Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression

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

[1] US EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC

[2] US EPA, National Exposure Research Laboratory, Atmospheric Modeling and Analysis Division, Research Triangle Park, NC

[3] National Park Service, Colorado State University/Cooperative Institute for Research in the Atmosphere, Fort Collins, CO

Correspondence to: H. Simon (Simon.Heather@epa.gov)

Abstract

Data from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network are used to estimate organic mass to organic carbon (OM/OC) ratios across the United States by extending previously published multiple regression techniques. Our new methodology addresses common pitfalls of multiple regression including measurement uncertainty, colinearity of covariates, dataset selection, and model selection. As expected, summertime OM/OC ratios are larger than wintertime values across the U.S with all regional median OM/OC values tightly confined between 1.80 and 1.95. Further, we find that OM/OC ratios during the winter are distinctly larger in the eastern US than in the West (regional medians are 1.58, 1.64, and 1.85 in the great lakes, southeast, and northeast regions, versus 1.29 and 1.32 in the western and central states). We find less spatial variability in long-term averaged OM/OC ratios across the US (90% of our multiyear regressions estimate OM/OC ratios between 1.37 and 1.94) than previous studies (90% fell between 1.30 and 2.10). We attribute this difference largely to the inclusion of EC as a covariate in previous regression studies. Due to the colinearity of EC and OC, we find that up to one-quarter of the OM/OC estimates in a previous study are biased low. Assumptions about OC measurement artifacts add uncertainty to our estimates of OM/OC. In addition to estimating OM/OC ratios, our technique reveals trends that may be contrasted with conventional assumptions regarding nitrate, sulfate, and soil across the IMPROVE network. For example, our regressions show pronounced seasonal and spatial variability in both nitrate volatilization and sulfate neutralization and hydration.

1 Introduction

Atmospheric measurements have shown that organic mass (OM) is a major component of fine particulate matter ($PM_{2.5}$), comprising over 50% of ambient $PM_{2.5}$ in some locations (Jimenez et al., 2009;Murphy et al., 2006;Zhang et al., 2007). OM can be divided broadly into two components: organic carbon (OC), and all other mass which we will hereafter refer to as non-carbon organic mass (NCOM). NCOM is the largest component of ambient $PM_{2.5}$ that is not routinely measured. To achieve mass closure in source testing and ambient particle measurements, an OM/OC ratio [denoted as k and R_{OC} in some earlier literature (Frank, 2006;Malm and Hand, 2007)] is often multiplied by measured OC to estimate total OM. This ratio is primarily affected by the oxygen content in the organic aerosol (Pang et al., 2006), although hydrogen, nitrogen, and sulfur also make small contributions to the NCOM.

The first estimate of OM/OC was made by White and Roberts (1977), who calculated an average ratio of 1.4 for specific organic compounds measured in Los Angeles. This value was used widely until Turpin and Lim (2001) analyzed a larger dataset to show that OM/OC is generally higher than 1.4. In recent years a range of techniques have been applied to quantify OM/OC, including gas chromatography/mass spectrometry (GC/MS) (Turpin and Lim, 2001;Yu et al., 2005), high resolution time of flight aerosol mass spectrometry (HR-ToF-AMS) (Aiken et al., 2008;Chan et al., 2010;Sun et al., 2009), Fourier Transform Infrared (FTIR) spectroscopy (Gilardoni et al., 2007;Kiss et al., 2002;Liu et al., 2009;Polidori et al., 2008;Reff et al., 2007;Russell, 2003;Russell et al., 2009), sequential extraction followed by gravimetric weighing and thermal optical measurement of carbon (El-Zanan et al., 2005;El-Zanan et al., 2009;Lowenthal et al., 2009;Polidori et al., 2008), and coupled thermal gravimetric and chemical analyses (Chen and Yu, 2007). Those studies have contributed substantially to our understanding of NCOM in many laboratory and field settings, but none of the techniques have been applied over a broad temporal and spatial range.

Numerous $PM_{2.5}$ constituents, including OC but not OM, are measured routinely across two large US networks: the Chemical Speciation Network (CSN) and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network. A technique for computing OM from these networks could yield a comprehensive dataset of OM/OC ratios covering a large spatial and temporal extent. Frank (2006) developed the SANDWICH method to estimate OM from measurements across the urban-centric CSN. He calculated total OM as $PM_{2.5}$ minus the sum of other components (sulfate, nitrate, ammonium, crustal material, and elemental carbon (EC)), while making adjustments for particle-bound water (not measured directly) and nitrate volatilization. Unfortunately, the uncertainty in OC data collected at CSN sites prior to some major network changes in

2008 is comparable to the uncertainty in OM/OC ratios (Watson, 2008). Therefore, although the SANDWICH technique is useful for estimating total OM, CSN data are not yet adequate for estimating OM/OC over large multiyear periods.

The IMPROVE network tracks visibility degradation in national parks and wilderness areas via routine measurements of $PM_{2.5}$ mass and composition (Malm et al., 1994). The network began with 36 monitoring sites in 1988, and currently reports data from 178 remote and 13 urban sites across the continental US, Hawaii, Alaska and the Virgin Islands (http://vista.cira.colostate.edu/improve/Data/IMPROVE/AsciiData.aspx). $PM_{2.5}$ is collected on filters for a 24-hour period (midnight to midnight) every third day. The filters are subjected to a gravimetric analysis that measures total mass and various chemical analyses that measure bulk composition. Specifically, OC and EC are measured by the Thermal Optical Reflectance (TOR) combustion method; SO_4^{2-}, NO_3^-, and Cl^- by ion chromotograpahy; and elements with atomic weights between sodium and lead by X-Ray Fluorescence (XRF). Table 1 summarizes the IMPROVE measurements used for this paper and the filter medium on which each particle component is collected. In addition to these direct measurements, the network reports a reconstructed fine mass (RCFM) concentration which is a weighted sum of selected chemical constituents. RCFM was first calculated using Eqs. (1) and (2) (Malm et al., 1994), though our notation differs slightly from the original publication.

$$RCFM = (NH_4)_2SO_4 + SOIL + EC + OM \quad (1)$$

$$SOIL = 2.20\,Al + 2.49\,Si + 1.63\,Ca + 2.42\,Fe + 1.94\,Ti \quad (2)$$

Ammonium sulfate ($(NH_4)_2SO_4$) was calculated as $4.125 \times S$ (sulfur was measured by Particle Induced X-ray Emission [PIXE] until 2002 and by XRF since then), SOIL was calculated with Eq. (2) (assuming the soil in $PM_{2.5}$ samples mimics the average composition of sedimentary rock), and OM was calculated as 1.4xOC. Changes to the RCFM equation since 1994 include the addition of more components (ammonium nitrate

(NH_4NO_3), non-soil potassium, and sea salt), modification of Eq. (2) to eliminate Al, and an increase of OM/OC from 1.4 to 1.8 (McDade, 2008).

