateful for TAS partners, in particular Sandrine Mathieu, LMD partners, in particular Vincent Cassé, Cyril Crevoisier and Olivier Chomette, and colleagues from LSCE: Yilong Wang, Elise Potier and all the Chaire TRACE team for fruitful discussions. [This work was performed using HPC resources from GENCI-TGCC \(grant no. 2019-A0070102201\).](#)

950 **References**

[AIRPARIF: Bilan des émissions de polluants atmosphériques et de gaz à effet de serre en Île-de-France pour l'année 2010 et historique 2000/2005. Méthodologies et résultats, Tech. rep., AIRPARIF Surveillance de la Qualité de l'Air en Île-de-France, Paris, France, available at: <http://www.airparif.asso.fr/pdf/publications/inventaire-emissions-idf-2010-rapport-130731.pdf> \(last access: 26 March 2015\), 2013.](#)

955 Bocquet, M., Wu, L. and Chevallier, F.: Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism, *Quarterly Journal of the Royal Meteorological Society*, 137(658), 1340-1356, 2011.

[Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO₂ emissions from space and related applications, *Atmos. Meas. Tech.*, 3, 781–811, <https://doi.org/10.5194/amt-3-781-2010>, 2010.](#)

960 Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO₂ emissions from atmospheric concentration measurements, *Atmos. Chem. Phys.*, 15, 1707–1724, <https://doi.org/10.5194/acp-15-1707-2015>, 2015.

965 Broquet, G., Bréon, F.M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., Chevallier, F., Wu, L., and Ciais, P.: The potential of satellite spectro-imagery for monitoring CO₂ emissions from large cities, *Atmos. Meas. Tech.*, 11, 681–708, <https://doi.org/10.5194/amt-11-681-2018>, 2018.

Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V., Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO₂ and CH₄ retrieval errors by error parameterization, *Atmos. Meas. Tech.*, 6,3477–3500, <https://doi.org/10.5194/amt-6-3477-2013>, 2013a.

970 Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O. and Gerbig, C.: Carbon Monitoring Satellite (CarbonSat): assessment of scattering related atmospheric CO₂ and CH₄ retrieval errors and first results on implications for inferring city CO₂ emissions, *Atmos. Meas. Tech.*, 6(12), 3477-3500, 2013b.

[Chevallier, F.: Impact of correlated observation errors on inverted CO₂ surface fluxes from OCO measurements. *Geophysical Research Letters*, 34 \(24\), 2007.](#)

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font color: Custom Color(RGB(34;34;34)), English (U.S.)

Chevallier, F., Feng, L., Bösch, H., Palmer, P. I., Rayner, P. J.: On the impact of transport model errors for the estimation of CO₂ surface fluxes from GOSAT observations, *Geophysical Research Letters*, 37(21), 20104.

~~Chevallier, F.: Description of the CO₂ inversion production chain. CAMS deliverable CAMS73_2015SC2_D73.1.5.5_201703_CO₂ inversion production chain_v1. <http://atmosphere.copernicus.eu/>, 2017a.~~

980 ~~Chevallier, F.: Validation report for the inverted CO₂ fluxes, v16r1. CAMS deliverable CAMS73_2015SC2_D73.1.4.2_1979_2016_v1_201707. <http://atmosphere.copernicus.eu/>, 2017b.~~

Ciais, P., Crisp, D., Denier van der Gon, H. A. C., Engelen, R., Heimann, M., Janssens-Maenhout, G., Rayner, P., and Scholze, M.: Towards a European Operational Observing System to Monitor Fossil CO₂ emissions, European Commission Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs Directorate I – Space Policy, Copernicus and

985 Defence, Brussels, Belgium, 2015.

~~Ciais, P., Wang, Y., Andrew, R., Bréon, F. M., Chevallier, F., Broquet, G., ... & Zheng, B.: Biofuel burning and human respiration bias on satellite estimates of fossil fuel CO₂ emissions. *Environmental Research Letters*, 15(7), 074036, 2020.~~

Corpetti, T., Héas, P., Mémin, E., Papadakis, N. : Pressure image assimilation for atmospheric motion estimation. *Tellus A: Dynamic Meteorology and Oceanography*, 61(1), 160-178, 2009.

990 ~~Crisp, D.: Committee on Earth Observation Satellites GHG White Paper, 2018.~~

~~Hobbs, J., Braverman, A., Cressie, N., Granat, R., & Gunson, M.: Simulation-Based Uncertainty Quantification for Estimating Atmospheric CO₂ from Satellite Data. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1), 956-985, 2017.~~

Houweling, S., Aben, I., Breon, F. M., Chevallier, F., Deutscher, N., Engelen, R., Marshall, J.: The importance of transport model uncertainties for the estimation of CO₂ sources and sinks using satellite measurements, *Atmospheric chemistry and physics*, 10(20), 9981-9992, 2010.

995 Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A. M., ..., and Huijnen, V.: The CAMS reanalysis of atmospheric composition, *Atmospheric Chemistry and Physics*, 19(6), 3515-3556, 2019.

Kaminski, T., Rayner, P. J., Heimann, M., Enting, I. G.: On aggregation errors in atmospheric transport inversions, *Journal of Geophysical Research*, Atmospheres, 106(D5), 4703-4715, 2001.

1000 Kang, J. S., Kalnay, E., Liu, J., Fung, I., Miyoshi, T., and Ide, K.: "Variable localization" in an ensemble Kalman filter: Application to the carbon cycle data assimilation. *Journal of Geophysical Research: Atmospheres*, 116(D9), 2011.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO₂ emission plumes of cities and power plants with the Copernicus Anthropogenic CO₂ Monitoring (CO2M) mission, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2019-180>, in review, 2019.

1005 Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., ... & Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO₂ exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochemical Cycles*, 22(2), 2008.

~~Menuet, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., ... & Mailler, S.: CHIMERE 2013: a model for regional atmospheric composition modelling. *Geoscientific model development*, 6(4), 981-1028, 2013.~~

Formatted: Font: 10 pt, Subscript

Formatted: Font color: Black

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: French (France)

Formatted: Font: (Default) Times New Roman, 10 pt, English (U.S.)

