26 August 2014

Revisions for “Estimating sources of elemental and organic carbon and their temporal emission patterns using a Least Squares Inverse model and hourly measurements from the St. Louis-Midwest Supersite” by B. de Foy, Y. Y. Cui, J. J. Schauer, M. Janssen, J. R. Turner, C. Wiedinmyer, Atmospheric Chemistry and Physics Discussion, 2014.

Dear Editor,

Thank you for the reviews of our manuscript. Below you will find a point by point reply to the comments from the reviewers along with a list of the changes made to the text. We feel that we have responded to the comments and that the paper has improved as a result.

The complete original comments from the reviewers are in black below, and the responses and modifications made to the manuscript are listed in blue.

Sincerely,
B. de Foy, Y. Y. Cui, J. J. Schauer, M. Janssen, J. R. Turner, C. Wiedinmyer

Anonymous Referee #1, Received and published: 17 June 2014

1. Overview

The manuscript by de Foy et al. uses least-absolute value regression to constrain emissions of EC and OC that contribute to year-long hourly measurements in St. Louis. There modeling setup allows them to specifically investigate temporal emissions patterns in some detail. Overall, the manuscript is very well written and easy to follow. The introduction and abstract might be enhanced a bit in terms of framing the value of their work in examining an already much studied dataset. I believe the biggest scientific issue I see is the discussion of sinks, which are mentioned by the other reviewers. Mostly I have comments and clarifications about the inversion methods. This manuscript will be suitable for publication after revision to address the comments and corrections noted below.

2. Comments

• Title (and throughout): I feel like using the term “least squares inverse” as the name of the method in the form of a proper noun is a bit odd. The least squares method is ubiquitous, and by definition it is an inverse molding approach. So it doesn’t seem to warrant capitalization in this form.

You are right, we have changed to lower case and/or reworded as appropriate to refer just to "the inverse model."

• 12032: Regarding the IRLS scheme, this is in general a method to perform least absolute value regression, i.e., L1 regression. The textbook by Aster shows this equivalence. It is thus further confusing that the authors would refer to their method as “Least Squares Inversion” when in fact it is actually a least-absolute value regression.

Yes, the weights in the IRLS scheme can be chosen to implement L1 regression as described in Aster et al., 2012. In this paper, we use the weights to eliminate the influence of outliers which is a form of robust least squares, but we do not approximate L1 regression. As described above, we use lower case and have reworded to refer more generically to "the inverse model."

• It might be useful if an introductory sentence was added to the beginning of the abstract to help emphasize the value of this study.
Thanks for the suggestion, we have added the following sentence:
"Emission inventories of Elemental Carbon (EC) and Organic Carbon (OC) contain large uncertainties both in their spatial and temporal distributions for different source types."

• 12021.13: A subtle point on methodology: it is not necessary for error covariances to be diagonal in order for a Bayesian inversion to be cast as a standard least squarest problem. See for example the textbook by Aster, wherein augmented matrices involving the square roots of the error covariances are used to turn the standard Bayesian cost function into a standard least squares regression (Chap 11 perhaps? Sorry, I don't have it with me.). Maybe it is just then not clear what the authors mean by "single" in this context.
Yes, we were wrong to imply that this was necessary. Although diagonal matrices make the math more straightforward, Aster et al., explain how to do this with non-diagonal matrices. The following phrase was removed from the abstract: "and by using diagonal error covariance matrices,"

• 12021.25: The text refers to "the inventory" as if we knew specifically of one being discussed (e.g., NEI, or LADCO), but we don't yet at this point.
Details added above in the abstract: "using known emissions inventories for point and area sources from the Lake Michigan Directors Consortium (LADCO) as well as for open burning from the Fire Inventory from NCAR (FINN)."

• Could the authors comment a bit more on the disconnect between the time periods covered by the different emissions inventories, and the observations? There have been significant trends (mostly reductions) in BC concentrations in the U.S. in the past decade. To what extent are inventories for years several after 2002 possibly impacted by these trends? Would this explain some of the deficiencies notes e.g., on lines 12038.23?
Yes, some of the discrepancy can be due to the temporal disconnect.
We have expanded the sentence starting "Although..." into its own paragraph as follows:
"EC and OC have experienced a downward trend in the US, with around 1% to 2% decreases per year Hand et al., 2013. This means that emissions calculated based on 2002 measurements could be expected to be 5%to 10% higher than an emissions inventory for 2007. Although emission inventories existed for 2002, it was felt that the considerable improvements and developments that went into the LADCO 2007 inventory meant that this would be a better choice for the prior, and that consequently the 2008 NEI was the most appropriate comparison point to the prior. Nonetheless, the temporal discrepancy should be borne in mind when interpreting the results."

We have added a caveat in the discussion:
"The large reduction in emissions during fall and winter is unlikely to be realistic, even accounting for the fact that the measurements are from 2002 and the inventory for 2007, and so it suggests that there is an issue with the current representation of the emissions in the inventory and/or with the simulated wind transport from the sources to the receptor site."

And in the conclusions we have specified that we are working with the 2007 LADCO inventory:
"The inversion was based on the 2007 LADCO inventory."