Although a network-wide OM/OC ratio is commonly used to compute RCFM, a few studies have estimated site-specific OM/OC ratios from IMPROVE data. El-Zanan et al. (2005) describe a mass closure technique for calculating OM/OC,

$$\frac{OM}{OC} = \frac{PM_{2.5} - ((NH_4)_2SO_4 + NH_4NO_3 + EC + SOIL + Other)}{OC} \quad (3)$$

in which "Other" is the sum of sodium, chlorine, and trace elements measured by XRF that are not associated with soil (Lowenthal and Kumar, 2003). Unfortunately, there are many uncertainties associated with a mass closure analysis of IMPROVE data. First, assumptions must be made about two unmeasured $PM_{2.5}$ components: ammonium and particle-bound water. Since ammonium is not routinely measured at IMPROVE sites, sulfate and nitrate are commonly assumed to be fully neutralized by ammonium. Estimation of water mass is complicated by the fact that filter samples are shipped at ambient conditions and weighed in a laboratory where relative humidity (RH) is not controlled. Second, nitrate measurements are made from particles collected on nylon filters downstream of a HNO_3 denuder, to which nitrate adheres well, whereas $PM_{2.5}$ weights are determined from Teflon filters, from which nitrate is known to volatilize (Hering and Cass, 1999). The amount of volatilization from the Teflon filter depends on which cation the nitrate is bound to as well as the temperature and RH during sampling, shipping, and analysis. Third, the IMPROVE soil equation relies on assumptions about the abundance and oxidation states of various trace elements. Since soil composition is spatially heterogeneous, this equation may not accurately estimate the soil contribution at all sites. Finally, OC measurement artifacts contribute additional uncertainty because OC is measured from quartz filters while OM is derived from gravimetric measurements on Teflon filters. Differing tendencies among these two filter materials at retaining OM and/or adsorbing semi-volatile organic gases may affect OM/OC estimates.

To overcome some shortcomings of the mass-closure approach, Malm and collaborators developed a multiple regression technique to estimate OM/OC from 1988 – 2003 IMPROVE data (Hand and Malm, 2006;Malm et al., 2005;Malm and Hand, 2007). They fit seven coefficients in Eq. (4) using ordinary least squares (OLS) regression at each monitoring site. Some notation in Eq. (4) has been changed from that of Malm and Hand (2007) for consistency with the present study.

$$\begin{aligned} PM_{2.5,i} &= \beta_0 + \beta_{OC}\, OC_i + \beta_{sulf}\, (NH_4)_2SO_{4,i} + \beta_{nit}\, NH_4NO_{3,i} + \beta_{soil}\, SOIL_i \\ &\quad + \beta_{EC}\, EC_i + \beta_{seasalt} \times 1.8\, Cl_i^- + \varepsilon_i \end{aligned} \qquad (4)$$

The subscript, *i*, represents a day-specific sample and β_0 represents a site-specific intercept. The remaining β coefficients represent ratios of the mass associated with a given $PM_{2.5}$ component on the Teflon filter when it was weighed to the mass of that same component determined (or estimated) via chemical analysis of a (possibly) separate filter. The residual error (ε_i) denotes the difference between the measured $PM_{2.5}$ mass and the estimated mass (based on fitted coefficients and measured chemical components) for a particular sample. The coefficient of most interest to us is β_{OC} because it represents OM/OC. This technique circumvents many of the assumptions needed for mass closure. For example, β_{OC} is insensitive to the degree of sulfate neutralization since the relative abundance of ammonium would mainly affect β_{sulf}. However, OC measurement artifacts can certainly introduce bias in β_{OC}.

In this paper we develop a nationwide dataset of seasonally- and spatially-varying OM/OC ratios across the IMPROVE network by extending the methodology of Malm and Hand (2007) while addressing some common pitfalls in multiple regression. We discuss new quantitative insights regarding the measurement artifacts associated with $PM_{2.5}$ components other than OC (e.g. nitrate volatilization and water associated with particulate sulfate), which are ancillary benefits of our methodology. Finally, spatial and temporal trends in OM/OC are reported and examined.

2 Methodology

Figure 1 shows a schematic of our methodology, with complete details provided in this section.

2.1 General equation and dataset selection

We begin by making three minor modifications to Eq. (4). First, we eliminate the intercept term (β_0) and reduce the number of explanatory variables (i.e., covariates) to four that constitute the majority of $PM_{2.5}$ and have large uncertainty in their coefficient: OC, $(NH_4)_2SO_4$, NH_4NO_3, and SOIL [Eq. (5)].

$$PM_{2.5,i} = \beta_{OC}OC_i + \beta_{sulf}(NH_4)_2SO_{4,i} + \beta_{nit}NH_4NO_{3,i} + \beta_{soil}SOIL_i + EC_i + 1.8Cl_i^- + 1.2KNON_i + \varepsilon_i \quad (5)$$

$$KNON = K - 0.6Fe \quad (6)$$

$$SOIL = 3.48Si + 1.63Ca + 2.42Fe + 1.94Ti \quad (7)$$

In contrast to Eq. (4), we assume that EC has no artifact and set its coefficient to 1 because treating EC as a separate explanatory variable can bias β_{OC} (see Sect. 3.3 and Supplement Sect. S3). Similar to Eq. (4), we estimate sea salt as 1.8 Cl^- (Pitchford et al., 2007;White, 2008) but do not treat it as an explanatory variable. Although 1.8 Cl^- has been deemed a good estimate of sea salt mass at coastal IMPROVE sites, it may underestimate sea salt concentrations at inland locations where Cl^- has been displaced from the aged sea salt. However, this underestimation should not substantially affect the regression results because sea salt contributes little to $PM_{2.5}$ mass at most inland locations. Second, we add KNON to Eq. (5) for consistency with the newest IMPROVE RCFM formula (McDade, 2008). KNON represents non-soil potassium (e.g., from wood burning) and is calculated using Eq. (6). The KNON coefficient is fixed at 1.2, the molar mass ratio of potassium oxide to potassium. Although KNON is influenced by soil composition (i.e., soil K/Fe ratio may deviate from 0.6), it contributes a small enough mass to total $PM_{2.5}$ that fixing its coefficient should not adversely affect the regression as a whole. Third, we use an updated IMPROVE soil equation (compare Eqs. (2) and (7)) which eliminates aluminum from the calculation because Al is not reliably measured by the IMPROVE XRF analysis (McDade, 2008).

We downloaded the IMPROVE data from http://views.cira.colostate.edu/web/DataWizard/ on 6 January 2010, and analyzed the measurements collected at 186 continental US sites between 1 January 2002 and 31 December 2008. All analyses are performed using the R statistical software package (R Development Core Team, 2010). Like Malm and Hand (2007), we segregate the data by monitoring site. In addition, we segregate data by season: quarter 1 (Jan, Feb, Mar), quarter 2 (Apr, May, Jun), quarter 3 (Jul, Aug, Sep), and quarter 4 (Oct, Nov, Dec), because we expect the coefficients (i.e., OM/OC and nitrate volatilization) to vary seasonally. However, we could not justify the seasonal variability in soil coefficients estimated from our initial analyses. For instance, the variability in β_{soil} was not correlated to Asian dust plumes or other seasonally varying dust sources. We therefore hold the soil coefficient constant throughout the year by first performing a multiyear regression at each site using all data from 2002-2008 and then fixing β_{soil} in each quarter-specific regression to the β_{soil} value obtained from the multiyear regression at that given site.