Formatted: Font color: Custom Color(RGB(34;34;34))

- 1010 [Monteil, G., Broquet, G., Scholze, M., Lang, M., Karstens, U., Gerbig, C., ... & White, E.: The regional EUROpean atmospheric transport inversion COMparison, EUROCOM: first results on European wide terrestrial carbon fluxes for the period 2006–2015. *Atmospheric Chemistry and Physics Discussions*, 1-40, 2019.](#)
- Moore III, B., Crowell, S. M., Rayner, P. J., Kumer, J., O'Dell, C. W., O'Brien, D. and Lemen, J.: The potential of the geostationary carbon cycle observatory (GeoCarb) to provide multi-scale constraints on the carbon cycle in the Americas, *Frontiers in Environmental Science*, 6, 109, 2018.
- 1015 Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO₂ emissions from individual power plants from space, *Geophys. Res. Lett.*, 44, 10045–10053, <https://doi.org/10.1002/2017GL074702>, 2017.
- O'Brien, D. M., Polonsky, I. N., Utembe, S. R., and Rayner, P. J.: Potential of a geostationary geoCARB mission to estimate surface emissions of CO₂, CH₄ and CO in a polluted urban environment: case study shanghai, *Atmospheric Measurement Techniques*, 9(9), 4633-4654, 2016.
- 1020 [O'Dell, C., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M., Fisher, B., ... & Merrelli, A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, 2018.](#)
- [Oda, T. and Maksyutov, S.: ODIAC Fossil Fuel CO₂ Emissions Dataset, Center for Global Environmental Research National Institute for Environmental Studies, <https://doi.org/10.17595/20170411.001>, 2015.](#)
- 1025 Parkinson, C. L., Ward, A., and King, M. D.: Earth science reference handbook: a guide to NASA's earth science program and earth observing satellite missions, National Aeronautics and Space Administration, 277, 2006.
- Patra, P., Law, R. M., Peters, W., Rödenbeck, C., Takigawa, M., Aulagnier, C., and Bruhwiler, L.: TransCom model simulations of hourly atmospheric CO₂: Analysis of synoptic-scale variations for the period 2002–2003, *Global Biogeochemical Cycles*, 22(4), 2008.
- 1030 Pillai, D., Buchwitz, M., Gerbig, C., Koch, T., Reuter, M., Bovensmann, H., Marshall, J., and Burrows, J. P.: Tracking city CO₂ emissions from space using a high-resolution inverse modeling approach: a case study for Berlin, Germany, *Atmos. Chem. Phys.*, 16, 9591–9610, <https://doi.org/10.5194/acp-16-9591-2016>, 2016.
- Pinty, B., Janssens-Maenhout, G., Dowell, M., Zunker, H., Brunhes, T., Ciais, P., ... , and Engelen, R.: An operational anthropogenic CO₂ emissions monitoring and verification support capacity. Baseline requirements, model components and functional architecture, doi: 10.2760/08644, European Commission Joint Research Centre, EUR 28736 EN, 2017.
- 1035 [Pison, I., Menut, L., & Bergametti, G.: Inverse modeling of surface NOx anthropogenic emission fluxes in the Paris area during the Air Pollution Over Paris Region \(ESQUIF\) campaign. *Journal of Geophysical Research: Atmospheres*, 112\(D24\), 2007.](#)
- Pregger, T., Scholz, Y., et Friedrich, R.: Documentation of the anthropogenic GHG emission data for Europe provided in the Frame of CarboEurope GHG and CarboEurope IP. *Institut für Energiewirtschaft und Rationelle Energieanwendung, Universität Stuttgart, Stuttgart, Germany*, 2007.
- 1040

Formatted: English (U.S.)

Formatted: Font: 10 pt, Subscript

Formatted: Font: (Default) Times New Roman, 10 pt, English (U.S.)

Formatted: Font: (Default) Times New Roman, 10 pt, English (U.S.)

Formatted: Font color: Custom Color(RGB(34;34;34)), English (U.S.)

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

- Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO₂ emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), *J. Geophys. Res.*, 121, 2015JD024473, <https://doi.org/10.1002/2015JD024473>, 2016.
- 1045 Schwandner, F. M., Gunson, M. R., Miller, C. E., Carn, S. A., Eldering, A., Krings, T., ..., and O'Dell, C. W.: Spaceborne detection of localized carbon dioxide sources, *Science*, 358(6360), eaam5782, 2017.
- Stauffer, J., Broquet, G., Bréon, F.-M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO₂ emissions based on atmospheric inversion, *Atmos. Chem. Phys.*, 16, 14703–14726, [https://doi.org/10.5194/acp-16-](https://doi.org/10.5194/acp-16-14703-2016)
- 1050 14703-2016, 2016.
- Stockman, G., Shapiro, L.G.: Computer Vision (1st ed.). Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001.
- Super, I., Dellaert, S. N., Visschedijk, A. J., & Denier van der Gon, H. A.: Uncertainty analysis of a European high-resolution emission inventory of CO₂ and CO to support inverse modelling and network design. *Atmospheric Chemistry and Physics*, 20(3), 1795-1816, 2020.
- 1055 Tarantola, A.: Inverse Problem Theory: Methods for Data Fitting and Parameter Estimation, Elsevier, 1987.
- Thiruchittampalam, B.: PhD thesis Entwicklung und Anwendung von Methoden und Modellen zur Berechnung von räumlich und zeitlich hoch aufgelösten Emissionen in Europa, PhD Thesis at the University Stuttgart IER, Band 118, 2012.
- Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O., Xia, Y., ..., Huang, Y.: Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmos Meas Tech*, 11, 5673-5686, 2018.
- 1060 [Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., ... & Tao, S.: Potential of European 14 CO₂ observation network to estimate the fossil fuel CO₂ emissions via atmospheric inversions. *Atmospheric Chemistry and Physics*, 18\(6\), 4229-4250, 2018.](#)
- [Wang, Y., Ciais, P., Broquet, G., Bréon, F.-M., Oda, T., Lespinas, F., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., Xu, H., Tao, S., Gurney, K. R., Roest, G., Santaren, D. and Su, Y.: A global map of emission clumps for future monitoring of fossil fuel CO₂ emissions from space. *Earth Syst. Sci. Data*, 11\(2\), 687–703, doi:<https://doi.org/10.5194/essd-11-687-2019>, <https://doi.org/10.5194/gmd-2019-326> Preprint. Discussion started: 2 January 2020 c Author\(s\) 2020. CC BY 4.0 License. 27 2019.](#)
- 1065 [Wang, Y., Broquet, G., Bréon, F. M., Lespinas, F., Buchwitz, M., Reuter, M., ... & Ciais, P.: PMIF v1. 0: an inversion system to estimate the potential of satellite observations to monitor fossil fuel CO₂ emissions over the globe. *Geoscientific Model Development Discussions*, 1-27, 2020.](#)
- [Worden, J. R., Doran, G., Kulawik, S., Eldering, A., Crisp, D., Frankenberg, C., ... & Bowman, K.: Evaluation and attribution of OCO-2 XCO₂ uncertainties. *Atmospheric Measurement Techniques*, 10\(7\), 2759-2771, 2017.](#)
- 1070 [Wang, Y., Broquet, G., Bréon, F.-M., Lespinas, F., Buchwitz, M., Reuter, M., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., and Ciais, P.: PMIF v1.0: an inversion system to estimate the potential of satellite observations to monitor](#)

