Dear Referee #1
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.

Comments:

The manuscript "Multi-species inversion and IAGOS airborne data for a better constraint of continental scale fluxes" by Boschetti et al. describes the effect of including the correlations between multiple species in a bayesian inversion framework in order to improve error reduction compared to solving for individual species independently. The experiment described in the manuscript uses synthetic observations based on measurements made during the IAGOS campaign in Europe in order to assess the potential for future measurements of CO2, CO and CH4 during this campaign to better constrain regional emissions of all three species. Finally, there is some discussion of the effect of different assumptions about the prior error of the emissions upon the level of error reduction achieved by the inversion.
Overall the manuscript is fairly well written, with few technical corrections necessary. The figures are generally quite clear and well chosen, although some further detail needs to be provided for some of them. The methods and models used within the manuscript are appropriate for such a study, and are able to provide some assessment of the potential for improvement supplied by future multi-species measurements as part of the measurement campaign. The paper is successful as far as it goes, and whilst it would have been nice to further examine the effect of different experiment set-ups within this paper, the authors acknowledge that this is the case, and may be the focus of a future manuscript.

We appreciate these positive remarks.

My main reservation with the study is that the results and discussion section is a little light on detail in places and feels like it was rushed, making the thread of the paper more difficult to follow than it should be. More details and deeper analysis of the results is needed in order to contextualize the findings of the experiment. The authors must make sure that all terms used have been explained or defined, and that they provide enough analysis of their results. See general comments for details. I suggest that this paper is suitable for publication in this journal after the following revisions are carried out and the results section is improved.

Thank you for the constructive comments. In the revised version …

Page 3 line 6: "Because most biogenic fluxes in Europe are influenced by human activities..." - reference?
We have added references and modified the sentence to:

"Because most biogenic fluxes in Europe are influenced by human activities, with 22% of Europe's land is dedicated to agriculture (FAO, 2013) and 45 % covered by forests, of which 80% are managed for wood supply (UNECE, FAO, 2011), understanding and managing these biogenic fluxes must also be a component of any policy to reduce anthropogenic emissions."

Page 4, lines 1-2: "proven to be important in the fields of..." - reference?
Two references were added; one for IAGOS and one for CONTRAIL (Zbinden et al, 2013; Sawa et al., 2012)

Page 8, line 18: the first term in equation (3) should be to the power of (-1).
The equation (3) was corrected accordingly

Page 8, line 31: the term "50% footprint" should be explained.
A reference to section 2.1.2 was added to remind the reader of the 'footprint'; the section now reads:
"As a spatial aggregation scale we chose an area from which fluxes have a significant contribution to the observations made at Frankfurt. For this we compute the temporally accumulated footprint values for the whole year 2011, and select those spatial pixels that correspond to 50% of the total (spatially integrated) footprint (Fig. 1)."

We also modified Section 2.1.2 (Pag. 6, Line 12) to better explain the concept of footprint:
"…so-called "footprints". Briefly, for each measurement location and time (also called receptor point), the model releases an ensemble of virtual particles that are driven back in time using simulated wind fields from ECMWF and turbulence as stochastic process; the residence time within the lower half of the mixed layer is used to determine the potential contribution from surface fluxes, and the cumulative sum of these contributions determines the footprint, that identifies the part of the domain with a certain influence on a single receptor point. This footprint is then matrix-multiplied with an emission map to derive the corresponding simulated mixing ratio in a given receptor point."

Page 8, line 15: is it fair to assume no correlation between months? You should comment here (or later in the discussion) on whether this would be the best set-up of the correlation matrix in an inversion using real observational data.
In page 10, line 16, the following was added:
In this study, we assume a certain annual total domain wide flux uncertainty, and then break it down by sectors, fuels, and months by inflating the error. By assuming no correlation between different months we ensure maximum flexibility in the system to retrieve month-to-month changes based on the observations. Assuming correlation between months would be possible, but has not been investigated here. It is unclear how good the seasonal variation in emissions from the inventories actually is, so in order to not rely too much on these we chose zero correlation.