Within site and quarter-specific datasets, the only data filter that we apply is completeness. If a major component in Eq. (5) (i.e., $PM_{2.5}$, OC, S, NO_3^-, Si, Ca, Fe, Ti, or EC) is missing from a single site and sample, we eliminate the whole date from that site. Missing data values for Cl^- and K are set to 0. All concentrations reported as negative values are left as is. Finally, sites that do not have an average of at least 15 days of complete data per quarter (i.e., 105 samples for each quarter over the 7 year measurement period) for all four quarters are eliminated from the analysis. This criterion eliminates thirty-three sites. As shown in Fig. 1, we perform one multiyear and four quarter-specific regressions for each of the remaining 153 monitoring sites (i.e., 765 separate regressions).

2.2 Physical interpretation of coefficients

When interpreting the coefficients in Eq. (5), it is important to note that all results may be affected by changes in measurement techniques and variability in the ambient conditions. Therefore, readers are cautioned against over-interpreting results from a single regression and instead are encouraged to use these results to understand spatial and temporal trends in the coefficients. For each $PM_{2.5}$ component, the regression coefficient represents the ratio of retained mass associated with that component on the Teflon filter (used for gravimetric $PM_{2.5}$ analysis) to the mass of that component derived from chemical analysis. Here we describe how values different than 1 may be interpreted and set bounds on physically reasonable values for each coefficient.

The OC coefficient, β_{OC}, should represent the OM/OC ratio. We expect its lower bound to equal 1, representing pure graphitic carbon with no associated hydrogen, oxygen, or nitrogen mass. We expect the upper bound to equal 3.8, which is at the upper end of OM/OC ratios for aliphatic dicarbonyls (Turpin and Lim, 2001). It is possible to have a higher OM/OC for some organic sulfates, but it is unlikely that these compounds would contribute enough mass to raise the overall OM/OC above 3.8. Typical OM/OC ratios for primary organic emissions are around 1.25 in vehicle exhaust and 1.7 in wood smoke emissions (Reff et al., 2009). Measurements of OM/OC from laboratory-generated secondary organic aerosol (SOA) range from 1.4-2.7(Chhabra, 2009;Kleindienst et al., 2007). Ambient measurements of OM/OC have shown a wide range of values for different types of aerosols in different locations. Aiken et al. (2008) report values between 1.4 and 2.5 in Mexico City and the surrounding areas during the spring of 2006. Sun et al. (2009) report values ranging from 1.75 to 2.83 at Whistler Mountain in British Columbia, Canada also in the spring of 2006. Finally, Huang et al. (2010) measured OM/OC between 1.3 and 1.78 in Beijing in 2008. Although we interpret β_{OC} as equivalent to OM/OC, this coefficient may also be skewed by two types of OC measurement artifact: negative artifacts occur when organic PM collected on the filter volatilizes before chemical analysis and positive artifacts occur when organic vapors adsorb to the filter surface (McDow and Huntzicker, 1990;Turpin et al., 1994). β_{OC} will be influenced further by differences in the sampling artifact on quartz filters (used to measure OC) versus Teflon filters. These artifacts are discussed further in Supplement Sect. S3. It should also be noted that OC is operationally defined. Here, OC is measured with the IMPROVE TOR protocol, which is now used at both CSN and IMPROVE network sites. Coefficients reported in this paper should only be applied to OC measurements derived using the same or equivalent methods.

A soil coefficient not equal to 1 could represent soil compositions differing from the average sediment used to develop Eqs. (2) and (7). β_{soil} represents the actual soil mass in the $PM_{2.5}$ sample divided by the soil mass calculated from Eq. (7). Simon et al. (2010) report that this ratio can range from 0.41 to 1.63 based on soil compositions in the literature, so these bounds are used to assess the physical reasonableness of β_{soil}.

A sulfate coefficient, β_{sulf}, below 1 would indicate that the assumption of dry ammonium sulfate over-estimates total sulfate mass in the samples. Incomplete neutralization could cause such an over-estimate. The molar mass of ammonium bisulfate (NH_4HSO_4) and sulfuric acid (H_2SO_4) are 87% and 74% of the $(NH_4)_2SO_4$ molar mass. Therefore, 0.74 would seem like a reasonable lower bound for β_{sulf}. However, the sulfate mass in our regression is calculated from an XRF sulfur measurement which can detect organo-sulfur atoms. A conservative lower bound could be calculated assuming that all sulfur mass associated with organic molecules would be included in the β_{OC}. Surratt et al. (2008) report that up to 20% of sulfur may be contained in these organic compounds, so we expect the lowest reasonable value of β_{sulf} to equal 0.59 (0.74×0.8) to capture an admittedly extreme scenario in which all inorganic sulfate is in the form of sulfuric acid and 20% of the total sulfur is contained in organic compounds. A sulfate coefficient above 1 would indicate that there is extra mass associated with the particulate sulfate. This extra mass could come from water if the aerosol remains hydrated during gravimetric analysis. During the history of the IMPROVE network, RH in the gravimetric measurement laboratory was only recorded intermittently. We obtained laboratory measurements of RH during the gravimetric analysis of filters collected from September 2003 to May 2005 and from May to December of 2008 (personal communication, Charles McDade, 2009). The maximum reasonable β_{sulf} is estimated using the 99th percentile of those measurements (i.e., 52% RH). At this humidity, the AIM model (Wexler and Clegg, 2002) (available at http://www.aim.env.uea.ac.uk/aim/aim.php) computes hydrated $(NH_4)_2SO_4$ to have 53% more mass than dry $(NH_4)_2SO_4$ and hydrated NH_4HSO_4 to have 32% more mass than dry NH_4HSO_4. Therefore, 1.53 is a reasonable upper bound for β_{sulf}.

Nitrate coefficients less than 1 likely represent volatilization of NH_4NO_3 from the Teflon filter prior to gravimetric analysis. Hering and Cass (1999) report that the absolute amount of nitrate volatilization is a function of RH and temperature, but not ambient nitrate concentration (unless the calculated nitrate loss exceeds the ambient nitrate available). Thus, a proportional coefficient captures the average volatilization behavior reasonably well. Because a value of 0 (complete nitrate volatilization) would imply no statistical relationship between nitrate mass and $PM_{2.5}$ mass, a slightly negative β_{nit} value caused by measurement error is just as likely as a slightly positive β_{nit} value. Consequently, for each regression performed, we set the lowest reasonable value for β_{nit} as 1.5 standard errors below 0 (calculation of standard errors is described in the Supplement, Sect. S1.1). There are 129site/quarter groupings exhibiting negative β_{nit} values within 1.5 standard errors of 0. To show that these negative values really represent slight variations around 0, we repeat each of these regressions without the nitrate term and find that β_{OC} and β_{sulf} coefficients change by less than 3% on average (no β_{OC} and six β_{sulf} coefficients change by more than 5%). A β_{nit} greater than 1 indicates that the assumption of dry NH_4NO_3 underestimates the actual nitrate mass on the Teflon filter at the time of weighing. This would occur either if the cation has a larger molar mass than ammonium (e.g. Na) or if there is water associated with the nitrate during weighing. Again a maximum reasonable value for β_{nit} is determined by computing increases in water mass at 52% RH with the AIM model for both NH_4NO_3 and $NaNO_3$. This analysis shows that hydration can add 35% extra mass to the nitrate, so 1.35 is a reasonable upper bound for β_{nit}.