• 12022.7: An additional (better?) citation for BC-specific health impacts is

Janssen et al., Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environmental Health Perspective, 119(12):1691-1699, 2012.
Reference added, thank you.

• 12024: At this point in the manuscript, it seems that many previous works have used this dataset to look at source attribution questions. It might be good to state here what the angle of the present work is in terms of questions that remained to be answered or additional analysis that will be brought to bear.
Thank you for the suggestion, we have reworded the paragraph to be clearer about what we are doing in comparison with the studies cited:
"In this paper, we study the same year-long hourly time series of EC and OC measured in East St. Louis. We seek to obtain improved estimates of the diurnal and monthly emission profiles of specific types of sources by combining forward simulations of EC and OC concentrations from emissions inventories with the measurements using an inverse model. This is carried out for five different source categories as well as for emissions from open burning."

• 12025: Given that later parts of the article emphasize the importance of micrometeorology, to what extent to the authors expect that the meteorological data from 15 miles away from the measurement site are relevant?
There are significant discrepancies, especially for the super stable events associated with the low-level jet, as described in Sec. 3.1. This is why we use KCPS which is only 3 miles away, and which was found to be in agreement with the onsite data, but to have fewer missing data.

• 12026: Could it be clarified how these were updated?
We have added two references with more details, text adjusted as follows:
"Point source emissions were specified using 2007 CEM data with updated temporal profiles to include adjustments for weekend/weekday emissions while still providing a solid platform for future projections (Edick et al., 2006)."
And:
"Non-Road emissions were updated to reflect higher agricultural equipment emissions during the spring and fall season rather than the default of a single summer maximum based on midwest crop calendars and tilling, planting, pesticide application and harvesting cycles Thesing et al., 2004."

• 12029: I'm not sure if CFA is a widely used technique. Can the authors explain, in a sentence or two, what this does?
Sentence added: "Concentration Field Analysis is based on scaling the Residence Time Analysis at each time step with the concentration at the measurement site. The sum over the entire measurement period is then normalized with the Residence Time Analysis. This highlights air flow patterns that are associated with high receptor concentrations."

• 12031.17: Another minor point about the methods: this statement is true only if the error covariance matrices can be reduced to alpha I, where alpha is a constant and I is the identity matrix. This is a more restrictive condition than just being diagonal.
Yes, this is true for a single value of alpha. In our case, we use a vector s containing different values of the regularization parameter, in which case we can represent any diagonal matrix, not just alpha times an identity matrix. The text was adjusted as follows:
"In practice, alpha can be replaced by a vector of parameters s that scales each term in x within the L2 norm. In this way, the method was shown to be equivalent to a Bayesian derivation when diagonal

error covariance matrices are used (de Foy et al., 2012, Wunsch 2006, Aster et al., 2012)."

• 12041.5: An alternative explanation is that estimates could be stabilized with more prior constraints, i.e., the current setup is under-smoother or ill-conditioned.
Yes, text added:
"but also that the estimates could be stabilized with more data, or with stronger constraints on the prior."

• 12042.2: I'm concerned about the large relative increases in emissions, factors of 20 and 30. This again seems like the system is under constrained (either to lack of data or lack of prior constraints). At the very least, these posterior estimates are vary inconsistent with the a priori uniform error assumption of 100% (12033.8).
Yes, these large adjustments definitely suggest the need for more work. Note that the uncertainty in the prior is equivalent to a factor of 33 for EC (see Sec. 2.4), which is in line with the results.
This section was expanded to include these concerns:
"As shown in Fig. 9, uncertainty estimates based on bootstrapping are largest for open burning, with 20%. However, adjustment factors of 20 to 30 suggest either that the uncertainties are underestimated, or that the inversion of these emissions are underconstrained. Overall, these results suggest that future work with more surface measurements and emissions estimates from more recent satellite sensors are needed to improve the inverse estimates, but that nonetheless emission factors in FINN should be revised upwards."

• 12033.8: It seems odd that all emissions would be ascribed equal a priori uncertainty. Wouldn't we expect some sectors to be constrained much more or less than others?
There are separate regularization parameters for the RTA grids, for the LADCO emissions simulations and for the open burning simulations. Within each category, we felt that there was insufficient information to ascribe different uncertainties, although in future work we could use different values for example for the point sources which are better characterized than the other categories.

• 12041.22: Alternatively, generating and using different meteorological fields from WRF using different physics schemes could provide some diversity to test the impact of the dynamics on the results.
Yes, text added:
"Alternatively, the uncertainty could be estimated by running the inverse model with different sets of WRF simulations that used different options, for example by generating input meteorological fields with different boundary layer schemes."

• 12034.24: Could the authors clarify which features of the inventory that they know about are being referred to here?
Yes, text clarified by relating the comment to Fig. 2: "but is puzzling given that southern Illinois does not stand out as a large source region in Fig. 2."

3. Corrections
• 12035: Low Level Jet -> low-level jet: Changed, thank you.
• 12040.6: has a more -> has more: Changed to "has a more pronounced annual variation" (not "variations").
• 12043: The phrase "LADCO inventory is slightly larger than the NEI" is written

twice in this paragraph.
Wording modified as follows: "For OC, the largest category by far in both inventories are the Other sources which are 17% higher in the LADCO inventory. These include residential wood and waste combustion, non-vehicle road emissions and food cooking (estimates of agricultural burning are high in the NEI but low in the LADCO inventory)."