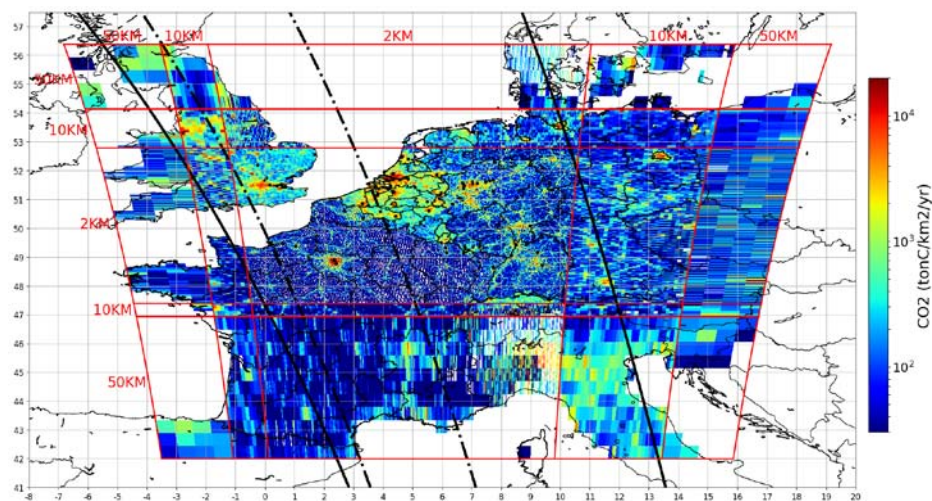
Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt, Subscript

Formatted: Font color: Custom Color (RGB(34;34;34))

Formatted: Font: (Default) Times New Roman, Font color: Auto

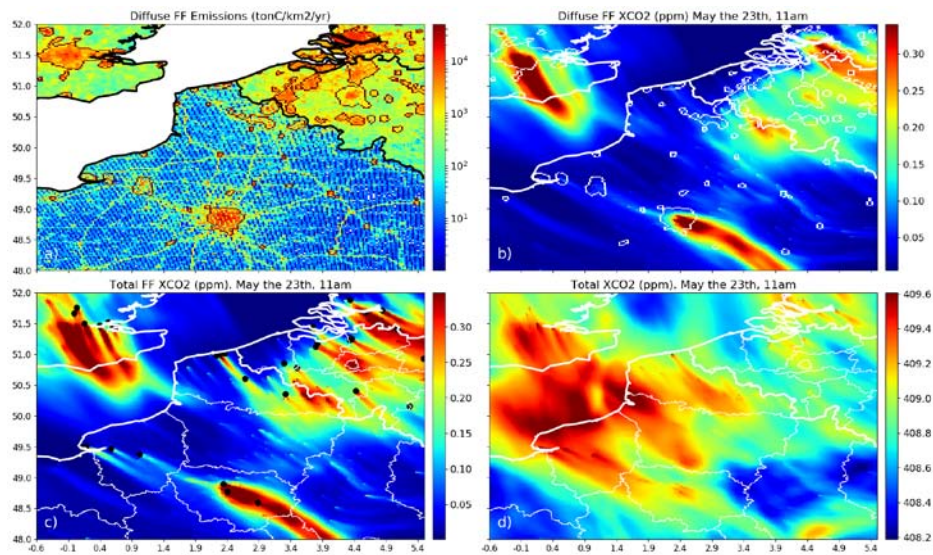
1075 fossil fuel CO₂ emissions over the globe, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2019-326>, in review, 2020.



1080 **Figure 1:** Maps of the IER annual emissions interpolated over the domain of the CHIMERE model. Point sources are indicated by black dots and administrative regions by thin black lines. The grid of the model is defined by sub-domains with several resolutions ($r \times s$ km² where r and $s=2, 10$ or 50 km) and whose boundaries are represented by the red lines. The thick black lines depict the edges of the satellite tracks corresponding to the synthetic data used in this study (300 km swath: dash-dotted line, 900 km swath: solid line).

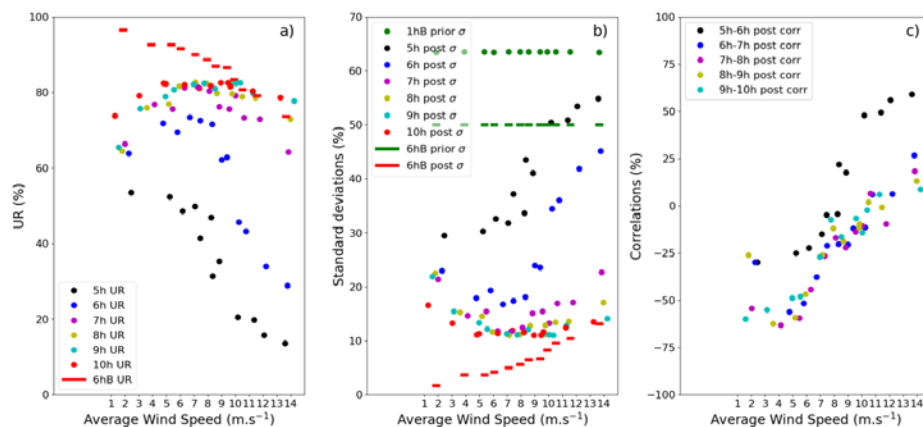
1085

1090



1095 | Figure 2: IER emission maps interpolated over northern France, western Belgium and the London area (a). We have represented the anthropogenic emission sources without point sources ("area emissions"). Red curves depict the boundaries of the city clusters defined by a pattern recognition algorithm (Section 2.2.3.). Panels (b) and (c) show the simulations of XCO₂ (ppm) on May the 23rd at 11am that are produced respectively by the area and total anthropogenic emissions between 5:00 and 11:00. Point sources are indicated by black dots in panel (c). In panel (d), simulations of XCO₂ (ppm) on May the 23rd at 11am that are produced altogether by the anthropogenic, natural and boundary fluxes between 5:00 and 11:00. For the sake of clarity, these figures do not show the whole 2 km-resolution sub-domain of CHIMERE, but illustrate the patterns seen over this subdomain well.