2.3 Effects of measurement uncertainty

Despite the aforementioned advantages of the regression method, it is subject to several pitfalls. One is that measurement uncertainty in the explanatory variables can bias the regression coefficients. An OLS regression assumes that explanatory variables are measured without error, but this assumption conflicts with the reality of our application in which measurement uncertainty is associated with all explanatory variables: OC, $(NH_4)_2SO_4$, NH_4NO_3, and SOIL. For regressions with a single explanatory variable that is uncertain, the coefficient is biased towards zero (Fuller, 1987;Saylor et al., 2006;White, 1998). With multiple explanatory variables, bias in the coefficients exhibits a complex dependency on the relative uncertainties in various components, the correlation between explanatory variables, the correlation between measurement errors, and other factors. White (1998) examined this problem in a simplified case with two correlated explanatory variables of which one was measured without error. For that case, he showed that the coefficient for the perfectly measured explanatory variable was artificially inflated while the other coefficient was diminished.

To evaluate this bias within the more complex conditions of the present study, we analyze synthetic datasets that mimic the IMPROVE data. Assuming that the actual values for each measurement were exactly equal to the reported value, we create 200 synthetic datasets for each site- and quarter-specific dataset that represent "observed" data with error in the explanatory variables. Errors are added by perturbing the reported values of OC, sulfate, nitrate, and $PM_{2.5}$ using the reported uncertainty and assuming that "observed" values would be normally distributed around the actual value. For each site- and quarter-specific dataset, we then perform an OLS regression on the reported dataset and the 200 synthetic datasets. The reported dataset is considered the "truth" in this exercise, so OLS regression yields "true" coefficients for comparison with the results from our synthetic datasets. Results from one such analysis for a regression with typical OLS biases (Gila Wilderness in New Mexico during quarter 1) are shown in the left half of each plot in Fig. 2. The dotted lines represent the "true" coefficients and the box plot shows the distribution of coefficients obtained from the 200 synthetic datasets. Although the true value could be accurately estimated from some synthetic datasets in this example, β_{OC} is typically under-estimated while β_{sulf} and β_{nit} are over-estimated.

To overcome the biases associated with the OLS assumption of error-free explanatory variables, a class of methods has been developed to explicitly account for the existence of such errors; these are often collectively called measurement error models or errors-in-variables (EiV) models. Such methods typically assume that for all observations of each covariate, the errors are independent, identically distributed and follow a normal distribution with mean zero and a fixed (possibly unknown) standard deviation. In the IMPROVE data, the standard deviation is not fixed because we have a different estimated error associated with each observation of a given covariate, which we take as the standard deviation of the error distribution. To accommodate this added complexity, we turn to an advanced measurement error model described by Fuller (1987) (Sec. 3.1.2). The following discussion is based entirely on Fuller's work, conforming to his original notation as much as is feasible.

To begin, we define Y_t as the value of the response variable for observation t, such that t=1, 2, …, n, with n representing the number of observations. For the multiyear regression, this response is given by $PM_{2.5} - (1.2\ KNON + 1.8\ Cl^- + EC)$, and for the quarter-specific regression it is $PM_{2.5} - (1.2\ KNON + 1.8\ Cl^- + EC + \beta_{soil}\ SOIL)$. The row vector X_t contains the observed values of the explanatory variables associated with observation t. The first element is the observed value of OC, the next element corresponds to $(NH_4)_2SO_4$, the third is NH_4NO_3, and the fourth is SOIL. (In the quarter-specific regression case, the SOIL component is omitted.) Note that the order of these explanatory variables mimics their order in Eq. (5) and is preserved in the various mathematical representations of their coefficients, errors, etc. which follow.

Additionally, we let $\sum_{uutt}$ represent the covariance matrix associated with X_t. Assuming that errors in each covariate are independent, this is a diagonal matrix. The elements along the diagonal contain the variance (square of the error standard deviation) associated with the explanatory variables, in the specified order. As an initial estimate for the regression coefficients, we use the method-of-moments estimator, the column vector $\widetilde{\beta}$, given by Eq. (8)

$$\widetilde{\beta} = \left[n^{-1} \sum_{t=1}^{n} (X_t' X_t - \Sigma_{uutt}) \right]^{-1} \left[n^{-1} \sum_{t=1}^{n} X_t' Y_t \right] \tag{8}$$

Having obtained this initial estimate, we work to refine it, as outlined by Fuller (1987). We define for each observation t the matrix $\sum_{aatt}$. This is also a diagonal matrix, with the elements along the diagonal consisting of the variance for the response followed by the variances for the explanatory variables in the specified order. We take the square of the reported measurement uncertainty for each chemical constituent in a particular sample as its variance. (Note that the $\sum_{uutt}$ featured in Eq. (8) is simply a submatrix of $\sum_{aatt}$.) We also let Z_t represent the row vector containing the observed response and the observed explanatory variables for each t; i.e., $Z_t = (Y_t, X_t)$. We then define the matrices M and A as

$$M = \sum_{t=1}^{n} \Sigma_{aatt} \text{ and } A = \sum_{t=1}^{n} Z_t' Z_t$$

With these defined, we can now obtain an estimate of the variance associated with the regression error, denoted σ_{qq}. We first solve for the eigenvalues of the matrix product $M^{-1}A$. If the minimum of these eigenvalues is less than one, then $\widetilde{\sigma}_{qq}$ is 0. Otherwise, $\widetilde{\sigma}_{qq}$ is given by Eq. (9):

$$\widetilde{\sigma}_{qq} = \sum_{t=1}^{n} \left[(n-k)^{-1} \left(Y_t - X_t \widetilde{\beta} \right)^2 - n^{-1} \left(1, -\widetilde{\beta}' \right) \Sigma_{aatt} \left(1, -\widetilde{\beta}' \right)' \right] \tag{9}$$

Both $\widetilde{\beta}$ and $\widetilde{\sigma}_{qq}$ are then used to obtain an estimate of the error associated with the linear relationship between the observed (with error) response and covariates, $\widetilde{\sigma}_{vvtt}$ [Eq. (10)]:

$$\tilde{\sigma}_{vvtt} = \tilde{\sigma}_{qq} + \sigma_{wwtt} + \tilde{\beta}'\Sigma_{uutt}\tilde{\beta} \tag{10}$$

where σ_{wwtt} is the measurement variance associated with the response at time t. To obtain our final estimate, $\hat{\beta}$, of the regression coefficients, we combine the previous elements to obtain Eq. (11):

$$\hat{\beta} = \left[\sum_{t=1}^{n} \tilde{\sigma}_{vvtt}^{-1}(X_t'X_t - \Sigma_{uutt})\right]^{-1} \sum_{t=1}^{n} \tilde{\sigma}_{vvtt}^{-1} X_t'Y_t \tag{11}$$

Here $\hat{\beta}$ is a column vector containing our estimates of β_{OC}, β_{sulf}, β_{nit}, and β_{soil} (for the multiyear regression). Fuller (1987) also provides an estimator for the covariance matrix of $\hat{\beta}$. We use the diagonal elements of this matrix to obtain the standard errors for our estimated regression coefficients. In the interest of brevity, we leave further discussion of this variance estimate to the supplement (Sect. S1). In addition, sample R code used to perform these regressions is also supplied in Sect. S1.