Anonymous Referee #2, Received and published: 10 June 2014
General comments
This study examines one year of hourly measurements of elemental carbon (EC) and organic carbon (OC) from the St. Louis Midwest Supersite. Using a least squares inverse model and atmospheric transport modelling, the authors estimate the emissions from different source types. In addition, the difference between weekday versus weekend emissions and the diurnal cycle are resolved. The authors find reasonably good agreement of the emissions with the prior estimate, but that open burning emissions are likely significantly underestimated in their prior. I recommend this manuscript for publication after the following comments have been addressed.

Section 2.4, which describes the inverse modelling method, is difficult to follow. In particular, there appear to be a number of inconsistencies in the definitions of the variables in Eq. 1 and the physical units of these. Details are given in the specific comments below. I suggest that the authors review this section carefully to make it clear to the reader exactly what was done.
Specific comments
P12021, L22: Specify whether this is emissions of EC, OC or both.
Text adjusted: "Emissions of EC and OC in the St. Louis region"

P12023, L27: What is meant by “smoking vehicles”. Is this synonymous with “vehicle emissions”?
We mean "High-emitting smoker vehicles," text replaced.
For more information, please refer to the papers cited in Bae et al., 2006.

P12024, L10: This sentence appears to be incomplete, for instance, which “Potential Source contribution function” or were there more than one. Please also add a reference.
Text clarified: "Lee et al., 2006, used the Potential Source Contribution Function method based on back-trajectories to show that sulfate levels at the site were impacted by the Ohio River Valley, while nitrate levels were associated with transport from the west and northwest."

P12024, L15: Do the authors mean at the St. Louis Midwest Supersite, if so, this should be specified.
Text changed, thank you: "measured at the St. Louis Midwest Supersite."

P12024, L18-20: Please rephrase this sentence to make it clearer that EC is a passive tracer whereas OC is produced also in the atmosphere. The way it is written it is not clear where OC is “created” and it implies that EC is not emitted, which of course is not the case.
Text clarified: "As discussed above, EC is not formed in the atmosphere but rather emissions are transported until they are removed by deposition such that they can be simulated as passive tracers. In contrast, OC is both emitted and produced in the atmosphere."

P12026, L7: Fig. 2 is referred to before Fig. 1, suggest either referring to Fig. 1 beforehand or reversing the order of the figures.
Thank you for pointing this out. Reference added to Fig. 1 in Sec 2.1: "Fig. 1 shows the location of the measurement site."

P12026, L9: Table 1 has not yet been referred to, suggest reversing the order of Table 1 and 2.
Thanks, Table #1 moved to #3.

P12026, L20: Suggest that the authors start this paragraph by mentioning that an alternative emission dataset was prepared, to compare with the LADCO one, from the NEI data. Otherwise it is difficult to follow the text.
Thanks for the suggestion, text changed: "In order to have an additional comparison to the LADCO prior emissions and the inverse model results, the 2008 National Emissions Inventory (NEI) version 3 was obtained from the US Environmental Protection Agency."

P12028, L20: Is dry/wet deposition accounted for in the FLEXPART simulations? If yes, please indicate the scheme for this. If no, please comment on how this may impact your residence time analysis.
RTA considers only air mass movements which are identified from passive tracers without deposition, as discussed in Ashbaugh et al., 1985.
Text added: "The particles were treated as passive tracers with neither wet nor dry deposition."

P12029, L12: What is the impact of not using the aerosol module in CAMx, for instance, does this mean that dry/wet deposition of aerosols is not accounted for. If so, how will this impact your simulations?
We used wet and dry deposition in CAMx, and apologize for failing to mention this in the text.
Text added: "Dry deposition was calculated using the Zhang et al., 2003 scheme, and wet deposition using the standard scheme in CAMx."

Text added to clarify the limitations of our study:
"Both EC and OC are therefore simulated as passive tracers with wet and dry deposition. This is adequate for EC, and so the inverse model results can be straightforwardly compared to the emissions inventories. In contrast to EC, there is extensive formation of OC in the atmosphere which is not simulated in our model. This means that the inversion will not distinguish between primary and secondary OC, and that results are therefore best interpreted as impacts at the measurement site rather than as emissions at the source location. It also means that we are not able to evaluate the non-linear interactions of different plumes together."