1100



1110 | Figure 3: Uncertainty Reductions (UR) from a 50% prior uncertainty on the 6 h-budgets of the Paris emissions for 12 days
 1115 | characterized by different average wind speeds over Paris (a). URs for hourly and 6 h-budgets of the Paris emissions are shown by
 color dots and red segments respectively. In panel (b) are shown prior vs. posterior uncertainties on 1h-emissions (color dots) and
 6h-emissions of Paris (green and red segments). In panels (a,b), the colors of the dots represent the hour of the corresponding 1 h-
 budget; the green dots are for the prior uncertainties on the 1 h-emissions (1hB prior σ) which are derived from 50% prior errors
 on the 6h-budgets (6hB prior σ) and by considering temporal prior correlations of 3 hours. Panel (c) shows correlations between
 posterior uncertainties in 2 consecutive 1 h-emissions (color dots). Results are computed with a retrieval resolution of 2 km \times 2 km,
 a precision of 0.6 ppm and a swath of 300 km.

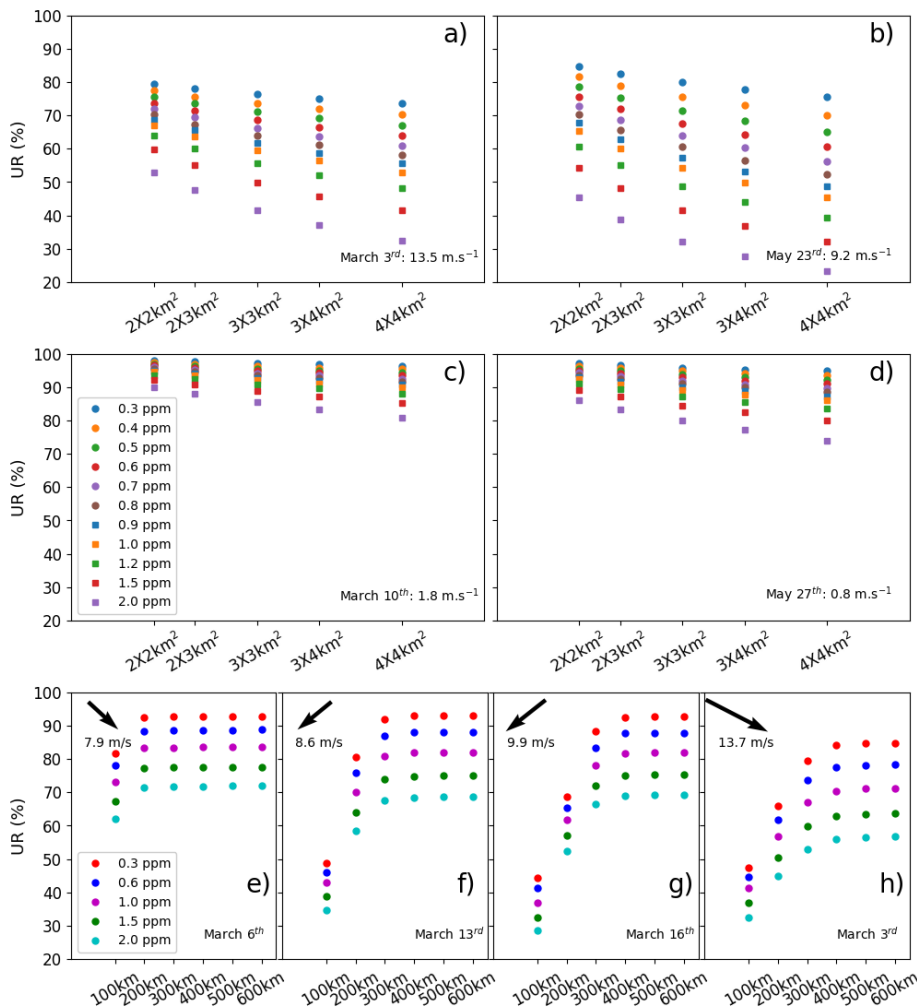
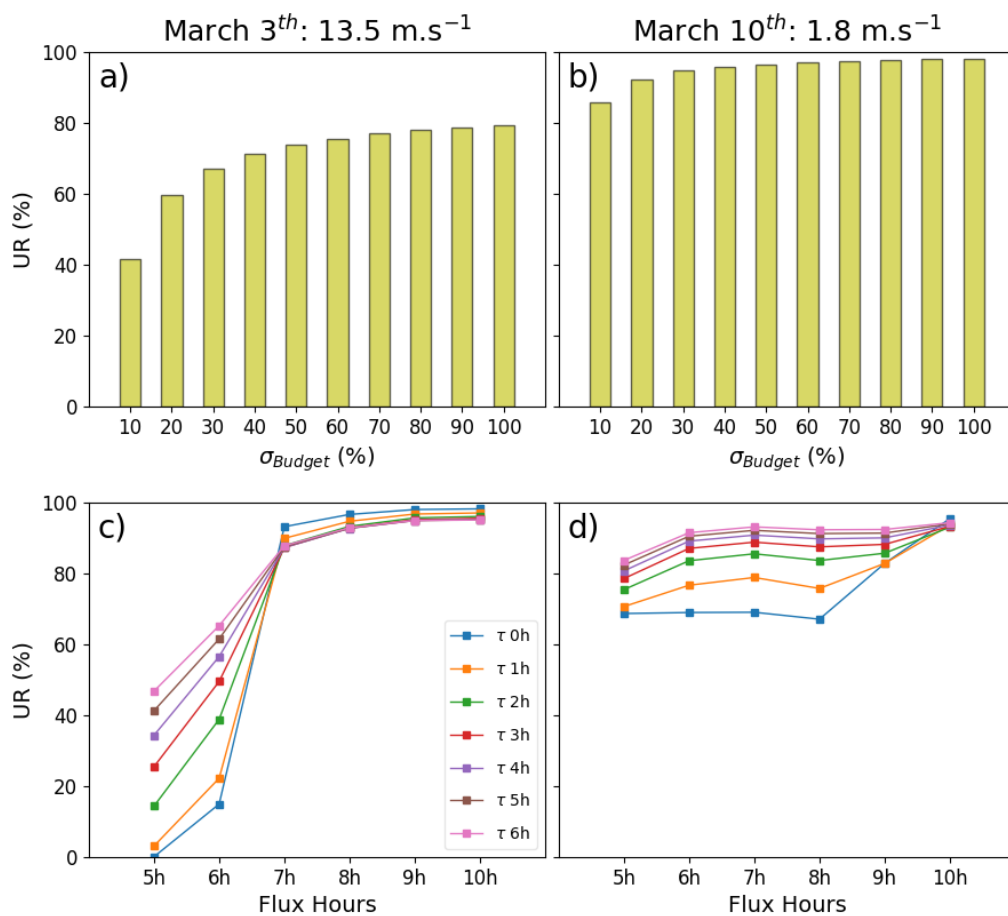
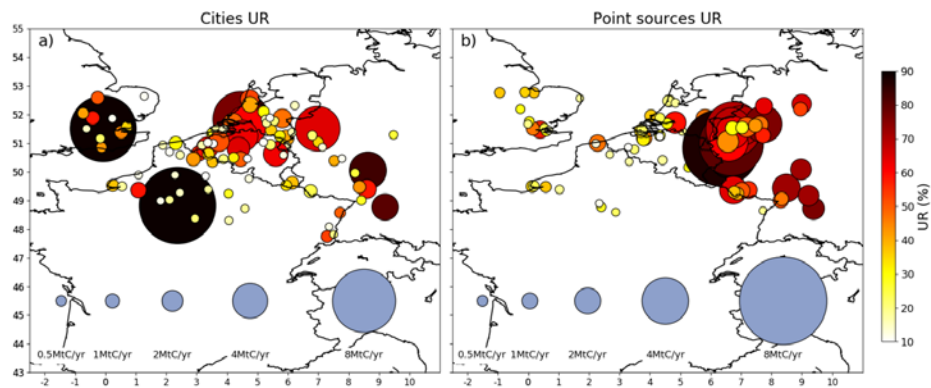


