Title: Estimates of European emissions of methyl chloroform using a Bayesian inversion method
Author(s): M. Maione et al.
MS No.: acp-2014-25
MS Type: Research Article

Responses to Referee 1

Major comments

COMMENT #1: I have some doubts about the results presented in figure 4. If I understand the method correctly, you assign uncertainties to the prior emissions that are a fraction of the emissions themselves. This implies that with zero or very small emissions the assigned uncertainty is small (page 8221).
REPLY: The referee is right. That piece of text was from a previous version of the paper and does not hold anymore. Therefore, the text has been rewritten as follows: "As there is no information about uncertainties, we used the uncertainty in the matrix diagonal elements as defined in Stohl (2009) i.e., for inversion box j, $\sigma_j = \max(p*x_j, 2*p*x_{surf})$ *with p being a properly chosen scaling factor,* x_j *the a priori emission flux in the inversion box j, and* x_{surf} *the global emission value, as estimated by Rigby et al., (2013), homogeneously distributed in the grid cells corresponding to land areas. We tested p values ranging from 50% to 500% of the prior emission estimate, balancing between (i) enough flexibility in emissions to allow adjustments that better fit the observations and (ii) not too high flexibility that might lead to over-fitting of the observations and to noisy and unrealistic emissions."*

COMMENT #2: Also, the resolution at which the emissions are derived varies. In a region with low sensitivity (e.g. Norway) the resolution should be about 36x36 degrees (page 8218).
REPLY: The referee is right. There was a misprint: 36°x36° are the box dimensions for the global inversion, covering the ocean (boxes not inverted). Using the three stations, the European domain is resolved in boxes whose dimension are 1°x1° lat long or 2°x2° lat long in Spain Portugal and Norway. We have corrected the text accordingly.

COMMENT #3: Figure 4 (and the accompanying text) shows that emissions are projected in Norway, exactly at the spots where "emissions to soil and water" are reported. And these were not included in the prior inventory, so I guess that the prior emissions are very low. The same holds for Madrid, Barcelona and hotspots in Northern Africa. Yet, the posterior emissions presented in figure 4 (right panel) show "hotspot" emissions rather far away from the measurement locations that are not in the prior (e.g. Madrid). This implies a tremendous skill of the model to pinpoint emissions in places where the prior has low values (with an uncertainly that is proportional to this low emissions). I cannot believe that inversions are capable to accomplish this, although also Appendix B (figure B1) shows that emissions are derived at the coast of Norway. So, I do not believe that the prior emission map in Figure 4 (left) and the described uncertainty was the basis for the posterior map in Figure 4b. So I would like to see more analysis. What station constrains the Norway emissions? What was the resolution of the emission grid in Norway, Spain, and Portugal?

REPLY: the boxes in remote regions have dimension 1°x1° or 2°x2° and the emission from each box is distributed according to the population. This allows us to allocate emissions also in remote hot-spots, using the a priori emission field and uncertainty as described above. This choice is particularly appropriate for a compound like MCF that is purely anthropogenic. It is important to highlight that the use of such artifice is essential in order to obtain the result reported in Fig.4. In such remote areas the analysis is definitely qualitative, since the emission intensity is affected by an uncertainty equal to the emission itself. The sensitivity maps of the three single stations reported in fig R1.1 can show the sensitivity to emissions in Norway of each station