Investigating the effects of different correlation set-ups for the seasonal cycle could be the focus of future research.

Page 12, line 4: What is *enh*?
Right after equation (11), the line *"…where enh indicates the modelled enhancement, and both the horizontal …"* **was added for clarity**

Page 12, line 7-8: You need to explain how you derive $\varepsilon_{␣}$ tran_v in more detail here.
The text (from Page 12, line 6) was edited as follows:
… where both the horizontal transport error ε_{tran_h} *and the vertical transport error* ε_{tran_v} *are characterized as percentage error;* ε_{tran_h} *is assumed to be 50% while* ε_{tran_v} *is profile-specific with mean value about 10%.*
"The vertical transport error accounts for the fact that the shallower the mixed layer is, the more difficult it is to model the atmosphere. We assume that after zi-correction the remaining error is on the order of 50 m (related to the vertical resolution of the profile data), so the relative error ε_{tran_v} *is assumed as the ratio of 50 m to the modeled* z_i*; in this way we obtain an error that gets larger the shallower the mixed layer is. "*

Page 12, line 18: What method do you use to invert **Sprior** and **Se** ?
We assume this comment refers to Pag 8, eq. (3) and (4). The error correlation matrices are inverted using the R-function "solve" of the base package. At pag. 8, line 26 the following was added:
In this study, the inverse of the matrix was calculated using the R-function 'solve' from the base package.

Page 13, line 5: Describe which version of the model output you are plotting in Figure 5. Does it use the prior emissions?
We edited the text at lines 3-4 as follows:
Figure 5 shows … for both observations and model outputs using prior emissions.

Page 13, line 9: Here, and in the caption of Figure 5, you say that the modelled CO is multiplied by a factor of 2.8. However, the legend of Figure 5 appears to say that the observations have been scaled. Which is correct?
The text and the caption were correct. The legend has been corrected accordingly

Page 13, line 12: Explain here what it is that is indicated by the performance of the model compared to the observations. Are you saying that the meteorology that you use and the correction to zi that you apply produce a good indication of the temporal variation of the ML enhancement? Does your choice of zi display an improvement over the original?

Thank you for pointing this out. We have now investigated the improvement brought about by using the zi correction. The text in line 10 and following were edited as follows:
"Mixing ratios are highly variable, but the model produces a good indication of the temporal variation of the ML enhancement; the squared correlation coefficient between observed and modeled CO enhancements is 0.62, while the standard deviation of corrected model and observation residuals is 85 ppb; note that by not accounting for the z_i correction, such values would be 0.56 and 87 ppb respectively. The median of the mixing ratio enhancement for the three trace gases is 2.8 ppm for CO_2, 18.6 ppb for CO and 26.6 ppb for CH_4."
-- Note: we found that in the uploaded version of the paper the zi correction was actually switched off. After switching the correction on, only Figure 5 is affected. The mean uncertainty reduction values are now 35% for CO2_ff, 48% for CO and CH4, 60% for GEE and 63% for respiration. We deeply apologize for the mistake --

Page 13, lines 25 and 26: You could probably add a little more detail to this one-sentence paragraph. Explain that Figure 6 is showing the prior and posterior emission error covariance matrices for the base multi-species inversion. Do the single-species matrices show a similar overall error reduction? Do you expect to see negative correlations in the posterior matrix? As it stands this sentence is disjointed and appears to come out of nowhere and doesn't relate to other text, making the manuscript unnecessarily difficult to follow.
We propose to replace the one-sentence paragraph with the following:
"Figure 6 shows the prior and posterior error covariance matrices for the base multi-species inversion. The posterior error covariance matrix for the multi-species inversion (Fig. 6b) shows lower values corresponding to an average uncertainty reduction of 23% across all state vector elements, while the posterior error covariance matrix for the single-species inversion (not shown) is characterized by a mean uncertainty reduction of 20%. This result implies that the multi-species inversion improves the uncertainty reduction by roughly 15%. Negative values in the posterior error correlation matrix are to be expected because different categories are bind together by correlations and therefore are not free to vary independently."