Figure 4: Uncertainty Reductions (UR) from a 50% prior uncertainty for the 6 h-budgets of the Paris emissions. In panels (a)-(d), results are displayed for 4 different days characterized by different wind speeds, for different spatial resolutions of the satellite data (x-axis) and for different precisions (color markers). For the panels (a)-(d), results are generated by considering a swath of 300 km. In panels (e)-(h), results are displayed for 4 different days characterized by different wind speeds, for different swaths of the satellite data (x-axis), for different precisions (color markers) and for a resolution of 2 km by 2 km.



1125 Figure 5: Uncertainty Reductions (UR) as a function of the prior uncertainty (x-axis) for 6 h-budgets of the Paris emissions (a,b).
 Correlations between the prior errors on hourly emissions have a temporal length of 3 hours (see section 2.4). Panels (c,d) show the
 URs for the hourly emissions between 5h and 11h (x-axis) for several temporal lengths defining the correlations between prior
 errors on hourly emissions (color dots), legend “ τ 0h” being for an absence of such correlations. Prior uncertainties on 6 h-budgets
 of Paris emissions are taken equal to 50% in panels (c,d). Columns represent 2 different inversion days: March 2016 the 3rd
 (strong wind) and March 2016 the 10th (low wind). All inversion results are computed with a retrieval resolution of 2 km \times 2 km, a
 1130 precision of 0.6 ppm and a swath of 300 km.



1135 | Figure 6: Mean Uncertainty Reductions (UR) for some city clusters (a) and some point sources (b) across the 62 inversion results of the days of March and May 2016. The areas and colors of the disks represent the annual emissions ($\text{MtC}\cdot\text{yr}^{-1}$) and the URs (%) respectively. The inversions are performed with a retrieval resolution of $2\text{ km} \times 2\text{ km}$, a precision of 0.6 ppm and a swath of 900 km. Prior uncertainties on 6h-budgets of clusters and point sources emissions are taken equal to 50% and prior error correlations have a temporal length of 3 hours.

1140

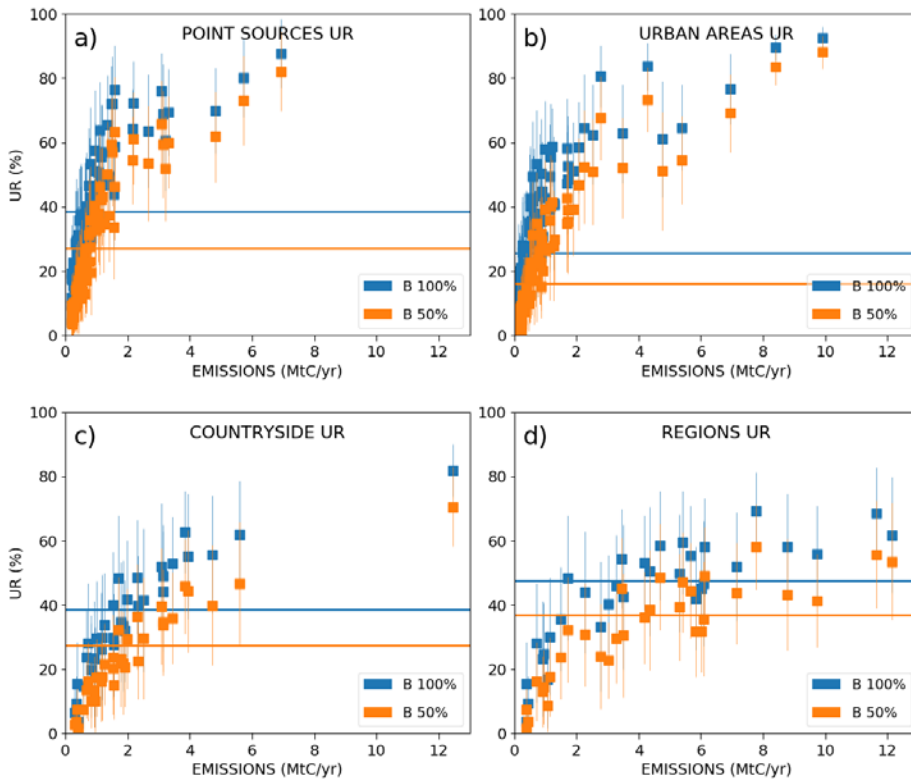


Figure 7: Mean values and standard deviations of the Uncertainty Reductions (UR) for the emissions of point sources (a), urban areas (b), countryside areas (c), and regions (d) across the 62 inversion results of the days of March and May 2016. The lines represent the averages of the temporal mean values across all sources of a given type. The emitting areas are chosen within the 2-km-resolution domain of the model so that they are covered by the satellite track in order to avoid swath effects. Results are given as a function of the annual emissions (x-axis) of the emitting areas. The inversions are performed with a retrieval resolution of 2 km × 2 km, a precision of 0.6 ppm and a swath of 900 km. Prior uncertainties on the 6 h-budgets of the point sources, urban and countryside areas are taken equal to 50% (orange squares) and to 100% (blue squares). Prior uncertainties on the regional budgets are derived by aggregation of the prior uncertainties on their constituent emitting sources for both cases. Prior error correlations between hourly emissions have a temporal length of 3 hours.