P12029, L25: Please specify the “two-step” method. In Rigby et al. and Rödenbeck et al., an Eulerian model is used to take into account the influence on the air masses which is not accounted for in the time frame of the back trajectories (in this study 4 days). It is not clear in this study, however, how the background influence is accounted for or how the Eulerian model simulations would provide the background influence information. Also, please note that Rigby et al. 2011 actually use a 1-step method.
Sorry for the mistake. It is not our intention to describe alternative methods at this point but rather to make sure that they are cited.
EC and OC levels are very low in St. Louis and the background can be assumed to be close to 0, see Fig. 4. We therefore do not worry about the influence of areas beyond those described in Sec. 2 - these

are negligible, especially when you take deposition into account. Both Rigby and Roedenbeck were looking at global emissions of inert gases.
The paragraph was expanded as follows:
"Inverse models based on back-trajectories alone include Stohl et al., 2009, Brioude et al., 2011 and Brioude et al., 2013. This work combines back-trajectories with Eulerian simulations, and in this respect is similar to the methods presented in Rigby et al., 2011 and Roedenbeck et al., 2009. The purpose of combining the Lagrangian and Eulerian simulations for Rigby et al., 2011 and Roedenbeck et al., 2009 was to combine global transport of inert species with higher definition impacts from specific locations. In our case, the background levels of EC and OC are very low (see Fig. 4), and we expect minimal impacts from sources outside the study area. The purpose of combining Eulerian with Lagrangian simulations is therefore to estimate adjustments to known emission inventories with the Eulerian simulations, and to estimate impacts from unknown area sources in an overlapping domain with the Lagrangian simulations."

P12030, L6-8: If I have understood correctly, the simulations are not made using actual meteorology of each hour/day. Are these then average hours and days for the given year (i.e. 2002) or other? Please specify. Also, please specify if this was using CAMx or FLEXPART.
We apologize for the lack of clarity in our explanation. In fact, we perform hourly CAMx simulations for each source group for each time chunk. Each of the 600 input times series is a year-long hourly timeseries. We hope that the following paragraph is clearer:
"Hourly Eulerian simulations with CAMx were performed for the five different source groups in the LADCO inventory: On-Road, Non-Road, MAR, Other and Point Sources. Because we are interested in evaluating the temporal profiles of the sources, we carry out separate simulations for emissions during different times of the day and different days of the week. The time slots were selected based on the diurnal profile used in the emissions inventory: 11:00 p.m. to 05:00 a.m., 05:00 a.m. to 08:00 a.m., 08:00 a.m. to 02:00 p.m., 02:00 p.m. to 06:00 p.m., and 06:00 p.m. to 11:00 p.m. Days of the week were split into a weekday group and a group containing Saturdays, Sundays and Holidays. As an example, an hourly time series of concentrations was obtained from a CAMx simulation with On-Road emissions only between 05:00 a.m. to 08:00 a.m. on weekdays. With 5 source groups, 5 time slots and 2 day types, this means that there were 50 CAMx simulations. We are also interested in the annual profile of the emissions, and so we divide the 50 resulting concentration time series into 12 months for a total of 600 input time series into the inverse model."

We also clarified that the open burning time series are obtained with CAMx:
"The open burning emissions are included in the inversion as 6 time series simulated by CAMx for the entire year for the 6 geographic sectors shown in Fig. 3. We also include a CAMx time series representing impacts from biogenic emissions, as discussed in Sec. 3.2."

P12030, L19-22: Please specify that these were the averaged timeseries since they are the averages of different the weekdays/weekends and 4 timeslots throughout 2002.
Hopefully the new paragraph clarifies this point. There are no averages, just lots of year-long hourly time series.

P12031, L1-2: This sentence does not follow from the preceding one. What does "these" refer to in this sentence, it is not clear.
Thank you for pointing out the ambiguities in the explanation, we have changed the paragraph as follows:
"The inverse model derives a posterior estimate of emissions based on the Eulerian simulations that

used the emissions inventory as a prior. In addition, the inverse model uses the Lagrangian simulations to derive an estimate of sources that may be missing from the inventory. This is done by using the polar grids of Residence Time Analysis that represent the impact that an emission in a given grid cell would have at the measurement site. As all the known sources were already included in the CAMx simulations with the emissions inventory, we use a field of zero prior emissions for the polar grids from the Lagrangian simulations."

P12031, L10: the elements of x' are not "factors" as there is no multiplication involved. Please use rather "vector of emission corrections" or other.
Changed, thank you.

P12031, L11: From the preceding paragraph I understood that you use the emissions estimates themselves and that these are not parameterized. Please clarify, are the elements of x the prior emissions or prior parameters?
The term parameter was meant in a very general way. A large range of different metrics of emissions can be used in x, so long as they yield concentrations when they are multiplied by the corresponding column in H. In this paragraph, we would prefer to keep the description general. In the following paragraph, we explain what types of parameters are actually used in the study (bearing in mind that in a broad usage, a parameter could be an emission in tonne/year, or a scaling factor of some kind).
Sentence added:
"The individual entries in x can take different forms: they can be actual emissions in units of mass per time, or they can be non-dimensional scaling factors."

P12031, L11 and L20: These sentences are inconsistent. In L11 the authors state that H is the operator to convert emissions to concentrations, while in L20 the authors state that the columns of H contain the timeseries' from the CAMx and FLEXPART simulations (in units of concentration for CAMx and emission sensitivity for FLEXPART).
We hope that having clarified that x can have different types of elements the inconsistency has been resolved. Please see changes to the text in the previous comment. What matters is that H times x yields concentration values that are consistent with y.

P12031, L20-25: Having read this section, I understand that the authors have made 606 + 2880 simulations hourly for all of 2002. Is this correct? I suggest that they re-write the section P12030, L17-27 to make this clearer as otherwise, the reader will think that these are averaged timeseries, in which case, the description of H does not follow.
Yes, your understanding is correct. Hopefully your suggestions for (P12030, L6-8) helped to clarify this up front. We have added the following to clarify things as you suggest, thank you:
" all of which are hourly time series for the whole of 2002."