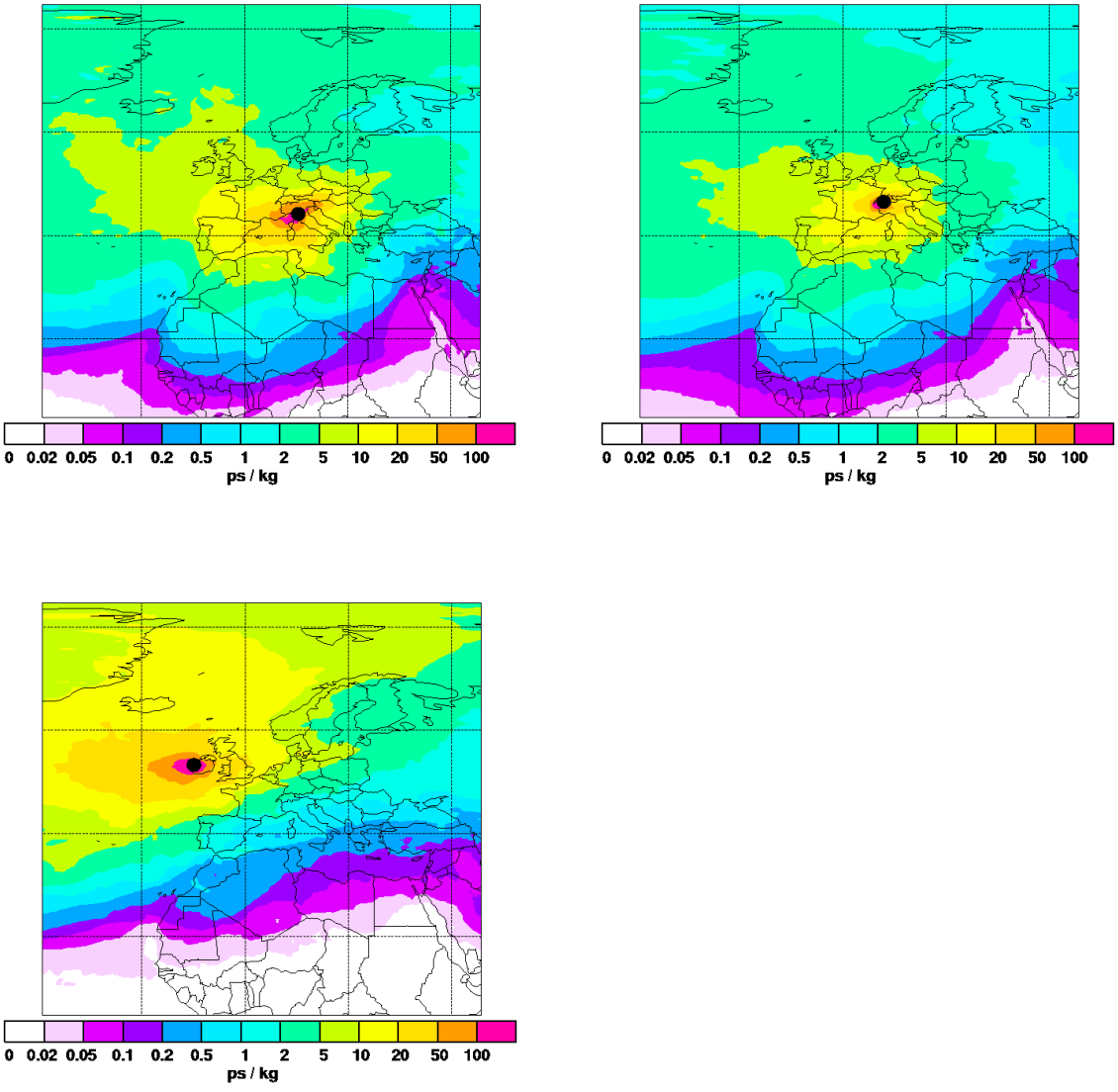


Figure R1.1 Sensitivity to emissions in Norway of each station used in this study

COMMENT #4 Another major issue is the posterior fit with the observations. The information is presented in Appendix A1.4. The relevant section is: "Given that the sources regions are scarce… being CMN with r2,ba = 0.5." In general, these results seem to be influenced by the fact that both the model and the observations correctly predict the general decline in MCF over the years. This naturally leads to a high correlation, which has nothing to do with the inversion. So, I think the analysis should be done on de-trended data. As a result, I would expect a very small r2.

What strikes me also is the small improvement of the r2eb compared to r2ea. Since this is the signal that drives the inversion, it would be instructive to present time series at the various stations (prior and posterior) just to see the skill of the model. Statistics do not look promising though. In Appendix A1.4 it is also written " the low values (of r2eb) obtained at MHD in this analysis confirms how scarcely is MHD affected by polluted air masses, despite the presence of numerous UK sources declared by the E-PRTR inventory". The correlation is not determined by the low number of data-points, but by the skill of the model to reproduce enhancements above the background. This has nothing to do with "scarcely". So, in general, I think Appendix A1.4 should be rewritten (text is not as well written as the main text), and more clearly present how the observations above the baseline drive the emission increments. After all, these are the data that drive the whole story in the main text.

REPLY: We re-wrote the Appendix in order to make it clearer. Furthermore, the referee notes that, in order to evaluate the model performance, the coefficients r^2_{ba} r^2_{bb} r^2_{a} r^2_{b} *should be evaluated on the detrended time series. Our reply is that the correlation between the observed and the not-detrended modelled signal can show not only the model capability but also the emission field characteristics (e.g.* r^2_{ba} *for MHD is 0.91, suggesting that the station is "remote" with respect to the main MCF sources in the SEF area). The model skill can be evaluated through the differences between the a priori and a posteriori coefficients.*