1145

1150

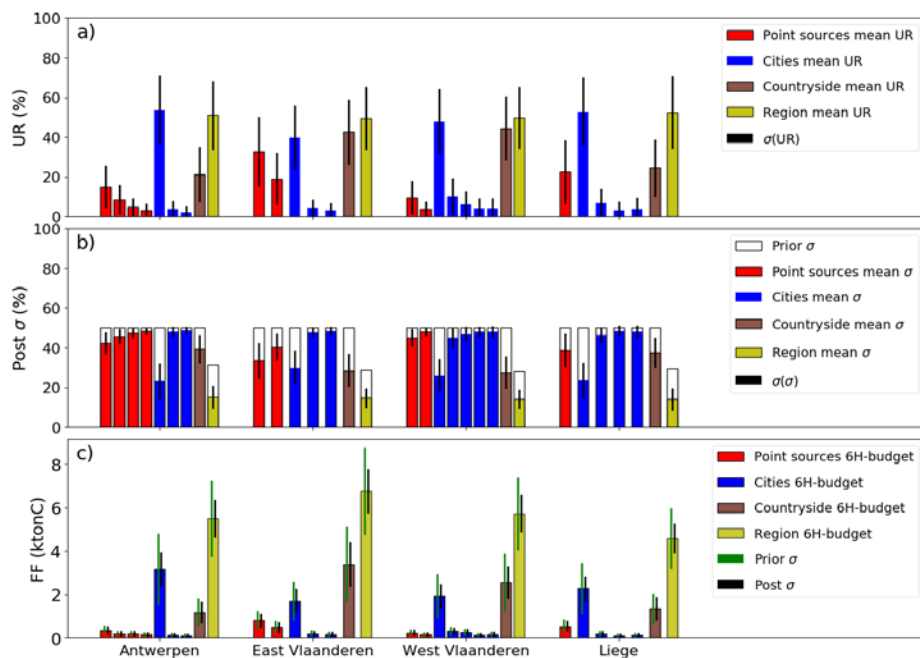


Figure 8: (a) Mean values and standard deviations of the Uncertainty Reductions (UR) for the regions, point sources, urban and countryside areas constituting 4 Belgium regions. Averaging is performed across the 62 inversion results of the days of March and May 2016. (b) Mean values and standard deviations of the relative post uncertainty for each emitting area. Prior uncertainties are represented as well. Prior uncertainties in the regional budgets are derived by aggregation of the prior uncertainties of their constituent emitting sources (mean value ~33%). Panel (c) shows the mean 6-hours budgets for each emitting area. The inversions are performed with a retrieval resolution of $2 \text{ km} \times 2 \text{ km}$, a precision of 0.6 ppm and a swath of 900 km. Prior uncertainties on the 6h-budgets of the point sources, urban and countryside areas are taken equal to 50% and prior error correlations between hourly emissions have a temporal length of 3 hours.

1155

1160

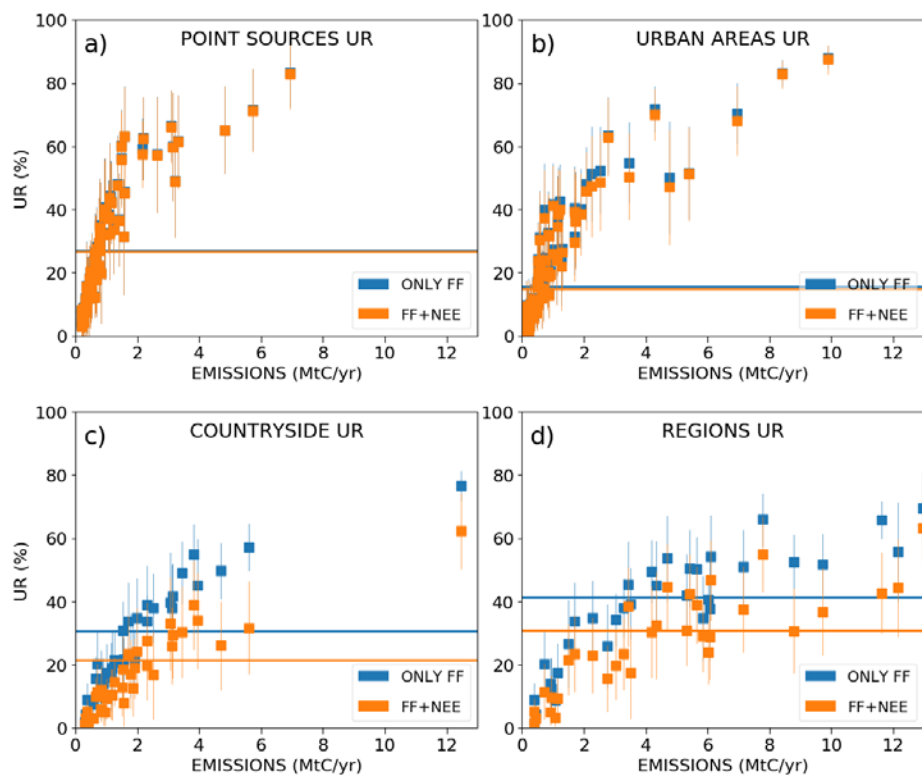
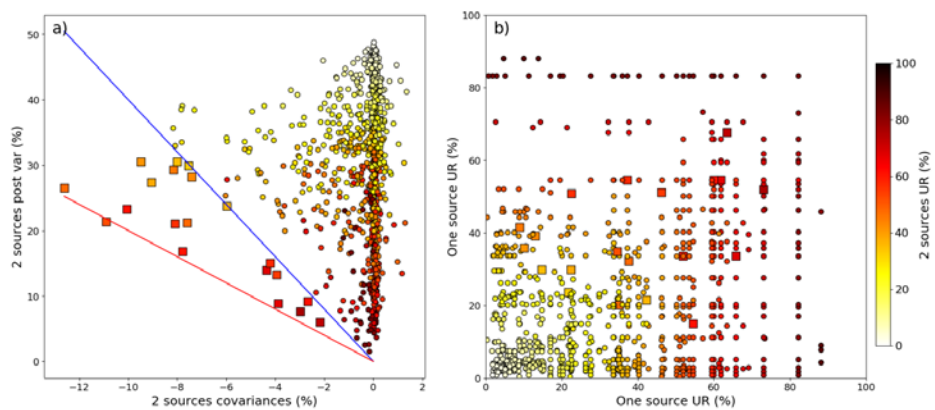