We recognize that our method includes several assumptions. Perhaps most notable is the assumption that the measurement errors are independent among all the covariates and the response measured at a given date and location. The method could be extended to include information about the correlation between measurement errors, if such were known. This would result in non-diagonal matrices $\sum_{uutt}$ and $\sum_{aatt}$. Another key assumption is that the measurement error distributions are normal. If this is an unreasonable assumption, we could explore more complex statistical models that allow for nonnormal measurement errors, which are currently a subject of statistical research.

To demonstrate that this new technique reduces the bias in coefficients, we reanalyze all of our synthetic datasets using the EiV regression methodology. The results for quarter 1 data from Gila Wilderness are shown in the right-hand box plots of Fig. 2. Clearly, the EiV method yields coefficients that are much closer to the "truth" than the OLS methodology. To confirm the generality of this result, Fig. 3 shows the distribution of bias across all site- and quarter-specific datasets. Substantial bias in β_{OC} (under-prediction), β_{sulf} (over-prediction), and β_{nit} (over-prediction) arise from the OLS regression, but these biases are greatly mitigated with the EiV technique. White (1986) provides a similar analysis of regression performance using measurements from the 1981 – 1982 Western Regional Air Quality Study. His analysis, which include three explanatory variables (sum of ionic sulfate, nitrate, and ammonium; organic carbon; sum of silicon dioxide and calcium oxide), also found that correcting for measurement uncertainty reduces bias in the coefficients.

Although the EiV methodology shows improved results, it should be noted that additional error arises if the measurement uncertainties are biased themselves. Hyslop and White (2008) report some systematic biases in the measurement uncertainty from XRF, ion chromatography, and TOR carbon measurements at IMPROVE sites. If future updates to the IMPROVE data include substantial changes to uncertainty estimates for these components, it may warrant some repetition of the present work. For all subsequent analyses discussed in this paper, we apply the EiV method (instead of OLS).

2.4 Statistical identification of high-confidence regressions

After applying the EiV method to each multiyear and quarter-specific dataset, it is tempting to begin examining spatial and temporal patterns in the regression coefficients. However, as emphasized by Malm and Hand (2007), "Regression coefficients are vulnerable to a variety of systematic and random errors." In this subsection, we establish some empirical guidelines for flagging or eliminating datasets that do not conform to Eq. (5). As summarized in the lower half of Fig. 1, these guidelines are subsequently applied to identify regression results that can be used with "high confidence" for applications such as air quality model evaluations, source-apportionment analyses, epidemiology studies, and radiative calculations.

2.4.1 Multicolinearity among explanatory variables

One requirement of our regression method (irrespective of choosing EiV or OLS) is that all explanatory variables be independent of each other. If any two $PM_{2.5}$ components are linearly related, the dataset is not suitable for regression analysis because the technique may over-estimate one coefficient and under-estimate another due to excess degrees of freedom. To identify such datasets, Pearson correlation coefficients (r_P) are calculated for all six couplings among the four explanatory variables (OC, $(NH_4)_2SO_4$, NH_4NO_3, and SOIL) in each site- and quarter-specific dataset. We examine all datasets having any $|r_P|$ values greater than 0.65 and look for cases in which the coefficient on one of the highly correlated explanatory variables appears to be over-estimated while the other appears under-estimated relative to the ranges established in Sect. 2.2. For example, sulfate and nitrate from 4th quarter measurements at the Puget Sound monitoring site are correlated with $r_P = 0.86$. In that regression, $\beta_{sulf} = 0.83$ (lower end of its physically reasonable range) and $\beta_{nit} = 1.28$ (higher end of its range). We regard these regression results as "suspect."

A summary of our analysis across all sites and quarters is shown in Fig. 4, from which we determine that $|r_P|$ values greater than 0.85 often indicate suspect results. We acknowledge that our empirical approach for setting this threshold value is not foolproof since 1) coefficients that appear skewed may actually be accurate, and 2) some regressions which are affected by co-linearity may not be identifiable if the estimated coefficients fall well within their physically reasonable ranges. However, our approach yields an easy-to-use procedure for screening out regression results that may be biased due to co-linearity in speciated $PM_{2.5}$ data. Seven quarter-specific datasets are eliminated from our analysis based on the $\max|r_P| > 0.85$ criterion (see list in Supplement Table S3).

2.4.2 Assessing the fit of the regression model

A second requirement for accurate regressions is that the Eq. used to fit coefficients is physically realistic. Based on our knowledge of ambient aerosol across the US, Eq. (5) includes all the essential $PM_{2.5}$ components. However, if a true coefficient for EC, Cl^-, or KNON is substantially different from our fixed coefficients for those species, the regression could be adversely affected. In addition, if the actual SOIL coefficient varies greatly throughout the year at any site, then our assumption of temporally-invariant β_{soil} could also degrade the regression results at that site. Finally, if the relationship between $PM_{2.5}$ mass and any major chemical component is nonlinear, our regression analysis will be inaccurate. For instance, if OC artifact corrections were biased high in clean conditions and vice versa, OC concentrations would be negatively [positively] biased in clean [polluted] conditions and the relationship between reported OC and total $PM_{2.5}$ would be nonlinear.

To identify cases influenced by one or more of these phenomena, we examine the residual errors [ε_i in Eq. (5)] resulting from each site- and quarter-specific regression. Spearman rank order correlation coefficients (r_S) are calculated between the ε_i values and each species used in Eq. (5): OC, S, NO_3^-, SOIL, EC, Cl^-, and KNON. Any strong correlation indicates that Eq. (5) is an inadequate representation of $PM_{2.5}$ at the given site/quarter. Examples are shown in Fig. 5. Following this analysis, a criterion of $|r_S| > 0.4$ is imposed to eliminate 12 quarter-specific datasets that are likely affected by the problems discussed above (see list in Supplement Table S4). Nine of these datasets exhibit a strong correlation between ε_i and Cl^-, largely due to an abundance of negative Cl^- concentrations in the underlying IMPROVE data. The negative Cl^- values in 2002 and 2003 were caused by variability in filter blanks and a change of filter suppliers in 2004 corrected this problem (White, 2008). This exemplifies a need to understand the underlying data before interpreting any results from a regression analysis.

2.4.3 Dataset selection and segregation

A third key element to obtaining meaningful regression coefficients from IMPROVE measurements is appropriate segregation of data. For this analysis, data are grouped by season and monitoring site with the intention that samples taken within each subset should yield fairly constant regression coefficients. However, sites that are strongly influenced by time-varying sources may not match our intent and therefore may not be ideal input for the regression analyses. For instance, a monitoring site that is impacted heavily on certain days by wildfires and on other days by diesel traffic will exhibit varying OM/OC ratios that violate our assumption of constant β coefficients by quarter.