However, the more crucial comment is that concerning the evaluation of the model performance based of the correlation coefficient of the enhancements above the baseline. The three stations used for the inversion show rather low r^2_{eb} r^2_{ea} *values, the two mountain sites because of the well known "incurable" errors affecting mountain stations, and MHD because too far away from the main source region (SEF). In particular, JFJ and CMN have* r^2_{eb} r^2_{ea} *values 0.05 0.07 (JFJ) e 0.12 0.16 (CMN), meanwhile MHD values are 0.06 0.1., and I don't think that, with the available data, we will be able to obtain a significant model improvement. Nevertheless, all the tests performed showed that the results are congruous*

Minor comments

Page 8211, line 4. I think it should be stressed here that the amendments to the Montreal protocol were much more stringent than the original protocol.
Page 8212, line 6. Same issue.
Reply: we changed the text accordingly

Page 8212, line 13: longitudinal? Do you mean latitudinal?
Reply: yes, text changed

Page 8213, line 20: on going. Should be one word (ongoing).
Reply: yes, text changed

Page 8214, line 11: given the fact that the other stations project emissions in Norway, I find it hard to believe that ZEP does not "see " these emissions. Thus, I find it rather strange that 1/4 of the SOGE network is not used in this analysis. At least it would be interesting to present the background concentration of ZEP in this paper.
Reply: As shown by the plot in Fig R1.2, reporting the MCF time series at ZEP, enhancements above the baseline are absent. We performed some tests also including ZEP for the two-year period in which the simulations were available for that station. The results didn't produce any significant

difference in the estimates of the European emission fluxes or any improvement in the uncertainty associated to the emissions from Norway. For this reason, we didn't consider useful for our analysis to prolong the computation time adding the simulations for ZEP.

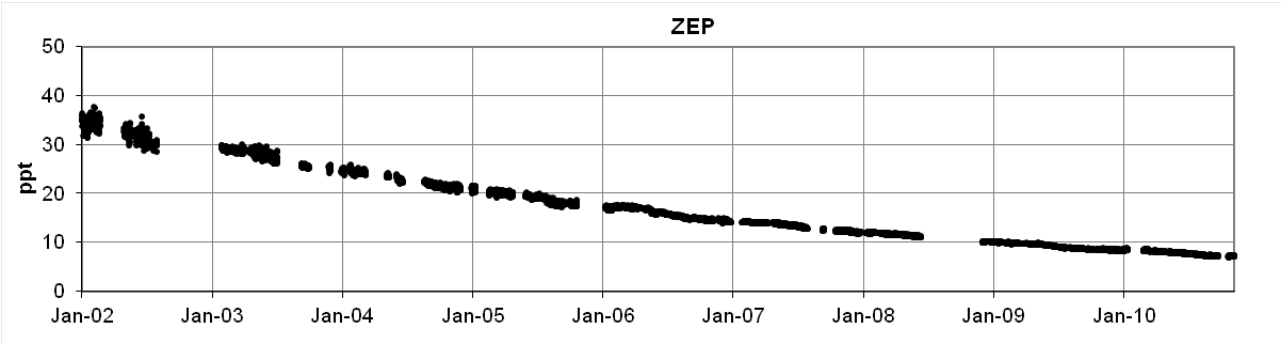


Fig R1.2. MCF time series at ZEP

Page 8214, line 22: Awkward sentence. Consider: ..and m/z values of .... are selected for detection and ...of MCF.
Reply: text changed accordingly

Page 8215, line 19: Space is missing I guess (SA-6RIX)
Reply: yes, text changed

Page 8216, line 22: consider "points" after data
Reply: yes, text changed

Page 8217, line 24: "analytical". This is in conflict with the iterative procedure described under (iii) at line 27.
Reply: yes, text changed

Page 8218, line 10: I think this does not reflect an outlier in the model simulation, but the fact that the model is not able to adequately represent a certain measurement (i.e. a representation error). By the way, in the remainder of the paper, I have not seen a quantification of the number of outliers, and their influence on the inversion.
Reply: the discussion on outliers is part of the inversion technique as described in Stohl et la., 2009 and Stohl et al., 2010. The text has been modified as follows:
"The method used also identifies "outliers" in the model-simulated mixing ratios and assigns them large uncertainties to prevent the solution being strongly influenced by large measurement and/or model errors (Stohl et al. 2009)."

Page 8218, line 25: I suggest to replace "but" by "and".
Reply: text changed

Page 8218, line 26: "is needed". I suggest something like: "The emissions at large distances from the measurement locations cannot be resolved at high spatial resolution"
Reply: text changed

Page 8219, line 3: At the start of this paragraph, it would be useful to write something about the reason why this second method is needed. Later on it becomes more or less clear, but this is the

location to highlight this. Also mention here that you can estimate time varying emissions with the point source analysis (now it is written later).
Reply: we cannot anticipate here, before the description of the inversion results, the reason why we decided to apply this method.

