“Estimation of volatile organic compound emissions for Europe using data assimilation”
M. R. Koohkan, M. Bocquet, Y. Roustan, Y. Kim, and C. Seigneur
acp-2012-786, J. West Ed.

Response to Reviewer #1

We thank the reviewer for his/her useful comments and suggestions.

1 Overview

The work by Koohkan et al. seeks to constrain VOC emissions over Europe using EMEP observations and inverse modeling. The content and scope are suitable for ACP. The authors are quite proficient in their application of inverse modeling methods, being well aware of potential pitfalls, and present interesting alternatives to frequently used methods that provide more valuable results. The methodological results related to the statistically consistent non-Gaussian approach are alone of potential wide interest. That being said, the paper would benefit from additional efforts to interpret the significance of the corrections to emissions. Further tests of the reduced model also appear needed. Lastly, discussion of relevant studies in the literature is lacking. These issues and other more detailed questions are described further below; they should be addressed prior to publication.

2 General comments

• sections 4.2, 4.3. These are great inverse modeling tests, and I just wish the authors could spend a bit more time trying to interpret the results. Does the grouping of errors, or forecast skill, by species indicate anything about the sectors responsible for the errors in VOC emissions? Or the timing (i.e., emissions from a particular season, day of week, or time of day)? Overall the quantitative interrogation of the results is fairly strong; the interpretation is a bit weak.

We agree with the reviewer, the interpretation of the results can appear a bit frustrating. However going much beyond our current conclusions appears difficult. On the basis of a Positive Matrix Factorisation (PMF) approach we have tried to disaggregate the global inversion to propose an information detailed by activity sectors. However this attempt has failed to provide robust results due to the relative collinearity of the emissions sectors chemical profiles in regard to the considered species. The use of source specific tracers, or a larger number of VOC species, could maybe overcome this issue, but such data are not available to our knowledge for this study. Concerning the timing, the relatively low number of observations available restrict the number of parameters that can be considered for the inversion. In our study the effective control variable $\boldsymbol{\alpha}$ does not depend on time. It means that the inversion rely on the a priori temporal distribution of the emissions (given by activity sectors) which is not modified. Our text has been modified to emphasize these restrictions.

In section 4.1.3, we added:
“Indeed, the relatively low number of observations available restrict the number of parameters that can be considered for the inversion. The independence of $\boldsymbol{\alpha}^s$ on time means that the inversion relies on the a priori temporal distribution of the emissions (given by activity sectors) which is not modified.”

We added in the conclusion:
“As an extension to this study, we have tried but failed to disaggregate the inversion among the different emissions sectors in this study using a Positive Matrix Factorisation approach. The chemical profile of the sectors in regard to the considered species were too collinear to get a robust segregation.”

- *I have some issues with the presentation and validation of the reduced model.*
    - *Figure 2: The log-scale shows the correlation across a wide range of values, which is great, but it does sort of hide the fact that errors on the order of 100% or more abound, for isoprene in particular. Can the authors also present the ratio of the direct to adjoint-based values, or the range of this ratio by species? Do any of the issues with the inversion results for isoprene likely relate to errors in the model estimates on the order of 100%? The authors only really consider lifetime to be an issue, but I think their linearization of the chemistry may also be contributing a lot.*

We agree that the linearisation of the chemistry is also contributing but to a minor extent, except for isoprene (see the answer to the next question). For isoprene and about the validity of the approximate adjoint, we believe that the impact of the short life-time is very strong. Think about an observation site and a very distant flux site. If the species is severely abated in between the two sites, then the adjoint-based and direct values of the measurement due to the flux will be small but potentially very different due to numerical truncations (to some extent even with an exactly derived adjoint). We have observed the same effect in (Bocquet, 2012) when a species is submitted to a very highly unrealistic dry deposition, hence with a very short life.

Please note that the scatter diagram (Fig. 2) tends to highlight the outliers since this is not a density plot that would be much more focused on the diagonal, and large errors are much less abundant that it seems.

