**Atmospheric
Chemistry
and Physics**
Open Access
Discussions

Interactive comment on "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 et al.

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 short-lived 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.

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?

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?

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.

5. Why was 1 ug/m3 uncertainty on the measurements chosen? Has a model representation error been included?

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)?

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.

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.

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)

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?).

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.

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)?

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 pervious 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.

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.

8. Page 12030 Line 29 – 'Area sources' is confusing. Sources unaccounted for in the inventories is more clear.

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.

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.

11. Page 12033 Line 18 – The claim that there are no systematic errors in the model is likely overstated.

12. Page 12038 Line 19 – 'explains why' should be 'low posterior emission causes the total emissions to decrease'

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)?

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. 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.

———————————————