Page 8219, line 23: "While individual emission values are noisy, their average can substantially reduce this noise". Individual data points cannot be noisy. I would describe the method (as I understand it) as follows: The PSA method takes each individual measurement above the baseline, and determines the emission that is needed in a predefined source region (X) to reproduce this measurement exactly. Due to inaccuracies in transport and other numerical errors, this normally results in a noisy emission time series that, upon averaging in time, provides an estimate of the emission in region
Reply: we have modified the text following the reviewer's suggestion

Page 8220, line 6: mixing ratio enhancements above the baseline. I would use either "mixing ratio enhancements" or "mixing ratios above the baseline". Also the caption of figure 2
Reply: text changed

Page 8221, line 1: "emissions to soil and water". Not clear to me what is meant exactly. Please clarify.
Reply: the E-PRTR database divides emissions into emissions to three environmental compartments: air, water and soil. Emissions to soil and water do not include subsequent and fugitive emissions from these two compartments to the atmosphere. The remarkable ability of the model to identify these emissions (not included in our a priori) in our opinion is to be considered a validation, although qualitative, of the reliability of the simulations.

Page 8221, line 10: "leading to", consider, "i.e. 50% to 500% of the prior emission estimate. "
Reply: text changed

Page 8221, line 11: "The emission variability and correlations between measured and a posteriori modelled data substantially increased over this range of values". Unclear sentence. I think the intention is to say that you try to balance between (i) enough flexibility in emissions to allow adjustments that fit the observations better (ii) not too much flexibility because this might lead to over-fitting of the observations and noisy and unrealistic emissions (e.g. negative emissions). I think this message should be better phrased here.
Reply: text changed

Page 8221, line 14: I think "changes" should be "improvements".
Reply: text changed

Page 8221, line 17: "allowing" should be "obtaining".
Reply: text changed

Page 8221, line 25: points (ii) and (iii). See major comment.
Reply: see replies above

Page 8222, line 4: within SEF are included. I suggest: "SEF includes"

Reply: text changed

Page 8222, line 10: We can extend: : :. I suggest: We extend (to reflect your choice).
Reply: text changed

Page 8223, line 10: not impacted: : :. I suggest: not severely impacted
Reply: text changed

Page 8223, line 26: cannot be excluded. I think "is likely" better reflects the situation.
Reply: text changed

Page 8224, line 6: MHD, meanwhile. I suggest full stop: ..MHD. Meanwhile : : :
Reply: text changed

Page 8225, line 6: (except MHD). I suggest to use: : :., but not for MHD.
Reply: text changed

Page 8225, line 7: The error estimates for JFJ and MHD are unrealistically low. How are these number derived? Later on it is suggested that the error bars reflect the scatter in the individually bi-hourly derived estimates. I cannot imagine such small scatter, given e.g. the uncertainty in transport modelling.
Reply: Error bars in the plot have been calculated as a 95% confidence interval for the mean:

$$\sigma = 1.96\frac{S}{\sqrt{N}}$$

Where S is the standard deviation of N sample data. In figure R1.3 is reported the time series of estimated emissions for JFJ for a SRRv threshold>1500. The time series shows how, despite the transport model uncertainties, the reported data are not too much scattered.

