Dear Referee #2

Thank you for thoroughly reading and commenting the manuscript. Please find below the replies to your suggestions; each of your suggestions is followed by the corresponding reply in bold letters and (where appropriate) actions taken to address it in the updated version of the manuscript in italics.

Major comments:

1. While I agree that inclusion of prior error correlations between different emissions can improve observation constraint, and help disentangle sources, improper characterization of the error correlation may result in systematic bias in the posterior estimate. So I suggest a more complicated OSSE is necessary, where perturbations are generated using different correlation parameters to exam how well the system will reproduce the 'true' fluxes, with incorrect correlation coefficients.

This is a very useful suggestion, which we followed now. We propose to add the following at Page 14, Line 29

"Improper characterization of the error correlation may result in systematic bias in the posterior estimate. As mentioned in Sect. 2.1.6, inter-species correlation, the correlation between different fuel types and the correlation between different emission sectors in ***Sprior*** *is assumed equal to 0.7 (Sect. 2.1.4). To assess how well the system will reproduce the 'true' fluxes with incorrectly specified correlations, a series of experiment was performed in which the inter-species correlation in* ***Sepsilon*** *remains equal to 0.7, while the three correlation coefficients in* ***Sprior*** *assume different values ranging from 0.1 to 0.9. Table 5 shows the residuals between total annual posterior fluxes and total annual true fluxes for the five simulated species, derived similarly as for Table 4.*

We found that for all species the uncertainty reduction increases with correlation. In addition, the posterior-truth biases are always lower than the prior-truth biases.

The posterior uncertainty values are almost always larger then the corresponding bias values, except for CO with prior correlation equal 0.8, and fossil fuels CO_2 with prior correlations equal to 0.6, 0.7 and 0.9. Follows that, except for these few cases, the posterior is not significantly different from the truth. Conversely, prior-truth biases (not shown) are statistically significant in the majority of cases for fossil fuel fluxes, and in some cases also for biogenic fluxes. The effect of assuming the incorrect error correlations appears to be in general small, possibly implying a relative robustness of our methods.

Correlation	Post-Truth CO_2 ff	Post-Truth CO	Post-Truth CH_4	Post-Truth GEE	Post-Truth Respiration
0.1	-17.2 ±17.6	-0.1 ± 0.3	-0.2 ± 0.4	-21.0 ±27.5	12.9 ±26.1
0.2	-13.6 ±14.4	0.1 ± 0.3	-0.2 ± 0.4	4.8 ±27.5	-5.8 ±26.1
0.3	-12.8 ±12.0	-0.2 ± 0.2	-0.2 ± 0.3	-1.2 ±27.4	0.5 ±26.1
0.4	2.9 ±10.1	-0.1 ± 0.2	-0.1 ± 0.3	-21.1 ±27.4	23.1 ±26.0
0.5	0.5 ± 8.6	-0.1 ± 0.2	-0.2 ± 0.3	17.8 ±27.3	-8.3 ±26.0
0.6	25.2 ± 7.3	0.1 ± 0.2	-0.3 ± 0.2	5.7 ±27.3	6.6 ±25.9
0.7	25.6 ± 6.1	-0.1 ± 0.2	0.1 ± 0.2	16.8 ±27.3	-7.2 ±25.8
0.8	-0.9 ± 5.0	-0.2 ± 0.1	0.1 ± 0.2	-5.8 ±27.3	23.3 ±25.7
0.9	13.8 ± 3.7	0.1 ± 0.1	0.2 ± 0.1	-10.2 ±27.2	14.5 ±25.5

Table 5: Residuals between total annual posterior fluxes and total annual true fluxes for the five simulated species (in MtC/yr) and different inter-species correlation values in the prior error covariance matrix (first column). The corresponding posterior uncertainty was added for each Post-Truth value.

-- Note to the Referee: the values for the correlation of 0.7 do not exactly reproduces the values in Table 4, as we realized that in the uploaded version of the paper, the zi-correction described in section 2.1.1 was mistakenly turned off. This has been fixed in the revised version, and the updated Table 4 has values matching the 7th row of Table 5.

For all of the experiments, the residuals between true and posterior fluxes are lower than residuals between true and prior fluxes for each of the simulated species; the difference between the cases with maximum and minimum residuals is around 4.2%. In addition, we found that the posterior aggregated fluxes in the nine experiments are not significantly different from each other, implying that the system is fairly robust against errors in the assumed inter-species correlation.

2. Discussions are more focused on the domain total. It is interesting to see how well the system will reproduce their spatial distribution.

1) Note that we do actually not focus on the domain total, as we believe it is not reasonable to constrain the whole European domain when pseudo-observations are focused only around a single city; for this reason we chose the region marked by the 50% footprint area, that contains most of the surface influence. We suggest to add the following sentence at page 8 – line 30:

As the pseudo-observations are clustered around a single location (Frankfurt), fluxes over the whole European domain can very likely not be constrained. Therefore, as spatial aggregation scale we chose a domain…

2) Regarding the reproduction of spatial distribution: Our modeling framework does not optimize the emissions in the individual grid-cells, but only the scaling factors for emissions from different sectors and fuel types. With this modeling framework it is not possible for us to evaluate how well the spatial distribution is reproduced.

Minor comments:

1. Line 5, Page 4: "This synergy follows from the fact . . ." Better changed to 'follows the fact. . .' or other phrase.

The text was edited according to the suggestion

2. Line 5, Page 14: ". . .have a magnitude of 6-11 Megatons of carbon per year (MtC y-1) in July" The unit of MtC/yr seems inconsistent with annual total presented in Table 5. I think it should be MtC/a.

In Table 4 (not 5), the total presented refers to overall residuals between (e.g.) total prior fluxes minus total true fluxes, aggregated over all emission categories. As such, they are not directly comparable with the amounts shown in Figure 7 (to which Line 5, Page 14 refers), which instead indicates the true fluxes for specific emission categories. Note that we chose to use the unit "MtC y^{-1}" over "MtC/a" as suggested by the journal 'Manuscript preparation guidelines for authors'.

3. Line 26, page 14: ".. .for the whole year between the prior and both posterior and the perturbed prior" The sentence is unclear.

The text was modified as follows:

"To do so, for each of the five simulated species we calculated the total annual fluxes for prior, posterior, truth, and perturbed prior. From these total fluxes we then derive the overall residual between prior and truth, posterior and truth, and perturbed prior and truth."

4. Figure 7: Please explain why for CH4 fluxes in December, their uncertainty has been significantly reduced, but the differences from the 'true' are not obviously improved.

It is normal for Bayesian inversion to have some elements of the posterior state space that are not obviously improved. The expectation is that the posterior values are in agreement with the true values within their respective uncertainty. As we use 1-sigma uncertainties, we expect about 36% to be even outside this uncertainty range. Note that the inversion for monthly fluxes solves for a total of 828 scaling factors for CH_4, of which about 70 contribute to 90% of the fluxes. However, the atmospheric signals associated with these 70 different sectors/fuel types are not observed directly, but only as a combined signal in CH_4.

5. Figure 6: I suggest the authors also provide the prior and posterior error correlation between CO2 fossil fuel emission and biospheric net flux in the main text.

A sentence was added to provide the correlations (Line 26, page 26)

Note that CO_2 from anthropogenic emissions is assumed to be independent from biogenic emissions; therefore prior error correlation between these categories is zero.

6. Figure 7: check the units for monthly fluxes (in main text as well).

The units in Figure 7 and 8 were changed to MtC/m. References to these figures in the main text were also given with this unit.