About the statistics of the direct / adjoint-based valued mismatch, the following table is added to the revised manuscript (in µg.m^{-3} for the means) and the NMSE:

	R	Mean(Sim)	Mean(Adj)	NMSE
C_3H_8	0.9981	0.5325	0.5580	0.0062
NBUT	0.9935	0.6287	0.6776	0.0229
IBUT	0.9967	0.3003	0.3279	0.0389
NPEN	0.9958	0.2369	0.2672	0.0528
IPEN	0.9955	0.2557	0.2825	0.0397
C_3H_6	0.9962	0.1274	0.1404	0.0693
TOLU	0.9927	0.2290	0.2441	0.0387
OXYL	0.9958	0.3167	0.3339	0.0314
MPXYL	0.9901	0.1429	0.1478	0.0526
ISO	0.9728	0.1800	0.1521	0.2540
ACE	0.9995	0.4108	0.4241	0.0037
C_2H_6	0.9986	1.0104	1.0437	0.0029
C_2H_4	0.9953	0.2609	0.2814	0.0236
BEN	0.9979	0.4003	0.4214	0.0081

Looking at NMSEs, it is clear that the case of isoprene is, as expected, problematic, whereas the statistics are quite good for the other species.

Reference:

Bocquet, M., 2012. Parameter field estimation for atmospheric dispersion: Application to the Chernobyl accident using 4D-Var . *Q. J. R. Meteorol. Soc.*, 138, 664-681.

– The inversion results include changes to the emissions by over an order of magnitude in many locations. Surely this will change the local OH and O_3 concentrations. These would then need to be updated periodically throughout the iterative process, at some frequency determined based on tests

(not yet performed) of the extent over which the concentrations in the reduced model respond similarly to emissions changes as in the full model. While the authors do test reduced model relative to the forward model around a single atmospheric state, this is not a test of response of the reduced model to changes in emissions, which is necessary.

The paragraph “4.1.1 A posteriori verification of the model linearisation” deals with this issue. However, we agree that the figure and the indicators meant to support our conclusions were not convincing enough. In the revised manuscript, we now provide the global mean of each oxidant concentration field and the standard deviation of the difference between the oxidant field a priori and a posteriori. The standard deviation is computed considering each time step of integration (10 mn) and each cell of the discrete domain. Moreover, to clarify the presentation we now show the validation results for case C (instead of B2) since it corresponds to the methodology we ultimately advocate. The paragraph 4.1.1 is modified as follow:

“In order to check that hypothesis the a priori and the a posteriori (case C) oxidant concentration fields calculated by the full CTM are compared. Because the inversion of isoprene emission will be later deemed unreliable, the EMEP emissions (a priori emissions) have then been kept for ISO to perform this test. The mean value of the OH concentration is of $1.882 \,10^6$ \# cm^{-3} for the a priori fields and of $1.877 \, 10^6$ \# cm^{-3} for the a posteriori fields. The Pearson correlation between the a priori and a posteriori concentration fields is about 1.00. The standard deviation of the difference between the a priori and the a posteriori fields (computed for all the time steps and cells) is of $2.705 \, 10^4$ \# cm^{-3}. Figure 5 shows the comparison of the concentration fields of OH before and after data assimilation for two representative periods of hundreds hours, the first one in winter, the second one in summer. For the species NO$_3$ and O$_3$, the average values of the a priori concentrations are $1.418 \, 10^8$ \# cm^{-3} and $1.090 \, 10^12$ molec.cm^{-3} respectively. They are $1.418 \, 10^8$ \# cm^{-3} and $1.092 \, 10^12$ molec.cm^{-3} for the a posteriori concentrations. The Pearson correlations between the a priori and a posteriori concentration fields are about 1.00 for both NO$_3$ and O$_3$. The standard deviation of the difference between a priori and a posteriori fields is of $2.627 \, 10^4$ \# cm^{-3} for NO$_3$ and of $4.218 \, 10^9$ molec.cm^{-3} for O$_3$. Further iterations of the linearisation and inversion cycle barely change these results.”

– Turner et al (GRL, 2012) consider the adjoint “footprints” of column observations of HCHO. Despite the short lifetime, chemical feedbacks lead to large (hundreds of km) regions of influence. Are such influences for short-lived species missing from the reduced model?

Thanks for pointing out this reference. This approximation of the reduced model should not (if done properly) impact significantly the extension of the footprints. The magnitude of OH concentrations is correct and life-time of isoprene should not significantly change.