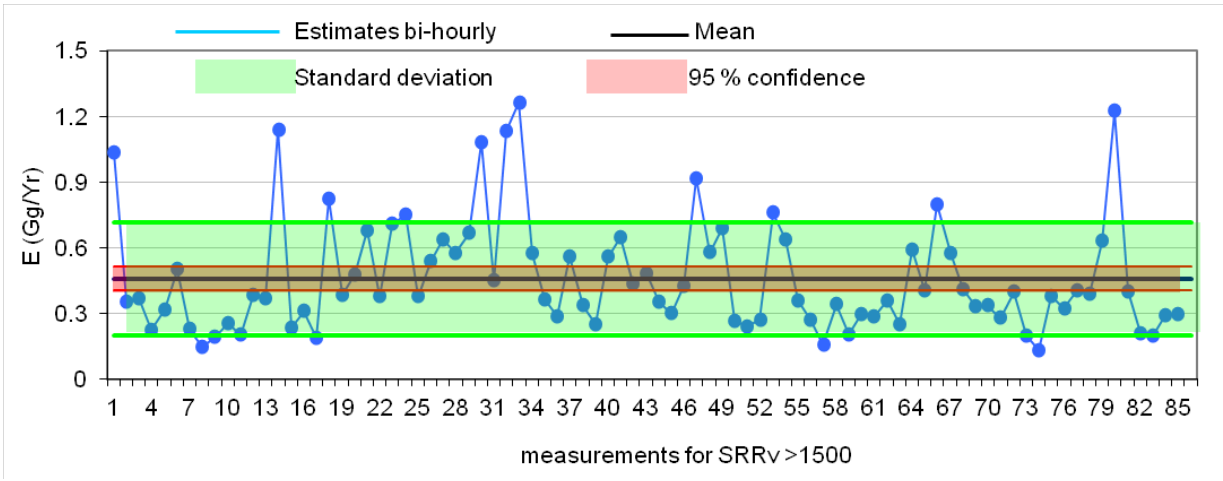
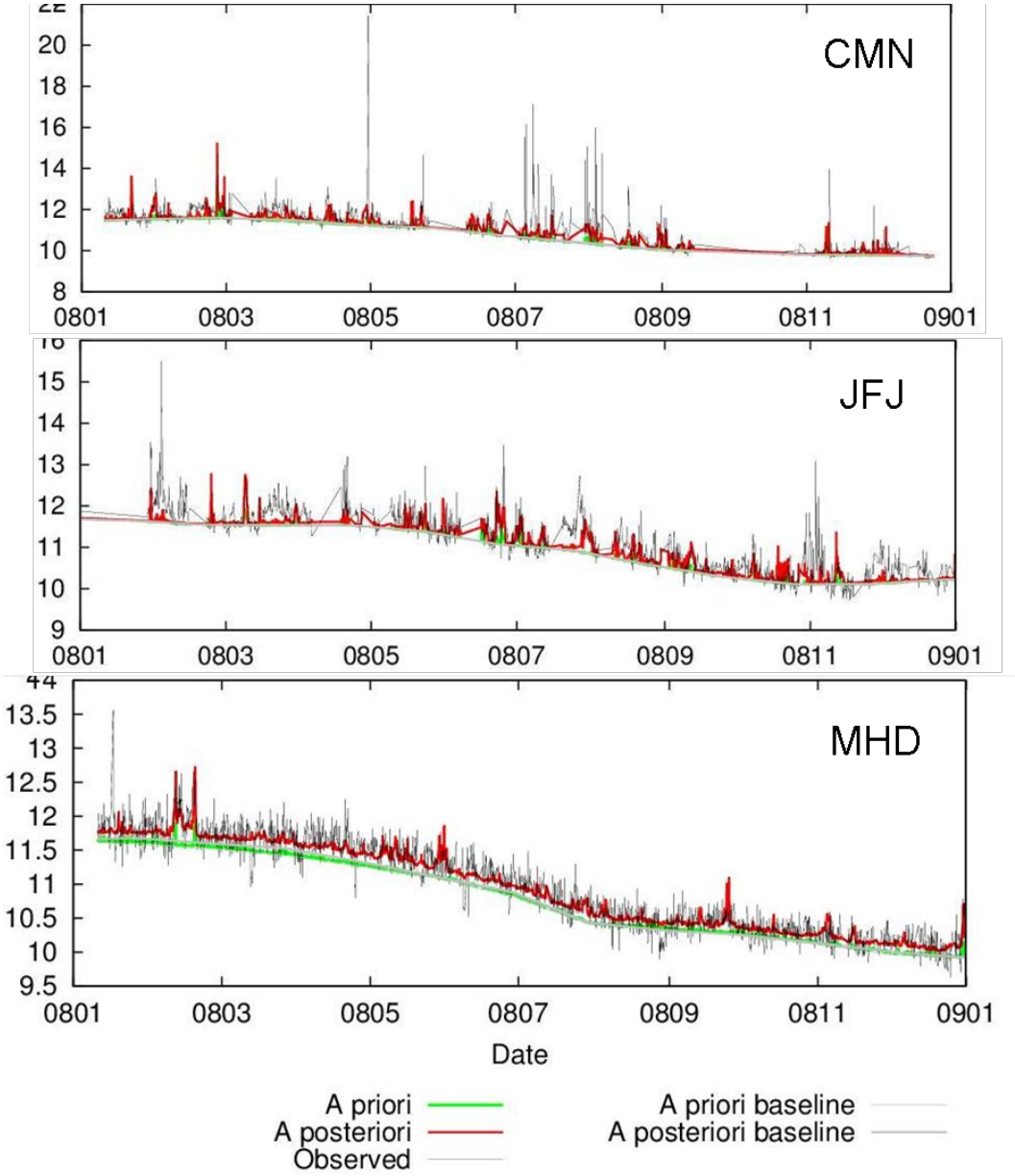