To check for temporal trends or irregularities during our 7 year study period, residual error values were binned by year and examined for each site- and quarter-specific dataset. This analysis was designed to identify three possible problems: 1) a one-time abrupt change in ε_i which could indicate a change in measurement methods, 2) a monotonic temporal trend in ε_i which could indicate changing aerosol characteristics at the site, possibly due to the implementation of regulatory controls on emissions, and 3) a single year which showed vastly different ε_i from other years indicating that a distinct and infrequent event (e.g., forest fire or abnormal meteorology) affected the monitoring site.

Visual inspection of all datasets shows no evidence of problem 1. There was a change in EC and OC measurement equipment between 2004 and 2005 (White, 2007) as well as a coincident change in the calibration of the XRF sulfur measurements (White, 2009a). (Details about these and other such changes to IMPROVE data can be found at http://vista.cira.colostate.edu/improve/Data/QA_QC/Advisory.htm.) Despite these changes in OC, EC, and sulfur, no shift in residual values is apparent between 2004 and 2005 for the network as a whole (see Fig. 6). That year-to-year change is no greater than other inter-annual variations. Though we found no observable effect, we acknowledge that any change in measurement techniques adds uncertainty to our final results.

Seven site- and quarter-specific datasets exhibit temporal trends in which median residual values or the inter-quartile range of residual values either increase or decrease monotonically from 2002-2008 (i.e., problem 2 outlined above). One example is shown in Fig. 7a and all seven are listed in Supplement Table S6. Further investigation of these datasets by people with site-specific expertise would be worthwhile. Though we report these 7 sets of regression coefficients, we do not regard them as high-confidence results.

Finally, sites affected by an infrequent event are identified using two criteria: the inter-quartile range of ε_i in a single year does not overlap the inter-quartile ranges from any other year; or the year with the broadest inter-quartile range is greater than two times the second broadest inter-quartile range. An example of each phenomenon is shown in Figs. 7b and 7c. We re-run these regressions without the errant year and report results from both the full and abridged datasets in Table S6 of the Supplement. Of the 28 cases flagged, we regard 10 as high-confidence results because none of their coefficients are perturbed by more than 0.1 when the outlier year is removed. These cases are shaded in gray in Supplement Table S6 and also appear in Table S5. In the remaining 18 cases, further examination of the underlying by site-specific experts is warranted.

3 Results

Table S2 in the supplement shows our multiyear regression results. Tables S5, S6, and S7 show coefficients for all quarter-specific regressions along with standard error values, normalized mean errors (NME), and normalized mean biases (NMB). NME and NMB are calculated using Eqs. (12) and (13). NMB and NME values are generally small (mean NMB for all regressions in tables S5, S6, and S7 = -0.2%, maximum absolute NMB = 2.6%, mean NME = 8.5%, maximum NME = 22.6%) indicating that the IMPROVE data fit Eq. (5) quite well.

$$NME = \left(\frac{\sum_{i=1}^{n} |\varepsilon_i|}{\sum_{1}^{n} PM_{2.5,i}} \right) \times 100\% \quad (12)$$

$$NMB = \left(\frac{\sum_{i=1}^{n} \varepsilon_i}{\sum_{1}^{n} PM_{2.5,i}} \right) \times 100\% \quad (13)$$

3.1 Physically unreasonable results

Only 7 of the multiyear regressions (i.e., < 5% of all IMPROVE sites) have a coefficient that is physically unreasonable (see Table S2). Of these, 2 have β_{soil} values (0.21 and

0.27) falling below those of known soil profiles (see Sect. 2.2). Both low β_{soil} values come from urban IMPROVE sites (New York City and Washington D.C.). In these locations, there are likely non-soil sources of Si, Ca, Fe, or Ti. For instance, residential wood combustion is a major source of all four elements, on-road vehicle exhaust is a major source of Si, Ca, and Fe, and surface coating operations are a major source of Ti (Reff et al., 2009). In urban areas where such sources may dominate, Eq. (7) would overestimate total soil mass and might yield an erroneously low value of β_{soil}. The other 5 problematic multiyear regressions have low β_{nit} values, for which the cause is unclear. We are nevertheless able to extract high-confidence values of β_{OC} at these sites by using the multiyear β_{soil} value in our quarter-specific regressions.

In total, 61 quarter-specific regressions (10%) have at least one physically unreasonable coefficient (see Supplement Table S7). The number of regressions with problematic coefficients is greatest in quarter 1 (n = 21) and quarter 3 (n = 22) and least in quarters 2 and 4 (n = 13 and n = 5 respectively). Problematic β_{soil} values from the multiyear regressions account for 8 of these (2 in each quarter).

Twenty of the 61 regressions with physically unrealistic coefficients are due to β_{OC} values less than unity, 17 of which occur in quarter 1. These low β_{OC} values may be caused by errors in OC artifact correction, as discussed in Sect. 3.3 and Supplement Sect. 3. Although the low β_{OC} values predominantly occur in quarter 1, this may be exacerbated by the fact that β_{OC} values are lower in quarter 1 than in other quarters (median β_{OC} in quarters 1, 2, 3, and 4 are 1.39, 1.83, 1.81, and 1.59 respectively). Therefore, a slight low bias would push more OM/OC ratios below 1 in the winter than in other seasons.

Eighteen of the 61 problematic regressions are due to negative β_{nit} values that are more than 1.5 standard errors below zero. Fourteen of these occur in quarter 3. There are two possible explanations for the high occurrence in quarter 3. First, nitrate concentrations are generally low in the summer. In quarter 3, network-wide median nitrate concentrations were only 3% of median $PM_{2.5}$ (versus 11% and 6% for quarter 1 and the annual average, respectively). When the mass of an explanatory variable is low compared to the mass of other $PM_{2.5}$ components, the model fit is not very sensitive to large changes in that coefficient. Second, these problematic β_{nit} estimations may be due to a large number of cases in quarter 3 when all the nitrate volatilized from the Teflon filter (see Sect. 2.2). The lower-bound for negative β_{nit} values (1.5 standard errors below 0) may be too conservative, leading us to flag regressions in which nitrate volatilization is 100% (i.e. β_{nit} is essentially 0) as problematic.

The third most frequent error comes from high β_{nit} values: 13 regressions estimate $\beta_{nit} >$ 1.35 In general these data points have higher than average standard errors (the mean nitrate standard error for these regressions is 0.50 while the mean nitrate standard error for all site-specific regressions is 0.21). These large standard errors indicate highly uncertain estimates of β_{nit}, possibly due to low nitrate concentrations.

Overall, 90% of our quarter-specific regressions yield physically reasonable coefficients for all four explanatory variables in Eq. (5). This leaves 511 high-confidence regressions (see Fig. 1) from which we can assess spatial and seasonal trends.