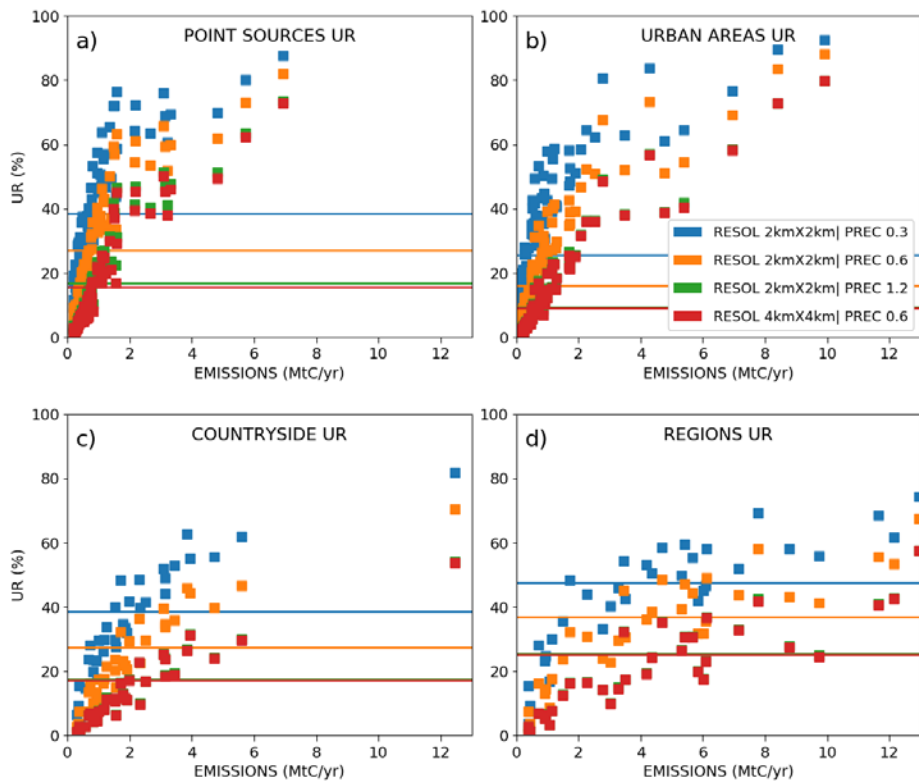
Figure 9: Same as Fig. 7 but results are derived from inversions considering the anthropogenic emissions only (blue markers) and from inversions considering the natural fluxes as well (orange markers). Prior uncertainties on the 6 h-budgets of the point sources, urban and countryside areas are taken equal to 50% and prior error correlations between hourly emissions have a temporal length of 3 hours. Prior uncertainties on the regional budgets are then derived by aggregation of the prior uncertainties of their constituent emitting sources (mean value ~33%).

1165



1170 **Figure 10:** a) Post variances of the emission estimates for pairs of anthropogenic sources contained within a same region function of the covariances between the individual sources of the pairs. The colors of the markers correspond to the UR of the pairs. b) For a pair of anthropogenic sources contained within a same region, UR of one individual source function of the UR of the other individual source of the pair. The colors of the markers correspond to the UR of the pairs. The regions corresponding to this plot are contained within the 2 km-resolution area of the model. Results are derived with an instrumental resolution of $2 \text{ km} \times 2 \text{ km}$, a precision of 0.6 ppm and a swath of 900 km. Prior uncertainties on the 6 h-budgets of the point sources, urban and countryside areas are taken equal to 50% and prior error correlations between hourly emissions have a temporal length of 3 hours. Pair of anthropogenic sources are made up of point sources and of urban and countryside areas.

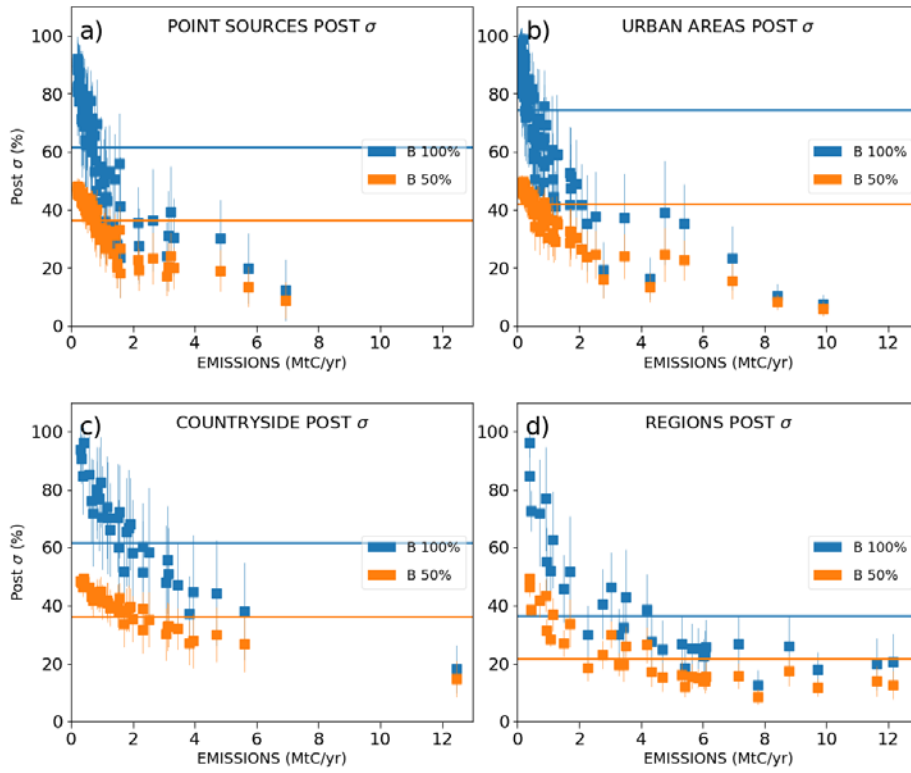
1175



1180 **Figure 11:** Same as Fig. 7 but the inversions are performed with different retrieval resolutions and precisions (swath=900 km).
 1185 **Prior uncertainties in the 6 h-budgets of the point sources, urban and countryside areas are taken equal to 50%.**

1185

1190



1195 | Figure A1: Same as Fig. 7 but relative posterior uncertainties are shown instead URs.