P12031, L27: There is an inconsistency here with L10-11 of the same page. In L10-11 the authors state that x0 are the prior emissions and x are emissions. However, in L27 the authors state that x are scaling factors such that posterior emissions are equal to x*x0. Furthermore, given the definition of H, Hx (for the 2880 emission sensitivities) cannot be equivalent to y if they are scaling factors. Lastly, again given the definition of H, the elements of x cannot all have the same units. Please clarify.
The sentence was replaced with the following one, which is in line with the changes made in response to the comments above:
"For the CAMx time series, the entries in x are scaling factors on the LADCO emissions that went into

the CAMx simulations. For the FLEXPART polar grids, the entries in x represent emissions."

P12031, L25-26: the vector x includes entries for the “gridded area sources” as well as the “group sources”. Please be clearer about which sources are gridded and which are grouped as to be sure that the source types are not being double counted.
The polar grids cover the same area but have a zero prior. They are meant to provide a way for the inverse model to identify areas with emissions that are not represented in the LADCO prior.
The following text was modified to clarify this point:
"As all the known sources were already included in the CAMx simulations with the emissions inventory, we use a field of zero prior emissions for the polar grids from the Lagrangian simulations."

P12037, L12-13: I do not understand this sentence. What is meant by “yielded the most consistent estimate of impacts in the inversion”?
We tested CAMx simulations using different biogenic species for different model configurations, and we selected the one that gave the best improvement to the overall match of the inverse concentration time series with the measurement time series. This was the CG5 category.
The sentence was changed as follows:
"For OC, we tested different biogenic components and found that condensable gases category 5 “CG5” yielded the best inverse time series of OC compared to the measured time series."

P12037, L19: Is the 19% for the “Other” category refer to the posterior emissions. This should be stated and possibly also mentioned at L4-5.
We have added an explanatory sentence at the top of the section to help clarify this:
"By impacts, we mean the surface concentration of EC or OC at the measurement site that are due to transport of particular emissions to the site."
In L4-5, we clarify that we are talking about contributions to the simulated impacts by either the prior or the posterior: "simulated" added in three places in the paragraph.
In L19, likewise, we mean that 19% of the posterior simulated concentrations are due to emissions from the "Other" category.

P12037, L23-29: Bootstrapping will provide an estimate of the uncertainty that comes from sub-sampling the data. However, there is also data selection in that outliers of more than 3 SD are removed. Have the authors investigated the sensitivity of the results to outliers and the selection criteria?
We have expanded the description of the bootstrapping and included a new figure to show the uncertainties in the results, please see the response to Reviewer #3, General Comment #6.

We have performed various tests with different selection criteria and found that while the least squares method is sensitive to outliers, the results of our inversion are robust relative to the different ways of handling those outliers. Since the method we employ is widely accepted and basic "textbook" material, we feel that it is not necessary to further justify it here.

Section 3.3: Suggest adding subheadings to this section to make it clear what type of emission is being discussed e.g. On-road emissions, Non-road emissions, etc.
Done, thank you for the suggestion.

P12044, L12: What is meant by “annualized” does this simply mean the emission for each period given as the emission per year? Please clarify.

This means that we take the emissions for 4 months and multiply by 3 to obtain an emission rate for a whole year that would be equivalent to having the same emissions as the 4 month chunk for the rest of the year.
The following has been added:
"The emissions rates are annualized by multiplying the emissions in tonnes per 4 months by 3 in order to have emissions in tonnes per year. This yields the annual emission rate that would be obtained if the emissions of the 4 months continue for an entire year."

Technical comments
P12021, L2: “a year-long” (since it is only one year). Changed "based on one year of hourly measurements" (There were 2 time series, one for EC and one for OC).
P12022, L20: replace “under-prediction” with “underestimation” as it is something can only be predicted or not and not “under” or “over” predicted. Changed.
P12025, L6: “mixed-use neighbourhood”. Changed.
P12025, L17-18: Please use SI units, i.e. metric units throughout. Changed.
P12029, L20: Please correct: Rödenbeck et al. 2009. Changed.
P12040, L2: “during daylight hours” (remove “the”). Changed.
P12044, L16: missing full-stop after “December”. Changed.

Anonymous Referee #3, Received and published: 6 June 2014
This manuscript presents an inversion for emissions of elemental and organic carbon using data from the St. Louis – Midwest Supersite. While the topic is important, I am concerned that the physical models are not appropriate for assessing emissions of EC and OC. With respect to both gases, it would seem vital to explicitly model their chemistry (formation, sinks). With respect to EC, which is not formed in the plume, has the sink been included (deposition)? With regard to OC, both formation in the plume and the sink would need to be modeled. In its current form, the manuscript does not appear to include these important processes. Therefore, the derived emissions will surely be biased. The study needs to be revised to model these processes.

