Response to reviewer comments supplement for:
Determining the spatial and seasonal variability in OM/OC ratios across the U.S. using multiple regression

Heather A. Simon[1], Prakash V. Bhave[1], Jenise L. Swall[1], Neil H. Frank[2], William C. Malm[3]

[1] US EPA, National Exposure Research Laboratory, Atmospheric Modeling and Analysis Division, Research Triangle Park, NC
[2] US EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC
[3]National Park Service, Colorado State University/Cooperative Institute for Research in the Atmosphere, Fort Collins, CO

Exploring alternate regression model set-ups to represent nitrate volatilization

The work done by Herring and Cass (1999) and Frank (2006) both show that absolute nitrate loss from Teflon filters is dependent on ambient temperature and RH but not on nitrate concentration (as long as the calculated nitrate loss is not greater than the ambient nitrate concentration in which case 100% of the nitrate would be lost). Therefore the regression equation could be rewritten so that nitrate volatilization is represented by an intercept instead of a multiplier. In that case the base regression equation (Eq. 1) would be replaced by Eq. 2, where the intercept term, c_{nit}, is expected to be negative and to represent the total nitrate loss from the Teflon filter converted to ambient concentrations (µg/m^3).

$$\begin{aligned} PM_{2.5,i} = {} & \beta_{OC}OC_i + \beta_{sulf}(NH_4)_2SO_{4,i} + \beta_{nit}NH_4NO_{3,i} + \beta_{soil}SOIL_i \\ & + EC_i + 1.8\times Cl_i^- + 1.2\times KNON_i + \varepsilon_i \end{aligned} \tag{1}$$

$$\begin{aligned} PM_{2.5,i} = {} & \beta_{OC}OC_i + \beta_{sulf}(NH_4)_2SO_{4,i} + (NH_4NO_{3,i} + c_{nit}) \\ & + \beta_{soil}SOIL_i + EC_i + 1.8\times Cl_i^- + 1.2\times KNON_i + \varepsilon_i \end{aligned} \tag{2}$$

However, in the case where nitrate loss is captured in the mass balance equation by an intercept, a multiplier could still be used to take into account extra mass from hydration and heavier cations such as Na. Therefore, the regression equation could be rewritten as shown in Eq. 3. For Eq. 3 to make physical sense, β_{nit} should always be greater than 1 since it represents added mass from hydration and heavy cations while c_{nit} should always be less than 0 since it represents loss of nitrate from the filter. For instance if the regression estimated β_{nit} = 1.1 and c_{nit} = -1, these values could be interpreted to mean that there is 10% added mass from hydration and 1 ug/m3 of nitrate lost due to volatilization.

$$\begin{aligned} PM_{2.5,i} &= \beta_{OC} OC_i + \beta_{sulf} (NH_4)_2 SO_{4,i} + (\beta_{nit} NH_4 NO_{3,i} + c_{nit}) \\ &\quad + \beta_{soil} SOIL_i + EC_i + 1.8 \times Cl_i^- + 1.2 \times KNON_i + \varepsilon_i \end{aligned} \tag{3}$$

Consequently we have performed two sensitivity analyses: one in which we use an Errors in Variables regressions with Eq. 2 and one in which we use an Errors in Variables regressions with Eq. 3. Figures 1, 2, and 3 show distributions of β_{nit} and c_{nit} by season for the base regression presented in the main paper and the two sensitivity analyses. As noted in the main paper, the base regression predicts that many IMRPOVE sites will have 100% nitrate volatilization during quarter 3 and greater than 50% nitrate volatilization in quarters 2 and 4. The regressions using Eq. 2 show little variation in predicted mean nitrate volatilization (around -0.1 to -0.2 $\mu g/m^3$). However, the 95th percentile nitrate volatilization was much greater in quarter 3 (1.2 $\mu g/m^3$) than in quarter 1 (0.7 $\mu g/m^3$). According to Figure 2 in Frank (2006), nitrate volatilization would be in the range of 0.5-1.75 $\mu g/m^3$ at typical wintertime temperatures and between 2.5-10 $\mu g/m^3$ at typical summertime temperatures. However, median quarter 3 nitrate concentrations are between 0.07-0.64 $\mu g/m^3$ at 90% of IMPROVE sites, meaning that 100% of the nitrate would volatilize at these sites. Therefore, median c_{nit} estimates shown in Figure 2 seem too close to zero.

The results from the regressions using Eq. 3 are hard to interpret as physically meaningful. Instead of β_{nit} values greater than 1 and seasonally varying nitrate volatilization (c_{nit}), we see seasonally varying β_{nit} values less than 1 and c_{nit} values with little seasonal variation. It appears that β_{nit} and c_{nit} are not truly independent, so that in this sensitivity analysis β_{nit} is capturing some of the nitrate loss that should be represented by c_{nit}.

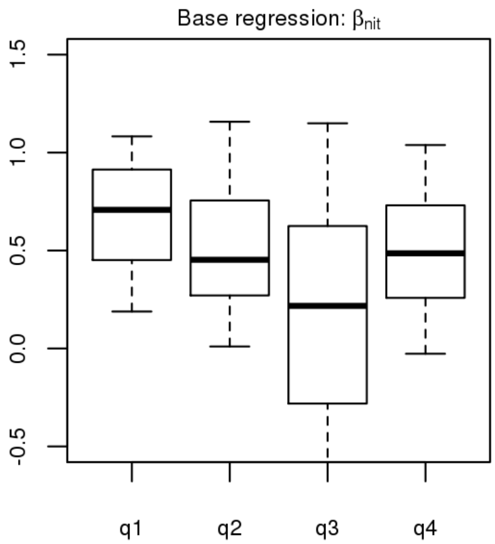

Figure 1: β_{nit} coefficients in the base regression analysis.

