

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.

Thank you for the reviews of our manuscript. Please find below a point by point reply to the comments along with a list of the changes made to the text.

The complete original reviewer comments 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. Their 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 modeling 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 squares 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 do 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 αI , where α 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 α . 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 α times an identity matrix. The text was adjusted as follows:

"In practice, α 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)."