3.2 Spatial and temporal trends in β_{soil}, β_{sulf} and β_{nit}

Figure 8 shows the spatial pattern of β_{soil} . Much of the country has β_{soil} values near 1, confirming that Eq. (7) does a reasonable job of estimating soil concentrations. Some notable departures from this are high values displayed in orange and red in the southwestern US and lower values (green and blue) in much of the Midwest. Both of these are consistent with the calculated β_{soil} values for different soil types (Simon et al., 2010). They report β_{soil} values for desert soil between 1.25 and 1.4 and β_{soil} values for agricultural soil between 0.78 and 1.10.

In order to evaluate spatial and temporal trends for β_{sulf} and β_{nit}, regression results are grouped by region, matching the organizations designated by the EPA to address regional haze (EPA, 2010). Hereafter, states included in WRAP, CENRAP, LADCO, MANE-

VU, and VISTAS will be referred to as the western, central, great lakes, northeast, and southeast regions, respectively.

Maps of β_{sulf} during each quarter are given in the supplement (Figs. S6-S9). Figure 9 shows a summary of β_{sulf} values from 593 quarter-specific regressions. Apart from the western region, β_{sulf} follows a seasonal trend in which values are lowest in the winter (median values in the central, southeast, great lakes, and northeast regions are 0.90, 0.92, 0.91, and 0.88, respectively) and highest in the summer (corresponding medians are 1.05, 1.04, 1.09, and 1.09). The median wintertime values less than 1 suggest that sulfate is not fully neutralized by ammonium in quarter 1. The summertime values greater than 1 suggest wet sulfate. Further analysis presented in the Supplement Sect. S2 suggests that the trends in Fig. 9 (excluding the western region) are reasonably explained by the seasonal variation in laboratory RH where samples were weighed and by the degree of sulfate neutralization.

Quarter-specific maps of β_{nit} are given in the supplement (Figs. S10-S13). Figure 10 summarizes the temporal and spatial trends. In general, β_{nit} values are lower (i.e. higher percentages of nitrate is volatilized from the Teflon filter) in locations and in seasons where temperature is higher. For example, the southeast is warmer, on average, than the rest of the country throughout the year. Median β_{nit} in this region are lower than all other regions in every quarter. Similarly, summer β_{nit} values are lower than winter values in all regions. In addition, regions which experience the most dramatic seasonal temperature variations (central, great lakes, and northeast) have the most dramatic variation in median β_{nit} values. Finally we posit that any site whose β_{nit} value is within 1.5 standard deviations of 0 is prone to total nitrate volatilization. The number of sites falling into this category increase from 6 in the winter to 71 in the summer, again showing that more nitrate volatilizes in warmer months. Since nitrate volatilization is governed by the temperature-dependent nitrate equilibrium (Hering and Cass, 1999), this behavior is expected. Figure 10 also exhibits a large range of β_{nit} values in quarter 3 which may be due, in part, to low nitrate concentrations. This large seasonal variation coupled with the large standard error for β_{nit} in quarter 3 (median = 0.34, versus 0.06, 0.16, and 0.08 in other quarters) indicate that the regression model is not precisely estimating β_{nit} in the summer months, though the seasonal variations in β_{nit} are believable. Furthermore, the median standard error for β_{nit} is much larger in quarter 3 (0.34) than in other quarters (0.06 in quarter 1, 0.16 in quarter 2, and 0.08 in quarter 4).

3.3 OM/OC results

Our analyses of spatial and temporal trends in β_{sulf}, β_{nit}, and β_{soil} show that they mostly can be explained by known aerosol properties and sampling artifacts. Those results build confidence in our estimates of the OM/OC ratio, β_{OC}. Table 2 summarizes the distribution of β_{OC} values across all regions for all quarters. Table 2 and Fig. 11a show that wintertime OM/OC ratios are generally higher in the eastern US than the West. Median β_{OC} values during quarter 1 in the great lakes, southeast, and northeast regions are 1.58, 1.64, and 1.51 respectively while the west and central regions exhibit 1.29 and 1.32 respectively. Higher OM/OC ratios in the eastern US may be a result of high residential wood smoke emissions (see Fig. S10f of Reff et al., 2009). In addition, high values in the southeast may be due to SOA, which is more abundant in this region than in other US regions during winter months (Yu et al., 2007). Figure 11b suggests that OM/OC ratios in the summer do not vary substantially by region; median β_{OC} values are 1.80, 1.81, 1.93, 1.87, and 1.81 in the west, central, great lakes, southeast, and northeast regions, respectively. The range of β_{OC} values within regions is also quite consistent across the US during quarter 3 (see Table 2). Maps of β_{OC} during quarters 2 and 4 are given in the Supplement Fig. S14.

Seasonal variations in β_{OC} can also be seen in Fig. 12, which shows β_{OC} values are generally higher during summer than in winter. Regressions at only 12 sites yield higher β_{OC} values in quarter 1 than 3 (out of 146 available pairs). This is consistent with higher SOA concentrations in the summer and more aging of primary OC due to higher oxidant concentrations than in winter. While the winter medians are low, β_{OC} is more variable than in other seasons: in quarter 1, 90% of β_{OC} values fall between 0.79 and 1.84; in quarter 3, 90% fall between 1.44 and 2.08. Although this seasonal trend is seen at the vast majority of IMPROVE sites, it is important to note that local conditions have cuased higher wintertime β_{OC} values in a small number of locations.

As mentioned in Sect. 2.2, β_{OC} is influenced by differences in the OC sampling artifacts on quartz versus Teflon filters. Whereas the literature is inconclusive regarding negative artifacts, quartz filters are more prone to positive artifact than Teflon filters. The IMPROVE data include a network-wide and month-specific correction for positive OC artifact on the quartz filter, but no correction for the Teflon filter. We evaluate the effects of site-to-site variability in positive OC artifact (quartz filter) on our regression results (see Supplement Sect. S3) and conclude that the network-wide artifact correction does not substantially affect our estimates of β_{OC}. However, the β_{OC} value could be skewed if (1) IMPROVE's back-up filter method does not completely capture all positive artifact on quartz filters, (2) Teflon filters have non-negligible positive artifact, or (3) the magnitude of negative artifact differs on the quartz and Teflon filters. An in-depth exploration of OC artifact is beyond the scope of this paper, but these uncertainties should be kept in mind when interpreting our regression results.

Our low wintertime β_{OC} estimates in the west and central regions (medians near 1.3) suggest an aerosol dominated by fresh, mobile-source emissions. Although oxidative aging and SOA formation are limited in these regions during winter, the US National Emissions Inventory indicates that other PM sources (e.g., wood smoke) increase β_{OC} to 1.5 or 1.6. Our low β_{OC} results may be a consequence of systematic biases in the reported measurement uncertainty, which the EiV regression is dependent upon (see Sect. 2.3). Another possibility is that our low β_{OC} results are somehow tied to the high wintertime β_{sulf} values in the western region, which we are unable to explain (see Sect. 3.2).