Page 14, lines 24 - 28: Explain what you mean by "a perturbed version of the prior" here. Also, does the multi-species inversion capture the "truth" any better or worse than the single-species inversion?
We propose to add the two following sentences at Line 25:
"Such perturbed version is obtained by adding realization of the prior error to the prior state space, similarly to how the "truth" is obtained. In addition, it was found that the truth-posterior bias of the multi-species inversion is always lower compared to the single-species inversion. Such difference is between 1.7% and 5.7%, according to the simulated species, with an overall value of 2.4%."

Page 15, line 19: How robust do you think the relative uncertainty reductions

that you derive are against different manifestations of the "true" fluxes?
What we investigate in Fig. 10 is not the uncertainty reduction, but the benefit from a multi-species inversion over a single-species one.
The following text was added at line 19:
The benefit of including inter-species correlations shown in Fig. 10 does not depend on different manifestations of the true fluxes, but only on the posterior uncertainty of the multi- and single-species inversions.

Page 15, line 26: Why do you think a smaller prior error for the CO2 FF fluxes compared to the other species leads to a greater uncertainty reduction for the posterior fluxes?
What we investigate in Fig. 10 is not the uncertainty reduction, but the benefit from a multi-species inversion over a single-species one. Uncertainty reduction for CO2 FF is actually greater in Case 1 (36%) compared with the other two cases (29% and 21% respectively), as in those cases (2 and 3) the prior is assumed to be known better already.

We have stated in the paper (page 15, line 26) that the benefit from a multi-species over a single-species inversion increases, when changing the prior uncertainty for CO2 emissions. We think that the reason is the following: changing the prior uncertainty in CO2 emissions means changing also the off-diagonal blocks linking the different species together (see Eq. 8). However, the diagonal block for CO2 in the prior uncertainty changes by a factor four in that case, while the off-diagonal blocks change only by a factor of two. This effectively ties the emissions of CO2 tighter to the emissions of the other species, resulting in more benefit from a multi- over a single-species inversion. Note that this is related to the required rescaling of the prior error covariance matrix described in section 2.1.5.

We suggest adding the following text at line 28:
…for this increase in benefit. The reason for both of these results is probably to be searched in Eq. 8. In fact, changing the prior uncertainty in CO_2 emissions means to also change the off-diagonal blocks linking the different species together. However, by reducing the anthropogenic CO_2 uncertainty from 20% to 10% (Case 2), the diagonal block for CO_2 in the prior uncertainty changes by a factor four, while the off-diagonal blocks change only by a factor of two. This effectively ties the emissions of CO_2 tighter to the emissions of the other species, resulting in more benefit from a multi- over a single-species inversion. Conversely, when all prior uncertainties are reduced by a factor 2 (Case 3), both diagonal and off-diagonal blocks are reduced by a factor four. This explains why Case 1 and Case 3 show similar benefit values.

Page 16, line 3: What makes CO sensitive to different correlation structures during different seasons?
To explain the issue, we added a couple of sentences at line 4:
What makes CO sensitive to different correlation structures during different seasons is that CO enhancement shows a stronger seasonal cycle compared to e.g.

fossil fuel component of the CO_2 enhancement, with average values for January of around 150 ppb (25 ppm for CO_2), and for July of 9 ppb (4 ppm for CO_2). This results in a much weaker constraint on the CO emissions from the CO observations during summer, but still some constraint through the other species such as CO_2 via the a priori correlation in the emissions.

Technical corrections:

Page 1, line 13: no comma needed in "for, GEE"
The text was edited according to the suggestion

Page 1, lines 17 and 18: the percentages reported in the abstract here are in some cases slightly different to those reported in the main text of the manuscript (on page 15).
The percentage values were checked and replaced where needed

Page 2, line 2: difference -> differences
The text was edited according to the suggestion

Page 5, line 10: Matherial -> Material
The text was edited according to the suggestion

Page 10, line 3: Section 2.1.6 -> Section 2.1.5
The text was edited according to the suggestion

Page 16, line 18: Delete "meaning" - or explain what it means.
The word "meaning" was removed as suggested