General comments:
1. As mentioned above, the chemistry and deposition needs to be included to accurately derive model sensitivities. For EC, this might be simpler if there is no chemical formation/destruction in the plume and only deposition would have to be modeled in CAMx/FLEXPART. In FLEXPART, it is important to include sink processes for shortlived species and it was not mentioned in the manuscript whether this was done. OC will require a full chemistry model and deposition, without which, emissions or impacts at the site cannot be assessed. If this is not possible, then OC should be removed from the analysis. The current manuscript assumes that the OC measurement at the site is indicative of emissions/impacts from the source. The plume will have a different distribution from the inventory distribution so it is not clear how that can be disentangled without having a chemistry model.

We apologize for failing to mention that the CAMx simulations used both wet and dry deposition. The following text was added:
"Dry deposition was calculated using the Zhang et al., 2003 scheme, and wet deposition using the standard scheme in CAMx."

CAMx simulations are used for estimates of known emissions, whereas FLEXPART is used to estimate

concentration impacts from unknown sources. For FLEXPART we therefore stick to Residence Time Analysis grids which do not have deposition included. The role of the two models was clarified as follows:
"The purpose of combining Eulerian with Lagrangian simulations is therefore to estimate adjustments to known emission inventories with the Eulerian simulations, and to estimate impacts from unknown area sources in an overlapping domain with the Lagrangian simulations."

Because EC is not formed in the atmosphere, our model includes the main processes involved in EC transport, and so the emissions estimates can be interpreted directly. For OC, you are right that chemistry is an important source. However, we feel that there is still valuable information in the OC analysis and that this merits inclusion in the present paper. Aerosols are a very complex subject, and we believe that having different results from different angles does not detract from alternative methodologies but rather contributes to the field as a whole.

The following text at the end of the introduction clarifies this situation:
"Our model is focused on transport and consequently the results for EC can be straightforwardly compared to emission inventories. For OC however, the model does not distinguish between primary OC that is emitted by a source and secondary OC that is created in the plume of that same source. The results are therefore best interpreted in terms of impacts at the measurement site rather than emissions at the source location."

2. I do not understand the purpose of using the CAMx model when FLEXPART could be used for the entire inversion (provided that the chemistry can be included) or vice versa. If the only sources were ones that exist within the domain (i.e. boundary conditions are negligible), FLEXPART would contain all of the necessary information for the inversion. What is the benefit of using the second model?
This was clarified in response to the comment above and the comments from reviewer 2. We use CAMx to simulate transport from the well-established LADCO inventory. We then use FLEXPART to estimate impacts from sources that may have been totally missed in the inventory. Clearly there are alternative choices that are also valuable.

3. If the source distributions are incorrect, then this would affect both the inventory scalings that are derived as well as the estimation of ‘missing’ sources. How well are the spatial distributions known for each source?
We believe that the LADCO inventory is the state-of-the-art for our region. However anyone working in emissions knows what a hugely complex task it is. We think that the results of our analysis provide a partial answer to your question: the Point Sources, "Other," MAR and Non-Road emissions seem to be adequately represented. The category needing the most work according to our inverse model is the On-Road category, bearing in mind that part of the problem could be related to inaccuracies in winter-time WRF winds. Please refer to Sec. 3.3 for a discussion of these issues.
Additionally, the uncertainties in the spatial distributions of the sources is the main reason why we believe that it is valuable to combine CAMx for the known sources and FLEXPART for the unknown sources.

4. Please provide a more in-depth description of the inverse method and the assumptions that go into this method and what they imply (diagonal errors, trust-region iterative algorithm). As one example, assuming independent hourly observations (though there is a mention that previous studies have diagnosed a correlation timescale of 12 hours) could lead to an over-weighting of the data. References are given for various aspects

of the method but the method should be justified in the context of this work.
We have added a figure of the WRF density functions and auto-correlation coefficients in Sec 3.1. This shows more clearly what we are referring to, and also addresses Specific Comment #11.
Note that the measurement errors can be assumed to be uncorrelated in time, as is done by all the studies we know of. For block-bootstrapping, we want to select separate episodes. The issue is not one of correlation of errors, but of length of meteorological events. By using 24 hours as our block-bootstrapping interval, we select independent weather events. The discussion of the uncertainties in the paper has been expanded, please see comment #6 below.

5. Why was 1 ug/m3 uncertainty on the measurements chosen? Has a model representation error been included?
This value was selected by expert judgment as a realistic estimate. Note that is only used in order to interpret the values of the regularization parameter. The model representation errors are included in the regularization parameters. Note that what matters in the inversion is the ratio of the model uncertainties to the measurement uncertainties. In our work, we determine these objectively in order to minimize the total error as explained in the text.

6. It would also be nice to see an outline of the inverse procedure for clarity (for example, is the vector of regularization parameter optimized in the same iterative routine as the emissions)?
The following was added at the end of Sec. 2.4:
"In outline, we first perform the optimization of the regularization parameters without bootstrapping for each set in turn: for the RTA grids, for the LADCO emissions, for the open burning emissions and for the biogenics. This is repeated to make sure the values are stable. We then use the set of regularization parameters to obtain inverse results with the full data set, and 100 realizations with block-bootstrapping."