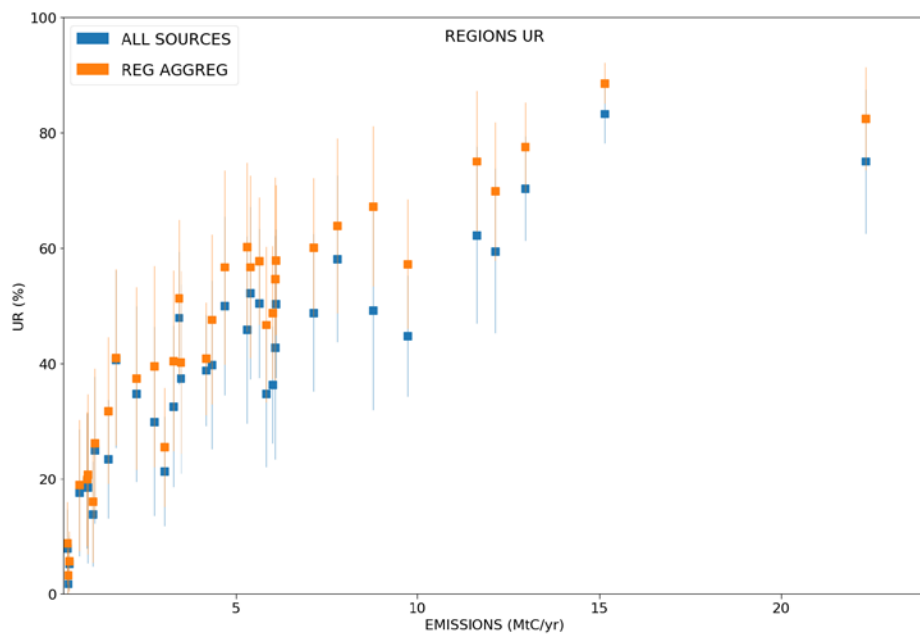


Figure A2: Mean values of the Uncertainty Reductions (UR) of the emissions at the regional scale across the 62 inversion results of the days of March and May 2016. Results are given function of the annual emissions (x-axis). Regional emissions are inverted with (blue markers) and without (orange markers) considering an internal separation of the region into cities, point sources or countryside areas.

1200

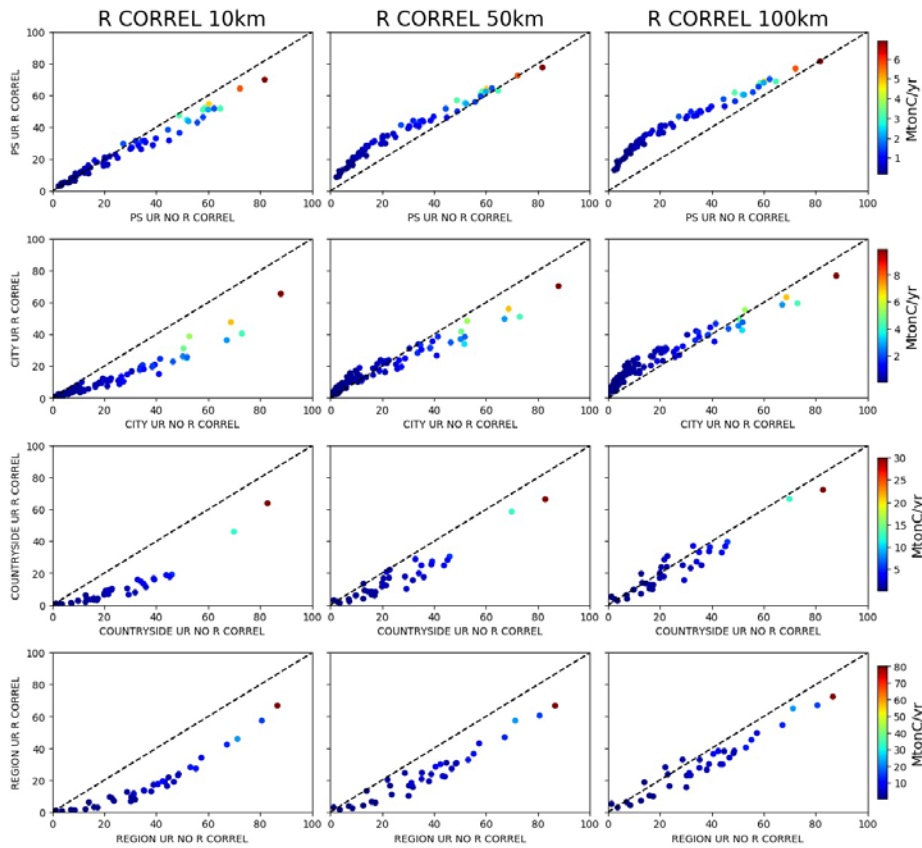
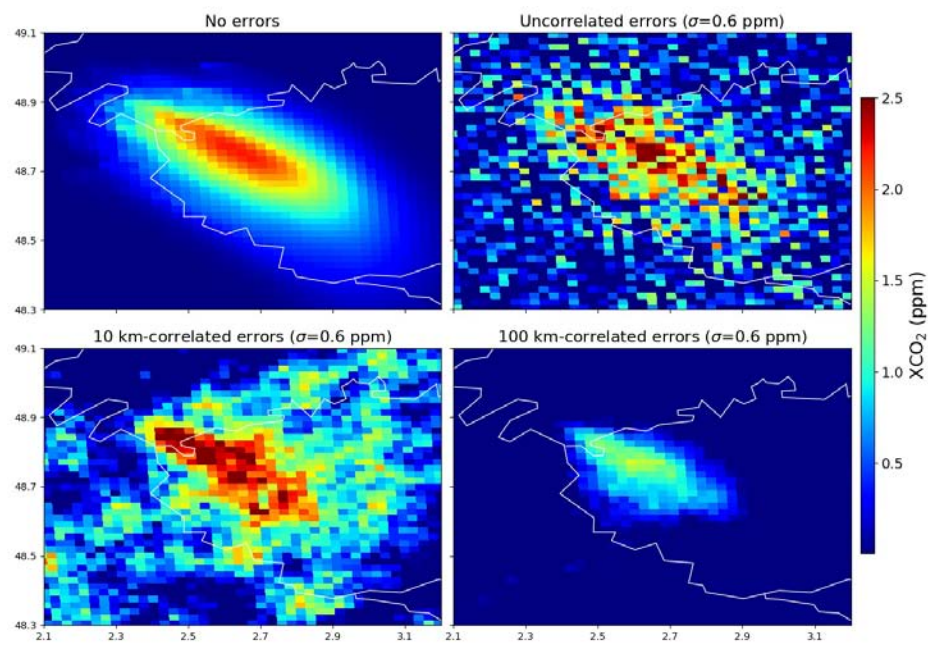


Figure A3: Uncertainty reductions (UR) for the emissions of the point sources (1st row), the urban areas (2nd row), the countryside areas (3rd row) and the regions (4th row) when considering spatial correlations (Y-axis) or no correlations (X-axis) between the observation errors. The 1st, 2nd and 3rd column correspond to correlations with a spatial scale of 10, 50 and 100 km respectively. The color of the dots corresponds to the annual budgets of the sources (color bar on the right of the figure). The dashed line is the 1:1 line.

1205



1210

Figure A4 : XCO₂ distributions produced by the emissions of the Paris area between 5:00 and 11:00 on March the 8th at 11:00: raw CHIMERE simulations (top left figure), raw CHIMERE simulations perturbed by an uncorrelated random noise of 0.6 ppm (top right figure), raw CHIMERE simulations perturbed by a multivariate normal distribution whose covariance matrix is characterized by an uncertainty of 0.6 ppm for the diagonal terms and a correlation length scale of 10 km (bottom left figure) and of 100 km (bottom right figure).

1215