However, we believe it is to be expected that the footprint of a column be much larger than the footprint of the in situ-observation. In situ observation are more precise but come with a much smaller footprint. We have for instance experienced this fact in Wang et al, 2013 where a footprint of a lidar measurement is compared to the footprint of a PM10 in situ measurement.

Besides, one should be careful when comparing with Turner et al. 2012. The large (several hundreds of kilometres) footprints of formaldehyde columns related to isoprene emission regions that are presented by Turner et al. are due to the fact that (1) the oxidation of isoprene leads to formaldehyde formation in a series of oxidation steps (i.e., over time) and (2) there are other sources of formaldehyde (most VOC) that complicate the relationship between formaldehyde and isoprene. This issue of chemical feedback (e.g., reservoir species such as PAN for NOx, formaldehyde formation in successive oxidation steps for isoprene) does not apply to our study because we are relating the emission flux of a given species to ambient concentration measurements of that same species. Since there are no chemical reservoirs for the VOC considered in our study, the radius of influence is solely related to the atmospheric lifetime of the species.

Reference:

Wang, Y., K. Sartelet, M. Bocquet, and P. Chazette, 2013. Assimilation of ground versus lidar observations for PM10 forecasting. Atmos. Chem. Phys., 13, 269-283.

- *There are several areas where discussion of the literature is lacking, such as:*
  - *33222.20: Discussion of previous efforts to constrain VOC emissions is missing many recent works. What about Stavrakou et al (2009, ACP) constraints on glyoxal? or Liu et al. (2012, doi:10.1029/2012GL051645) with glyoxal constrained aromatic emissions? or Zhang et al. 2011 Atmos. Environ. using HCHO remote sensing to constrain VOC emissions in Texas? These are just a few; the authors should include a more comprehensive survey of the literature, and the statement “have not yet included VOCs” should be removed.*

Our statement « such studies… have not yet included VOCs » pertained to the use of ground-level ambient concentrations and not to satellite data, which were discussed earlier in that same paragraph (p. 33222, lines 1-6). Nevertheless, to avoid any confusion, we have rewritten this sentence to be more specific: “such studies…have not yet included a large number of speciated VOC ground-level concentration measurements”. There is a large body of literature on the use of VOC (formaldehyde, glyoxal) satellite data and our objective is not here to provide a complete review of such studies, but to provide some examples to describe this possible approach to estimating VOC emission fluxes. Accordingly, a few additional references have been added: “Use of glyoxal satellite data (e.g., Liu et al., 2012) and a combination of formaldehyde and glyoxal satellite data (e.g., Stavrakou et al., 2009) have also been used to estimate emission fluxes of isoprene and other biogenic and anthropogenic VOC precursors of those oxidized VOC”.

References:

Liu, Z., Y. Wang, M. Vrekoussis, A. Richter, F. Wittrock, J.P.P. Burrows, M. Shao, C.C. Chang, S.C. Liu, H. Wang, C. Chen, 2012. Exploring the missing source of glyoxal (CHOCHO) over China, *Geophys. Res. Lett.*, 38 (10) doi/10.1029/2012GL051645.

Stavrakou, T., J.-F. Müller, I. de Smedt, M. van Roozendael, M. Kanakidou, M. Vrekoussis, F. Wittrock, A. Rihcter, J.P. Burrows, 2009. The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, Atmos. Chem. Phys., **9**, 8431-8446.

  - *33225.23: earlier, more general works describing the adjoints of chemical transport models exist.*

To the best of our knowledge, the use of the adjoint (of the idea of using the adjoint) in an air quality context dates back to Uliaz, 1983. There have been dozens if not hundred of papers using adjoint modelling in the same context since then, especially since the late 90's. The manuscript has been changed acordingly.

Reference:

Uliasz, M., 1983. Application of the perturbation theory to the sensitivity analysis of an air pollution model, Z. Meteorol., 33, 169 – 181.