How are uncertainties and correlations derived in the inversion accounted for? Have the authors analyzed the correlations in the bootstrapping results (e.g. from the realizations of 'x' that are derived)? Are uncertainties in the observations and uncertainties due to the prior (from the regularization parameter) propagated into emissions and associated uncertainties? Some of this material could go into the Supplemental section.
Thank you for bringing up this important point.
In addition to block-bootstrapping, we have performed a Monte Carlo error propagation to further understand the uncertainties and the cross-correlations in the model. A new figure was added to show the uncertainty in the results and the cross-correlations for EC using bootstrapping. The corresponding figures for the Monte Carlo error propagation in the EC inversion and for OC are included in a supplemental section. This shows that the results are not unduly correlated with each other, and that we have made a reasonable attempt at characterizing the errors.

New text at the end of Sec. 2.4:
"We estimate uncertainties in the inverse model by two different methods. The first is to use expert judgment to determine an uncertainty on the measurements (y) and on the model sensitivities H) and to use Monte Carlo error propagation. We perform 100 realizations of the inversion with randomized scaling of the entries in y and H in order to estimate the uncertainties in x. In practice, we assume that entries in y vary by plus or minus 20% and those in H by plus or minus 50%.

An alternative method is to assume that by randomly sampling the data included in the inversion we are randomly sampling both the measurement errors and the simulation errors at the same time.

This can be done with the bootstrap algorithm. Although measurement errors are assumed to be uncorrelated in time, meteorological events vary on the order of hours to days. In order to obtain samples that have different meteorological conditions, we perform block-bootstrapping with a block length of 24 h. We therefore perform 100 inversions with random selection with replacement of the days included in the analysis. In this way, the bootstrapping yields an estimate of the combined uncertainty due to measurement errors and due to transport modeling errors."

New text added at the end of Sec. 3.2:
"We used both Monte Carlo error propagation and bootstrapping to estimate the uncertainties in the emissions estimates. Fig. 9 shows the histogram of total emissions for each of the main categories in the inversion, along with correlation scattergrams of the results for the bootstrapped simulations for EC. For EC, the standard deviation of the contributions is between 3% and 5% of the mean contribution for all emission categories except for open burning where it is 20%. There is little correlation in the emissions estimates from the different source groups. The highest r2 is 0.22 for realizations of the On-Road and Other emissions. Overall this suggests that our results are not excessively impacted by cross-correlation terms.

The results of the Monte Carlo error propagation are included in the supplementary material. The uncertainties vary between 1.5% and 3% except for open burning where they are 6%. These are noticeably lower than the bootstrapping estimates as well as what we expect from knowing about emission inventories and from the values of the regularization parameters that were determined from the inversion themselves. These suggest that using block-bootstrapping provides a better estimate of the uncertainties.

The results for OC are included in the supplementary material. The bootstrapped standard deviations are between 5% and 10% of the mean contributions for all emission categories except for open burning where they are 18%. This suggests that the emissions estimates are robust with respect to uncertainties in the model inputs."

Specific comments:
1. Abstract – This sentence (The inverse model combines forward Eulerian simulations with backward Lagrangian simulations to yield estimates of emissions from sources in current inventories as well as from area emissions that might be missing in the inventories.) is confusing if you haven't first read the paper. Perhaps reword 'area' emissions to something like emissions unaccounted for in the inventories.
The word "area" was removed, which leaves the following text which is similar to the one you suggest:
"from emissions that might be missing in the inventories"

2. Page 12029 Paragraph 1 – it would helpful to have a short description of what Concentration Field Analysis is and what it shows (as was done for the Residence Time Analysis)
New text added:
"Concentration Field Analysis is based on scaling the Residence Time Analysis at each time step with the concentration at the measurement site. The sum over the entire measurement period is then normalized with the Residence Time Analysis. This highlights air flow patterns that are associated with high receptor concentrations."

3. Page 12029 Line 24 – An explanation for why the two models are used together would be helpful. At present, it is unclear what the need is for using both (i.e. couldn't

FLEXPART be used alone?).
Please see the clarifications added to the text, outlined under General Comments #1 and #3.

4. Section 2.4 Least Squares Inverse Model – Please provide a short description of the lifetimes of these species and whether it is assumed that the boundary conditions to the Lagrangian domain are negligible.
Text added: "In our case, the background levels of EC and OC are very low (see Fig. 4), and we expect minimal impacts from sources outside the study area."

5. Section 2.4 Least Squares Inverse Model - How are CAMx model sensitivities calculated (emissions from the inventory of that particular source/time period are perturbed)?
We have clarified the explanation (see also comments from reviewer #2):
"Hourly Eulerian simulations with CAMx were performed for the five different source groups in the LADCO inventory: On-Road, Non-Road, MAR, Other and Point Sources. Because we are interested in evaluating the temporal profiles of the sources, we carry out separate simulations for emissions during different times of the day and different days of the week. The time slots were selected based on the diurnal profile used in the emissions inventory: 11:00 p.m. to 05:00 a.m., 05:00 a.m. to 08:00 a.m., 08:00 a.m. to 02:00 p.m., 02:00 p.m. to 06:00 p.m., and 06:00 p.m. to 11:00 p.m. Days of the week were split into a weekday group and a group containing Saturdays, Sundays and Holidays. As an example, an hourly time series of concentrations was obtained from a CAMx simulation with On-Road emissions only between 05:00 a.m. to 08:00 a.m. on weekdays."

