**Atmospheric
Chemistry
and Physics
Discussions**

Interactive comment on "Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression" *by* H. Simon et al.

Anonymous Referee #3

Received and published: 16 December 2010

In this paper, a modified multiple linear regression approach is applied to data from the IMPROVE network for estimation of OM/OC ratios across the United States. This study is a continuation of a previous data analysis using ordinary least squares (OLS) regression. OLS regression does not account for errors in the explanatory variables. To overcome this limitation so called error-in-variance models are applied in the present paper. The authors know regression modelling very well and the data analysis is done very carefully. However, there are some open questions that need to be addressed prior to publication in ACP (see below).

Major comments:

1. There is no instrumental section. The authors give some information about measurement techniques used within IMPROVE in section 1, but this is incomplete (e.g. no information about method for determination of OC and EC). I would appreciate a very short section with the basic information. I would also appreciate a table summarizing what constituents were determined from which filter material.

2. I have a major concern about interpretation of the absolute values of the determined OM/OC ratios. My concern comes from the fact that thermal-optical determination of OC and EC as applied in this study must be regarded as a conventional approach for determination of OC rather than a correct determination of a chemically well defined PM fraction. The split between OC and EC strongly depends on the detailed realization of the measurement. The chosen temperature protocol and the method for charring correction (reflectance or transmission) have a rather large impact on the obtained value. On the other hand, the absolute values for OC have a direct impact on the estimate for coefficient b_OC when used in a regression model (the PM mass concentration explained by OC remains constant). As a consequence, the OM/OC ratio would be different from the reported values when e.g. a NIOSH temperature protocol instead of the IMPROVE temperature protocol would have been used. The authors should very clearly address this limitation.

The authors might argue that best solutions were found for b_EC to be equal to 1 (discussion in the supplementary material), however, other values seem also possible and the "uncertainty" in fixing b_EC to 1 must be considered as being rather large.

In my view, the regional and seasonal differences in b_OC are interesting and merit publication. However, the authors should avoid the interpretation of the absolute values. On p. 24658 the readers are cautioned against over-interpreting the regression coefficients which is good but not sufficient.

3. The authors argue that volatilization of nitrate leads to b_nit < 1 (p. 24660 line18, and section 3.2). This is plausible since losses of nitrate due to volatilization should be

different for nylon and Teflon filters. If volatilization of nitrate is important, the assumed linear effect of measured nitrate on measured PM2.5 is questionable. The authors should address this issue.

4. In Eq. 5, KNON represents the mass of wood burning related potassium. The coefficient of KNON is fixed at 1.2 in order to account for potassium oxide from wood burning. However, there should be more mass related to emissions from wood burning than potassium oxide (and OM and EC that is represented in Eq. 5 by the corresponding terms). So it is not so obvious that the model as given in Eq. 5 is (a) appropriate (are all effects linear?) and (b) complete. Why is a model without intercept used? An obvious option would be to use a model with intercept, if no intercept is needed the confidence interval for the estimated intercept would then include zero. There are some questions related to the model selection, I therefore would very much suggest to include a careful analysis of the residuals for justification of the selected model (in the supplementary material, justification of selected model in main text).

5. The discussion of the estimated coefficients in section 3 is lengthy and not easy to read. The manuscript would gain if this section could be reduced to the most important facts. Please try to revise accordingly.

Minor comments

Abstract, line 6: I would prefer the expression "variable selection" instead of "dataset selection".

Page 24654, line 11: Please mention here on what type of filter PM2.5 was collected.

Page 24674, line 12: It is stated that b_nit cannot be precisely estimated in summer – what about the stand-ard errors (or confidence intervals) for b_nit?