Figure R1.3. Time series of estimated emissions for JFJ for a SRRv threshold>1500

Page 8225, line 19: "This is expected as the inversion results are bound towards a priori emissions that are clearly too low for the SEF area. This leads to a low bias also in the obtained a posteriori emissions". I disagree with this sentence. If you ascribe all the measured concentration enhancement to emission from the SEF region (as done with PSA), you will

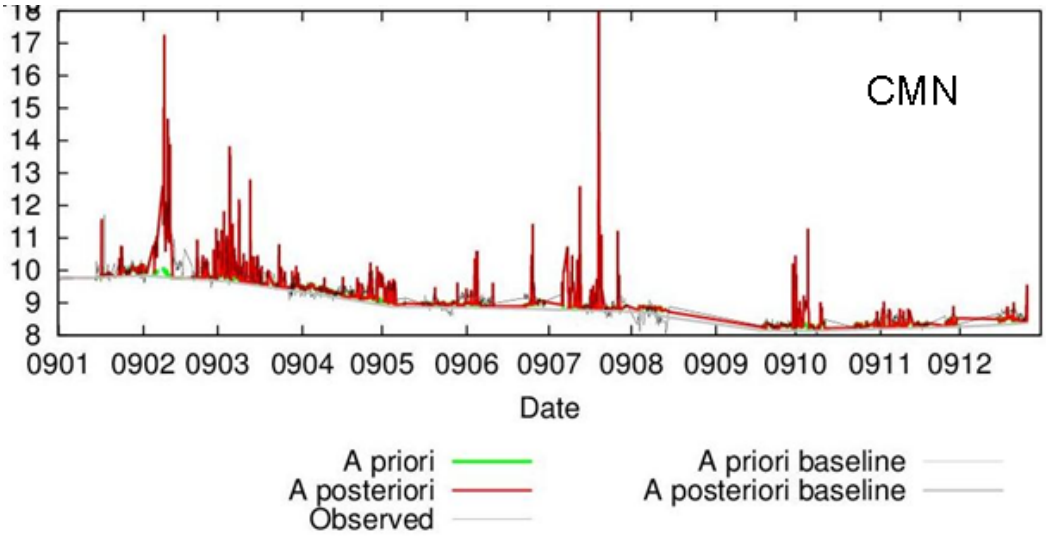
overestimate the SEF emissions (i.e. you project emissions elsewhere also to this region). So I think that both the Bayesian and PSA analysis are biased, but the latter more obviously.

Reply: Probably there is a misunderstanding. Actually, nor the Bayesian inversion, nor the PSA assigns all the spikes to the SEF area. The Bayesian inversion assigns a higher uncertainty to the SEF area than to the rest of the domain. As shown by the comparison between the observed and modelled time series (reported below), the observed concentration maxima intensities sometimes are not well reproduced. We think that this suggests that emissions from the SEF area are sometimes underestimated.



We do not claim that the PSA assigns all the enhancements to the SEF area; as stated at page 8224 line 24: "[...] excess concentrations that cannot be explained by the a priori emission field are assumed to originate exclusively from the SEF area." Nonetheless, emissions could be overestimated due to the incomplete information used to derive the a priori (sources' number and intensity).

Below we report a comparison between the observed and modelled time series using the PSA.



Page 8226, line 19: Conclusions. I suggest, "Discussion and Conclusions", since some discussion is provided, e.g. on non-reported emissions.
Reply: text changed

Page 8227, line 2: "With a less accurate a prior field". I have not read this in the main paper. It is presented in Appendix B, but no reference to this section can be found.
Reply: Reference added

Page 8227, around line 13: I find this hard to believe. See major comments.
Reply: see reply above

Page 8227, around line 15: "It is thus shown". I think this is not convincingly shown in the paper. Maybe there are "signs" of emissions from Norway, but omitting ZEP from the analysis is strange then. Which station picks up the signal, and what is the resolution at which the emissions are calculated? The reported errors are anyhow 100%, so "shown" is way to strong.
Reply: text changed

Page 8227, line 26: "even higher". See remark above.
Reply: text changed

Page 8228, line 13: grids. Should be grid.
Reply: text changed

Page 8228, line 15: Rp. How defined? I understand to varied the grid on which emissions are derived (between 2000 and 6000?, how many simulations?). This gives a mean emission and variations. But what is meant with "maximum error"?
Reply: the Rp parameter quantifies the emission uncertainty for each year using three different tests: i) A priori emission field modulation, ii) Station network geometry e iii) Meteorological data resolution described in appendixes A1.1, A1.2 e A1.3. These tests produced 8 independent estimates, whose variations have been quantified and defined as the maximum error (or maximum semi-dispersion) defined as:

Maximum error = (maximum emission estimate – minimum emission estimate)/2

Page 8228, line 20: I think "comprised" can be removed. Does Rp now relate to the variance between the different scaling factors or to the variance calculated on the different emission grid (between 2000 and 6000)? Also, in table A1 I cannot find the reported values for the SEF region. Under which header?
Reply: we removed comprised. As stated above Rp includes all the uncertainties described in Appendixes A1.1, A1.2 e A1.3. Table A1 is referred to the SEF region only (see Table caption) and shows the uncertainties associated to the three tests evaluated separately for two years. They are: Geometry: Station network geometry; *Intensity: A priori emission field modulation*; *Wind field: Meteorological data resolution. We modified the table accordingly.*

Page 8229, line 1: I thought the analysis was done only for 2008 and 2009?
Reply: In paragraph "A1 Sensitivity tests", we say that the "A priori emission field modulation" has been performed for the entire study period. However, Figure A1 reports 2008 results only.

Page 8229, line 5: In which column?
Reply: see our reply to comment Page 8228, line 20

Page 8229, line 12: For the other regions the averaged Rp is around 40% (Table A1). Table A1 is about SEF.
Reply: we changed the text accordingly

Page 8229, line 14: Figure A2 is a mystery. The caption mentions 5 regions, while the number of points is much larger. Also, the numbers in the text do not seem to correspond with the figure. So I doubt if the correct figure is shown.
Reply: the figure is correct but there is a mistake in the caption, because we have 9 (not 5) areas. The plot includes 9 areas X 3 geometries, i.e. 27 data points.

Page 8229, line 20-21. Again the values in the text do not correspond to table A1.
Reply: we changed the text accordingly

Page 8230, line 9. "The relative error reduction 1Ea/Eb (see Table A2) for CMN and JFJ were 0.18 and 0.17, respectively, and for MHD was 0.42, showing that the two mountain stations are more influenced by the sources present in the study domain". I do not see from these numbers that the mountain stations are more influenced by the sources. I would expect actually a larger error reduction at CMN and JFJ if emissions are updated, because there is more to gain.
Reply: we agree with the referee comment. We modified the text of Appendix A1.4, stating that this result shows the better performance of the model at MHD. An improvement was likely to be achieved for CMN and JFJ whose difference between the a priori and a posteriori emission fields was higher than at MHD. If this is not occurring is clearly to be ascribed to the poorer performance of the meteorological and dispersion model at the mountain sites with respect to MHD.

Page 8232, line 4-5. Note here that since the emissions are homogeneous, also the assigned errors are homogeneous. I find it rather surprising that specific hotspots are retrieved in this inversion (see also major comment #1). For instance, Greece, northern Africa and Norway. These

areas are quite distant from the stations, so transport errors will be considerable. Again, at which resolution have these inversions been done?
Reply: see replies to major comments

Table 1: unit (%)? Should that not include a time unit?
Reply: units added

Table A2: units are missing in general.
Reply: units added

Responses to Referee 2

p.8215, line 23. Since there is some detail given about the analytical procedure, I was curious about the statement about the standard humidification and linearity. Multi- level calibration and linearity are not discussed in the method description. How is this done and how often is linearity checked? Is there an issue with non-linearity in the measurement of the higher level signals? Also, it's not clear to me how a uniformly humidified standard ensures "a close similarity in composition" to air at a coastal station and also a mountain site. Probably not a big deal, but if not then just leave out all the boiler plate that is not supported nor particularly relevant to the current manuscript.
Reply: Multilevel calibration and linearity tests are a hard to solve but strictly needed issue for any analytical method. There are several source of possible non-linearity of the whole method that could lead to a deviation of the ideal constant sensitivity of the detector towards the varying concentrations.
A brief but exhaustive description is reported by Miller (Miller et al., 2008) and can be extended to our analytical instrument. It's not an easy / feasible task to prepare a multi-level calibration curve due to the costs and intensive labour involved in preparing and maintaining a set of multi-level mixtures. A workaround is to test the system by the same mixture but with increasing the sample volume. This test has a limited concentration range verification, but give enough information on the system behaviour and performances. Also the use of real air mixture as a working standard for the systematic calibration and for tests limits the overall bias of the method. For that reasons, the whole system was tested (by the increasing sample volume technique) on the lab prior setting it on the field and is regularly checked every time a sensible part is changed for some reason, such as every time the adsorbing trap is replaced (to verify the breakthrough volume and the recovery/desorbtion efficiency of the new trap), or the MS is opened for the service (change of the electromultiplier horn), or when a new quaternary/working standard tank is installed (for which there is a composition change from the old one). Methylchloroform, fortunately doesn't show any linearity issue to solve -at least within the limit of the test capability described above-; that kind of problems are most likely affecting very low boiling point compounds, for which the system has been anyway optimized, that are analysed at the station with the same instrument. By the way, as a general agreement, the arising non linearity, in case of large mole fraction samples, are accounted for by an increase of the associated uncertainty, since it is calculated as constant percentage of the absolute concentration.