And further down:
"For the CAMx time series, the entries in x are scaling factors on the LADCO emissions that went into the CAMx simulations."

6. Section 2.4 Least Squares Inverse Model – There could be significant temporal correlation. There is a brief mention of 12 hours being the correlation timescale from previous studies, but hourly observations are used and are treated as independent. This could lead to over-weighting of observations in the inversion. Can the least squares method be reformulated to deal with a full covariance matrix? Otherwise, using daily averaged observations may be better.
Yes, we could use a full covariance matrix with this method. In practice, most inverse models of emissions in the atmospheric sciences use diagonal matrices and so we are following common approaches to this question. The brief mention of 12 hours is to do with the correlation time of meteorological events. We use this to justify the selection of block-bootstrapping on chunks of 24 hours in order to increase the variability of the meteorological conditions in our bootstrapped sample. We moved the mention of the auto-correlation to Sec 3.1 as it is misleading in this context. Please refer to the new text described for Comment #6 above.

7. Section 2.4 Least Squares Inverse Model – Are there assumptions that go into converting equation 1 to equation 2? Can you describe what an augmented H", x" and y" are (what are the dimensions)? Describing the inverse methodology in more detail is needed and can go in the Supplement.
There are no assumptions, please refer to Aster et al., 2012 for more details along with our previous papers on the method.
Text added: "H'' has dimensions of (7091 + 3486) by (3486), and y'' has dimensions of (7091 + 3486)."

8. Page 12030 Line 29 – ‘Area sources’ is confusing. Sources unaccounted for in the inventories is more clear.
The paragraph was rewritten as follows (see also comments by reviewer #2):
"The inverse model derives a posterior estimate of emissions based on the Eulerian simulations that used the emissions inventory as a prior. In addition, the inverse model uses the Lagrangian simulations to derive an estimate of sources that may be missing from the inventory. This is done by using the polar grids of Residence Time Analysis that represent the impact that an emission in a given grid cell would have at the measurement site. As all the known sources were already included in the CAMx simulations with the emissions inventory, we use a field of zero prior emissions for the polar grids from the Lagrangian simulations."

9. Page 12031 Line 25 - What are the 606 emissions elements? Are they scaling factors of the prior distribution for that source/time? Please provide some text to clarify this.
This was clarified as follows:
"For the CAMx time series, the entries in x are scaling factors on the LADCO emissions that went into the CAMx simulations."

10. Page 12033 First Paragraph – Why are the results of the inversion for the regularization parameter described here rather in the Results section? Also why are single values given? Isn’t ‘s’ a vector of values? It would also be good to discuss these results more, for example, about which components of the inventories are most uncertain. The derived regularization parameter should give an indication of the relative uncertainties of various parts of the prior.
We appreciate the suggestion of moving some of this section to the results section which would be a logical place to find it. However, we felt that when we did this it broke up the flow of discussing the emissions, and it separated into two parts something that is best understood when it is kept in a single part. We would therefore prefer to keep this section as it is.
In principle, one can have as many values of s as there are entries in x. In practice this is neither feasible nor desirable. We have therefore elected to use common values by emission groups. The logic is that the estimate of the uncertainties for example of the 6 open burning parameters are similar to each other, but different than the ones for the LADCO emissions.

The text was modified as follows:
"While in principle we can ascribe different values for each entry in the sensitivity matrix, we decided to use common values by source groups. The values of s were therefore determined separately for the emissions inventory sources, for the open burning sources and for the emissions based on back-trajectories."

11. Page 12033 Line 18 – The claim that there are no systematic errors in the model is likely overstated.
We have included the figure for KCPS in Sec. 3.1 and added the following text:
"Fig. 6 shows the probability density function for both the measurements and the simulations at KCPS. The distributions are very similar, and all variables passed the Kolmogorov-Smirnov test to much lower than the 1% significance level, showing that the model does not suffer from significant systematic biases."

The original sentence was entirely removed as part of an improved discussion of model uncertainties, see comments #6 above.

12. Page 12038 Line 19 – ‘explains why’ should be ‘low posterior emission causes the total emissions to decrease’
We have replaced "explains" with "is": "which is why ..."

13. Page 12040 Line 19 – why is the inversion not able to simulate winter concentrations? Are there ‘missing’ sources at this time that are compensating for the lack of agreement with the inventories (if posterior is showing scaling from inventories are showing near 0 emissions)?
This is an area of future research - any statement we would make would be speculative at this point.

14. Page 12041 Line 1 – Are these swings statistically significant based on the derived uncertainties? The phrase ‘This suggests that there are large uncertainties in these estimates’ should be rephrased using the evidence from the uncertainties that are presented.
The following text was added: "These swings are mostly contained within the 90% confidence range displayed in the figure which suggests that they are not statistically significant."

Also, the conclusion that ‘more data could stabilize the emissions is too narrow. There are other areas that could contribute such as in the spatial distribution of the inventories and lack of chemistry being modeled that are hard-wired into the system.
We have added the following text: "or an improved model that considered in-plume chemistry."