- *33221.12: The statement "cannot be derived from mass balances. . . conducted at the source of emissions " is an oversimplification. Many works have used mass-balances to constrain VOC emissions based on ambient measurements. It is indeed a complicated process, but often ratios of different species can be used to generate useful results even for non-conservative tracers. See for example any number of papers where aircraft data or measurements from ships downwind of urban areas are used to constrain VOCs during field campaigns. Too many to list here, but I'm sure the authors could find several related to recent campaigns such as CalNex or MEGAPOLI.*

The reviewer may have misunderstood our statement concerning mass balances and we have now clarified this point in the introduction. We are specifically referring to emission estimates performed by conducting a mass balance on the emission process: for example, knowing the sulfur content of gasoline, one may estimate the sulfur oxide (SO_2 and sulfate) emissions associated with gasoline combustion. The reviewer refers to mass balance estimates conducted from ambient measurements and assumptions on the dispersion of the pollutants from their emission point to the point of the ambient measurements. Such estimates involve significant uncertainties although those uncertainties can be minimized (e.g., tunnel measurements where the air flow is well known or ambient measurements calibrated using a pollutant with a known emission rate); those mass balance estimates conducted in the atmosphere with ambient concentrations are discussed elsewhere in the introduction (p. 33221, lines 24-28; p. 33222, lines 6-10). This point is now clarified in the text: "Furthermore, VOC emissions cannot be derived from mass balances of the emission process (as can be done for example for sulfur or heavy metal emissions)…".

- *There has been a lot of inverse modeling work on CH4, which is a VOC. So perhaps the authors should in some contexts refer to NMVOCs.*

Ok. That is what we meant. We clearly state this in the revised manuscript (abstract and introduction).

• A lot of effort is spent dealing with the fact that direct inversion of emissions can lead to negative emissions, and inversion using L-BFGS-B imposes bounds on quantities assumed to be normally distributed. While it is nice to see the development of case C, where the statistics are adjusted to account for this, it seems like a lot of work compared to a much easier, and common, formulation of this problem: use scaling factors = ln(es/eb). The emissions are likely to be log normally distributed to begin with (I'm guessing a simple query of the EMEP inventory would demonstrate this to be the case). Further, the authors can use L-BFGS (i.e., not bounded) as this ranges from minus to positive infinity. At least it would be a nice comparison to B2 and C, hopefully not too much more work.

We agree that the log-normal statistics offer a better match to the true errors met in inventories. But the derivation of the statistically consistent hyper-parameters (as in case C) is a much more difficult matter with log-normal statistics. Actually we have tried it in the context of Winiarek et al, 2012, but failed to do so within a couple of weeks. Indeed, it is much more difficult to sample from the likelihood derived from a multivariate log-normal than from a semi-Gaussian (where efficient sampling techniques already exist). That is the reason why we based our analysis on Gaussian and semi-Gaussian distributions. Had we an efficient technique to compute the likelihood related to the log-normal multivariate distribution, we would have chosen the log-normal for the whole study following Elbern et al., 2007. We agree that for the long term we (or any other team) should address the issue and find a solution (we are considering it).

- *The framework assumes 0% error for the boundary conditions. How much might errors in the boundary conditions be impacting the analysis and projecting onto biases in the inferred emissions? What would the authors assume the uncertainty in their boundary conditions actually are, or how much this matters (especially for long-lived species)?*

The method does implicitly assume errors in the boundary. It is hidden in the estimation of the magnitude of the error covariance matrix (parameters r_s). The errors are estimated as a whole without having to look into specific sensitive factors. However the method does assume the absence of bias, for instance in the boundary conditions.

That being said, to get a rough idea on those uncertainties, we have performed a comparison of the results of the Mozart 2 model, from which are extracted the boundary conditions in our study, to the observations used for the inversion. From this comparison, rough average relative errors are estimated, ranging from 1% (for C_3H_8) to 116% (for C_3H_6). For the long lived species (ACE, C_2H_6 and C_3H_8) rough average relative errors are respectively of -60%, 8% and 1%. From this, we can infer that the inversion of acetylene emissions, with a life-time of 110 days and very uncertain boundary conditions, must be interpreted with caution.

This discussion is reported in the revised manuscript (section 3.2), and a word of caution about ACE emission inversion has been added.