Also, it's not clear to me how a uniformly humidified standard ensures "a close similarity in composition" to air at a coastal station and also a mountain site.

Reply: A partial response to this comment is contained in the above reply: humidification reduces uncertainties and non-linearity effects of the analytical method (the whole system is tested and calibrated with real air mixtures). Furthermore, it has been demonstrated that the humidification procedure of the air during the pumping of standard tanks ensures the stability of the compounds avoiding or minimising the degradation on the pump surfaces and on the surface of the cylinders.

p.8220, line 1. The baseline attribution seems reasonable to me, but I wonder if there are any comments about the baseline irregularities, such as the concentration dip at CMN in Jan 2006, or the longer term cycle at MHD throughout 2002-2003.
Reply: The irregularities in the baseline can be ascribed to the interaction with the OH, whose concentration is not constant in different years and seasons, as well as by changes in global emissions at a temporal scale shorter than a year.

p. 8220, line 19. Obviously it is not just the location of the MCF source nearer to CMN, but also the wind direction and transport.
Reply: yes, we agree with this comment and we modified the sentence accordingly.

p. 8217, line 25. Please add "emission estimate" or similar after "a priori". Try to avoid modelling jargon.
Reply: we modified the text accordingly.

p. 8223, section 3.2.3. I was struck by the distinction of the influence of SEF on the observed mixing ratios. Visually, all of the excursions, with the exception of a period in early 2009 at CMN, are due to the influence of SEF. This would seem to indicate that all of the other information on other emission sources are contained in mixing ratio excursions above background that are in the 0.1 – 0.3 pptv range, which is close to the uncertainties in the background itself. Is this true? It seems remarkable to me. Can this aspect be discussed a bit further?
Reply: The plot in Figure 6 is obtained applying an on-off filter, i.e. it shows when an air mass that reaches the receptor site has first passed over the SEF region (red points). This plot shows how nearly all the most intense enhancements are associated to a contribution from SEF, but this does not mean that all the enhancement is due to a contribution from the SEF region. In other words, any single enhancement could in principle contain also contributions from other regions, but has anyway passed through the SEF region. Anyway, the referee correctly points out how the contributions from other sources are very small at the receptors, being the other sources weaker by order of magnitudes than the SEF and, in the case of CMN and JFJ, further or less upstream than SEF.

p. 8221, line 1. It is noted that the MCF emissions to soil and water are not included in the a priori emission estimates (though it is suggested that these sources are revealed in the analysis). It would be useful to understand the relative magnitude of these sources, since I assume they are just a slower release to the atmosphere compared to direct emission.
Reply: we do not have enough information to quantify the emissions to the atmosphere deriving from the release of MCF to soil or water. The E-PRTR inventory does not provide information on the amount of the released MCF. Furthermore, these sources are too small and remote from the stations to be quantified through the model. Therefore, even if this is an interesting topic, we cannot examine it in depth in this study.

p. 8221, Figure 4. Though I could see the location of the sites on other Figures, the black dots that designate the measurement sites tend to disappear in the dark purple.
Reply: yes, we agree with this comment and we modified the black dots accordingly.

General comment: The appendices provided little information that I thought was particularly important for the analyses. I would expect that the basic findings of the Appendices could be incorporated into the main text and reduce the overall length of the manuscript. Some parts of the appendix discussion, too, e.g. A1.4/Table A.2, were rather opaque and unintelligible to someone (me) with no expertise in evaluating model statistical tests.
Reply: we would prefer to leave these contents in the Appendices in order to make the main text more concise. However, in order to improve clarity, we made substantial changes in the Appendices (See replies to comments from Referee 1).

General comment 2: It would interest me to know what other of the suite of trace gases measured at CMN (especially) can also be traced back to the SEF sources?
Reply: we didn't find any significant correlation with any of the other compounds that are measured at CMN. Furthermore, most of the compounds monitored at CMN are still in use and, as a consequence, they have more distributed sources across the domain.

General comment 3: The emission estimates should be reported with some clear statement of the associated error. If stated in the text, I missed it.
Reply: we included in the text a reference to the Appendix where the associated errors are described in detail.