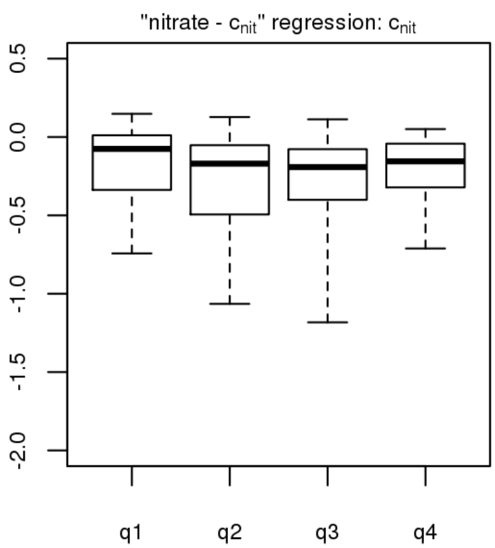

Figure 2: c_{nit} in regression analysis using Eq 2.

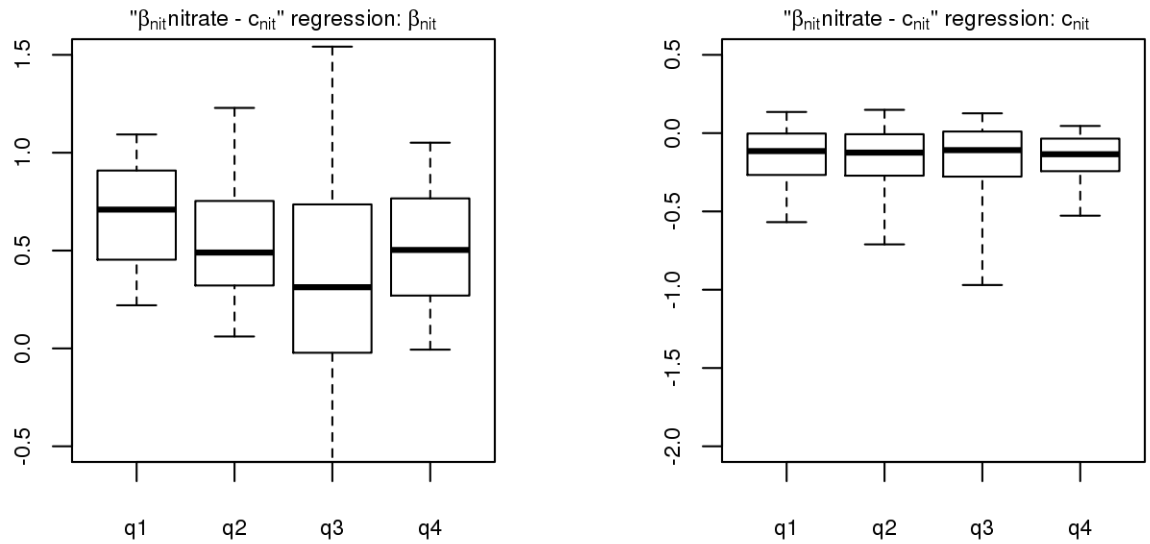

Figure 3: β_{nit} (left) and c_{nit} (right) in the regression analysis using Eq. 3.

In addition, in each of the sensitivity analyses there are some problematic estimates of c_{nit}. Nineteen percent of the regressions using Eq. 2 have c_{nit} values greater than 0, meaning that there is more mass on the Teflon filter than estimated from the various chemical constituents multiplied by their coefficients. Therefore, these c_{nit} values cannot be interpreted as implying anything about the amount of nitrate volatilization. In the analysis using Eq. 3, 23% of regressions have c_{nit} estimates greater than 0. There is one final complicating factor which makes these sensitivity analyses difficult to interpret. The regression estimates one “average” c_{nit} value, but this c_{nit} estimate can clearly not be applied to every sample within the dataset since some samples have lower nitrate concentrations than the “average” volatilization. For the sensitivity analysis using Eq. 2, 16% of samples typically have nitrate concentrations less than the estimated nitrate volatilization in quarter 1 and 72% of samples typically have nitrate concentrations less than the estimated nitrate volatilization in quarter 3. A similar problem can be observed in regressions using Eq. 3. There is no easy way to account for this problem (i.e. using max(c_{nit}, nitrate) is not easily implementable in a regression analysis). Therefore, we believe that the model formulations represented by Eq. 2 and Eq 3 do not offer a clear advantage over our base model formulation given in the main paper.

Nevertheless, we have evaluated how the estimates of β_{OC} differ when using Eq. 2 instead of Eq 1. When comparing results by region in quarter 1, the regression using Eq. 2 leads to more site-to-site variability in β_{OC} estimates within each region and more unrealistically low β_{OC} estimates in the west and great lakes regions. Median predicted β_{OC} estimates in each region are fairly similar between the two regressions except that the median β_{OC} estimate is slightly lower in the central region when using the base regression (1.3 vs 1.5). Both regression methods predict little variability among regions in quarter 3 with median β_{OC} estimates between 1.8 and 1.9.

Frank, N. H.: Retained nitrate, hydrated sulfates, and carbonaceous mass in Federal Reference Method fine particulate matter for six eastern US cities, J. Air Waste Manage. Assoc., 56, 500-511, 2006.

Hering, S., and Cass, G.: The magnitude of bias in the measurement of PM2.5 arising from volatilization of particulate nitrate from teflon filters, J. Air Waste Manage. Assoc., 49, 725-733, 1999.