3 Specific comments

- *33221.17: In terms of recent reductions in uncertainties, can the authors provide specific citations?*

We now provide examples of improvements in COV emission inventories based on tunnel measurements conducted in California by various groups over the years: Pierson et al. (1990) for example concluded from 1987 tunnel measurements that there was an underestimation of VOC emissions from on-road mobile sources of a factor of 4; tunnel studies conducted in 1994 led to VOC underestimations by factors of 1.5 to 2.5 (Kirchstetter et al., 1996). The text has been rewritten as follows: “Uncertainties in anthropogenic emissions have been reduced over the years as a result of better characterization of major emission sources; for example tunnel VOC measurements conducted in California have shown underestimates in VOC emissions from vehicles decreasing from a factor of 4 in 1987 (Pierson et al., 1990) to a factor of 2.5 in 1994 (Kirchstetter et al., 1996). Nevertheless, uncertainties still remain (Sawyer et al., 2000).

References:

Kirchstetter, T.W., B.C. Singer, R.A. Harley, 1996. Impact of oxygenated gasoline use on California light-duty vehicle emissions, Environ. Sci. Technol., 30, 661-670.

Pierson, W.R., A.W. Gerlter, R.L. Bradow, 1990. Comparison of the SCAQS tunnel study with other on-road vehicle emission data, J. Air Waste Manage. Assoc., 40, 1495-1504.

33223.19: Stable / accurate advection operators are usually nonlinear. Is the one used here really linear? If so, how is transport accuracy affected?

No, it is a third-order direct space-time scheme with a Koren-Sweby flux limiter. The impact of the nonlinearity on the approximate adjoint is contained in Fig. 2, and is therefore relatively controlled.

(The intend of a nonlinear flux limiter is to stably model a linear process.)
An interesting (?) discussion on this very topic with numerical results is given by Bocquet, 2012.

Reference:

Bocquet, M., 2012. Parameter field estimation for atmospheric dispersion: Application to the Chernobyl accident using 4D-Var . *Q. J. R. Meteorol. Soc., 138, 664-681.*

- *It is never mentioned what the temporal resolution of the observations is. Hourly? Daily? And at what temporal scales are they used in μ_s? There must be some limitations to the methods used here which rely on the full jacobian H given that d_s adjoint calculations are required.*

The duration of the observations is variable, from 20 minutes to 24 hours. We now provide in the text the range of sample duration existing in the used observation database. Each observation is used with this own duration, both for comparison to model and adjoint computations. As discussed in paragraph "2.2 The source receptor model", the method requires d adjoint simulations to build the Jacobian H. This means that the computational cost of the method increases linearly with the number of assimilated observations. However since each adjoint calculation is independent, it is trivial to parallelise this method.

This has been added to section 3.1:
"The sample duration ranges from 20 minutes to 24 hours depending on stations and species (see http://www.nilu.no/projects/ccc/index.html for details)."

- *33230.7: I appreciate that calculation of Eq. 15 is computationally intensive, but the phrase "integral over the positive cone" doesn't mean anything to me. Can this be explained more carefully for the lay audience?*

Ok. Thanks for the suggestion. The positive cone is the positive quadrant, the set of all positive-valued vectors. Geometrically, it forms a cone. This is the reason why the integral is not trivial to compute. This is made plain in the revised manuscript.

- *33227.15: Why is this not expected? I would assume that emissions within particular geo-political boundaries may be correlated, or that emissions for specific sectors would be correlated across broad spatial scales. Also, since the model resolution has yet to be introduced here, it wasn't clear when readying this what might be considered long-range.*

That the emissions are correlated is an issue. That their errors also are is another one. The two issues may be connected. Anthropogenic inventories are the results of aggregations of statistics over sectors often made of point-wise statistics (factories, urban areas, roads). In the absence of a bias in the way the survey is carried out, the correlation should be short-ranged. Biogenic errors in inventory are much more prone to correlation because they depend on vegetation models making systematic errors over large biomes.

We have further precise: "Because of the way the anthropogenic inventory surveys are performed, the anthropogenic emissions of VOCs are not expected to induce long-range correlation (a few tens of kilometers) in the errors."

- *33231.24: How are boundaries at the top of the model handled? Are these also taken from MOZART 2?*

Indeed, the boundaries at the top of the model are derived from MOZART 2. The text is modified to explicitly mention this point.

“The initial and boundary conditions (top and lateral faces of the domain) concentration fields are obtained from the global”...

- *33232.5: Could the authors say a bit more here what is meant by approximate adjoint? Is there similarity to the approximate adjoints of Singh and Sandu (2012 Computers and Geosciences)?*

We meant that the adjoint of the CTM is approximated by the discretisation of the continuous adjoint equation which in the linear case is easy to obtain. Besides it can be run backward in time using the original CTM on the condition that wind fields are reversed. This is based on ten years of practice in our group (some references of which could have been mentioned in Singh and Sandu, 2012). All of these relate to the same idea. One distinction can be made on whether the approximate adjoint is used to compute the gradient (Bocquet, 2012; Singh and Sandu, 2012; but also many references in the implementation of operational 4D-Vars in meteorology), of the full Jacobian of linear systems (see Krysta and Bocquet, 2007 for a comparison between an exact adjoint and an approximate adjoint of a CTM).

We have added: “The approximate adjoint is built by discretising the adjoint of the continuous transport equation. Since, in general, discretisation and adjointisation are operators that do not commute, this adjoint is only approximate. However, it turns out that its implementation re-use the forward CTM but with reversed wind fields and a backward integration in time.”

Reference:

Krysta, M. and M. Bocquet, 2007. Source reconstruction of an accidental radionuclide release at European scale .*Q. J. R. Meteorol. Soc.*, 133, 529-544.

- *Figure 6: The wording and description of this essential figure is a bit odd. “Normalized” and “correction” may be redundant here. Regardless, I’m not clear if it is showing correction factors, or if showing inversion emissions normalized by EMEP emissions. Further, discussion in the text mentions the large negative source for ISOP in case B1. But from my understanding of this figure, there are also small negative corrections for all inversions for other species, e.g., toluene.*

How are we to tell from this figure if these negative corrections are large enough to lead to actual negative emissions? Would that be at the -1 level? Again, clarifying what is actually being shown in this figure would be useful in this regard.

The caption of the figure is correct. Let ΔE be the correction to the emitted mass and E_0 be the original emitted mass, the normalized (with respect to the original emitted mass) emitted mass correction is $\Delta E/E_0$. There is no correction if it is equal to zero. If it is equal to 1, the corrected emitted mass is twice the original emitted mass. If it is equal to -1, the corrected emitted mass is zero. If it is less than -1, then the corrected emitted mass is negative (i.e., an atmospheric sink instead of an emission flux; it is the case only for isoprene with case B1, which as mentioned in the introduction of section 4 can lead to negative emissions fluxes). To clarify Figure 6, we have added the following explanatory text in the caption: “The normalized emitted mass correction represents the correction to the emitted mass ΔE divided by the original emitted mass E_0. Therefore, there is no correction if $\Delta E/E_0$ is equal to zero. If it is equal to 1, the corrected emitted mass is twice the original emitted mass. If it is equal to -1, the corrected emitted mass is zero. If it is less than -1, then the corrected emitted mass is negative”.

- *332326.8: Exactly! This cannot be said enough, and is unfortunately often omitted from many inverse modeling studies, so I commend the authors here.*

Thanks!

- *Eq 17: Can the origin of this equation be shown?*

Yes. But we have removed this section and we have replaced it by an estimation of the posterior uncertainty.

- *Figure 7: This would have a lot more meaning if the authors included a third column showing the base emissions themselves. Otherwise, it is hard to tell if the relative changes are significant or not.*

Figure 7 has been updated according to your suggestion. We have added a column showing the base emissions and we now present results for acetylene, ethane, isobutane, ethylene and isoprene. Our text in the section “4.1.4 Spatial distribution” is modified as follows:
“...the prior EMEP emissions for the species acetylene (ACE, with an average lifetime of 110$\,$days), ethane (C$_2$H$_6$, with an average lifetime of about 60$\,$days), isobutane (IBUT, with an average lifetime of about 7.5$\,$days), ethylene (C$_2$H$_4$, with an average lifetime of about 1.45$\,$days) and isoprene (ISO, with an average lifetime of about 1.7$\,$hour), respectively. Obviously the corrections extend much farther from the monitoring stations for the long-lived species, such as ACE and ETH than for the short-lived species...”

- *Figure 7: For isoprene in France, emissions near the measurement locations show large, localized increases, surrounded by more broad decreases. What is the reason for this?*

There is a bias to correct in the difference between simulation and observation. Because of the short lifetime of isoprene, the corrections to the flux cannot be corrected too far away from the stations, so that the corrections are located near the stations. These corrections might be overestimated, since it is cheaper (from the regularisation of the inverse modelling system point of view) to add stronger fluxes close to the observations site than to spread them over the whole domain.

This has to do with the so-called co-location effect, see in particular Saide et al., 2011 for a discussion.

Reference:

Saide, P., M. Bocquet, A. Osses, and L. Gallardo, 2011. Constraining surface emissions of air pollutants using inverse modeling: method intercomparison and a new two-step two-scale regularization approach, Tellus B, 63, 360-370.

- *33221.19: How large? 50%? 500%?*

Uncertainties have been estimated to be about a factor of two for isoprene emissions (Guenther et al., 2000, 2006; Warneke et al., 2010). We modified the text as follows: “Furthermore, uncertainties on the order of a factor of two are associated with isoprene emissions, and greater for other biogenic organic compounds (Guenther et al., 2000, 2006; Warneke et al., 2010), due to the difficulty…”.

References:

Guenther, A., C. Geron, T. Pierce, B. Lamb, P. Harley, R. Fall, 2000. Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America, Atmos. Environ., 34, 2205-2230.

Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P.I. Palmer, C. Geron, 2006. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210.

Warneke, C., J.A. de Gouw, L. Del Negro, J. Brioude, S. McKeen, H. Stark, W.C. Kuster, P.D. Goldan, M. trainer, F.C. Fehsenfeld, C. Wiedinmyer, A.B. Guenther, A. Hansel, A Wisthaler, E. Atlas, J.S. Holloway, T.B. Ryerson, J. Peischl, L.G. Huey, A.T. Case Hanks, 2010. Biogenic emission measurement and inventories determination of biogenic emissions in the eastern United States and Texas and comparison with biogenic emission inventories, J. Geophys. Res., 115, D00F18, doi:10.1029/2009JD012445.

Editorial comments

- *33220.4: comparison to a standard*

Done.

- *33221.9: strategies become implemented*

Done.

- *33222.11: The wording here implies that only the papers discussed after this point used outputs from air quality models as part of the inversion, when such models were as well integral to the satellite-based inversions mentioned at the beginning of the paragraph. Some adjustment to the wording or organization of this paragraph would be useful.*

This paragraph is not focused on inversion methods for estimating VOC emissions but on all types of methods (including of course inverse modelling) used to evaluate the error in VOC emission fluxes. Some of those methods use only observations (e.g. tunnel experiments), others use a combination of models and observations. Nevertheless, we have modified the sentence on satellite observations to clarify that they are used in combination with models.

“Satellite measurements in combination with various types of models have been used to assess VOC emissions. However, such techniques…”

- *I suggest picking one of the names “prior”, “a prior”, “background” or “first guess” and sticking with this throughout the paper.*

To some yes, we agree. But these terms do not necessary mean the same:
- Prior applies to observations and background errors (prior to using observations).
- Background pertains to any prior information that is not related to the observation.
- First guess is the mean (a state/control vector) of the background.

We have tried to make uniform the terminology throughout the manuscript, as much as possible. We have tried to avoid first guess (replaced by background, except for one occurrence), and avoid the excessive use of prior/a priori.

- *33223.4: the observations and*

Done.

- *33224.8: Would recommend eliminating “relevant to our study: they are not ”*

Done.

- *33228.3: It wasn't clear from reading this paragraph alone how the second solution was obtained. Only later, in discussion of the inverse modeling results, did it become clear that the second approach is a gradient-based iterative minimization of L_s using L-BFGS-B. It would be nice to revise / expand this paragraph to make the method clear here.*

Thank you for the suggestion. The paragraph has been clarified:

"The second solution is obtained by minimising Eq.(7) on $\boldsymbol{\alpha}^s$, under the positivity constraint of each one of its entry $\boldsymbol{\alpha}^s_l$. A bounded quasi-Newton gradient-based minimisation scheme will later be used. As opposed to the Gaussian case, the retrieved scaling factors $\boldsymbol{\alpha}^s_l$ cannot be negative. This solution implicitly assumes a truncated Gaussian distribution for the background error statistics."

- *33232.22: Can the phrase "value screening" be clarified?*

Done (this remark has also been made by Reviewer 3).

- *33237.11: One way to understand this is that . . .*

Done.