3.4 Differences with previous regression estimates of OM/OC

Differences between our methodology and that used by Malm and Hand (2007), referred to hereafter as MH07, are summarized in Table 3. A major difference is that we emphasize seasonal β_{OC} values, whereas MH07 focused on multiyear regression results. Beyond that, it is interesting to explore which of our subtle revisions to the MH07 methodology cause substantial changes in β_{OC}. Figure 13 compares β_{OC} results from our multiyear regressions (Supplement Table S2) with the MH07 results. Our β_{OC} estimates at 37% of sites differ from MH07 by more than 0.2, and 61% differ by more than 0.1. Within each region, our β_{OC} estimates exhibit less site-to-site variability than MH07. For example, our low β_{OC} values in the great lakes and southeast regions (5th percentile = 1.7 and 1.5, respectively) are higher than MH07 (1.4 and 1.3) despite similar medians. In addition, 95th percentile β_{OC} values in the west and central regions are lower in our multiyear regressions (1.9) than in MH07 (2.1).

To isolate the main cause of these different β_{OC} results, we perform a series of regressions, beginning with the approach of MH07, that incrementally incorporates each methodological revision listed in Table 3. The three parameters which have the largest effect on β_{OC} are the dataset download date, the years analyzed (i.e. 1988-2003 vs. 2002-2008), and the choice of explanatory variables (i.e. differences between Eqs. (4) and (5)). The use of EiV rather than OLS affects β_{OC} to a smaller degree. Using S instead of SO_4^{2-} to calculate ammonium sulfate and Eq. (7) instead of Eq. (2) to compute SOIL have almost no effect on the β_{OC} estimates. The download dates are important because the IMPROVE data archive is updated whenever errors are found. For example, historic chlorine data were adjusted in November 2009 because the original blank correction was deemed too low (White, 2009b). The large effect of the years analyzed may indicate a long-term trend in β_{OC} (about 64% of the sites have higher β_{OC} values when using 2002-2008 data than when using 1988-2003 data), or result from changes to measurement protocols and hardware which occurred during these time periods. Taken together, the effects of download date and years analyzed indicate a sensitivity of β_{OC} to changes in the measurements and data processing methodology.

Next, we analyze which specific changes between Eqs. (4) and (5) cause the largest difference in β_{OC} values. We find that accounting for KNON, removing the intercept (β_0), and fixing the Cl^- coefficient to 1.8 have almost no impact on β_{OC}. However, fixing the EC coefficient to 1 changes β_{OC} by more than 0.2 at 15% of the sites. We attribute this sensitivity to the fact that EC and OC are highly collinear in the IMPROVE data (r_P

exceeds 0.85 and 0.65 at 20% and 88% of sites, respectively). These high correlation coefficients imply that inclusion of EC as an explanatory variable will likely attribute some EC mass to β_{OC} or some OM to β_{EC}. In the Supplement (Sect. S3), we investigate our assumption of $\beta_{EC} = 1$ and find that it has little impact on our β_{OC} estimates. However, we also find that MH07 grossly underestimated β_{OC} at about ¼ of the IMPROVE sites due to unrealistically large values of β_{EC}. This helps explain why our 5th percentile β_{OC} values are higher than MH07.

4 Summary and future work

This work has helped to develop a robust technique for estimating OM/OC ratios that can be applied to an expansive dataset, such as the IMPROVE monitoring network data. Our ability to estimate physically reasonable spatial and seasonal trends in β_{sulf}, β_{nit}, and β_{soil} builds confidence in our β_{OC} results. Furthermore, our major methodological improvements include the use of an errors-in-variables regression and the elimination of EC as an explanatory variable. These two changes provide more realistic results and eliminate substantial biases from approximately ¼ of the regressions performed by Malm and Hand (2007). The reader is cautioned that all of our conclusions about OM/OC ratios rely on quartz and Teflon filter measurements and, hence, depend on accurate and complete OC artifact corrections on both filter types. Techniques for quantifying these artifacts are still an active area of research. Comparison of our β_{OC} results with other OM/OC estimation methods will be the subject of future work.

In addition, this work has identified future areas of research into the IMPROVE data. First, our analysis shows that sulfate is often not fully neutralized so ammonium measurements will greatly assist future mass closure efforts. Second, nitrate volatilization appears to vary substantially by site and season. A measurement study could be performed to verify the nitrate volatilization estimates made here. In addition, samples could be shipped in refrigerated conditions to prevent nitrate volatilization during transport. At a minimum, these results demonstrate the importance of recording the temperature and RH that filters are exposed during sampling, transport, and measurement. Most importantly, this work has identified general temporal and spatial trends in OM/OC ratios. We find that summertime OM/OC ratios are larger than wintertime values across the US and that winter values are larger in the eastern US than in the West. Considering this work plus the results of Malm and Hand (2007) and El-Zanan et al. (2005), users of the IMPROVE data should relax the common assumption of a fixed OM/OC ratio when calculating reconstructed fine mass.

Acknowledgements

The authors thank Warren White for helpful information on IMPROVE chloride measurements and regression methods, Chuck McDade for IMPROVE laboratory RH data, Lowell Ashbaugh for IMPROVE OC back-up filter measurement data, Doug Lowenthal for details of his mass closure methods, and Mark Pitchford, Venkatesh Rao, Ann Dillner, and Steve McDow for miscellaneous feedback and encouragement. The United States Environmental Protection Agency through its Office of Research and Development funded and managed the research described here. It has been subjected to Agency's administrative review and approved for publication.

References

\bibitem{}
Aiken, A., Decarlo, P. F., Kroll, J. H., et al.: O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry, Environ. Sci. Technol., 42, 4478--4485, doi:10.1021/es703009q, 2008.

\bibitem{}
Chan, T. W., Huang, L., Leaitch, W. R., Sharma, S., Brook, J. R., Slowik, J. G., Abbatt, J. P. D., Brickell, P. C., Liggio, J., Li, S.-M., and Moosm\"{u}ller, H.: Observations of OM/OC and specific attenuation coefficients (SAC) in ambient fine PM at a rural site in central Ontario, Canada, Atmos. Chem. Phys., 10, 2393--2411, doi:10.5194/acp-10-2393-2010, 2010.

\bibitem{}
Chen, X. and Yu, J. Z.: Measurement of organic mass to organic carbon ratio

in ambient aerosol samples using a gravimetric technique in combination with chemical analysis, Atmos. Environ., 41, 8857--8864, doi:10.1016/j.atmosenv.2007.08.023, 2007.

\bibitem{}
Chhabra, P.S., Flagan, R.C., Seinfeld, J.H.: Elemental analysis of chamber organic aerosol using an aerodyne high-resolution aerosol mass spectrometer, Atmost. Chem. Phys., 10, 4111--4131, doi:10.5194/acp-10-4011-2010, 2010.

\bibitem{}
El-Zanan, H. S., Lowenthal, D. H., Zielinska, B., et al.: Determination of the organic aerosol mass to organic carbon ratio in IMPROVE samples, Chemosphere, 60, 485--496, doi:10.1016/j.chemosphere.2005.01.005, 2005.

\bibitem{}
El-Zanan, H. S., Zielinska, B., Mazzoleni, L. R., et al.: Analytical determination of the aerosol organic mass-to-organic carbon ratio, J. Air Waste Manage., 59, 58--69, doi:10.3155/1047-3289.59.1.58, 2009.

\bibitem{}
EPA regional planning orgnizations: \url{http://epa.gov/visibility/regional.html\#thefive}, last access: 14~October, 2010.

\bibitem{}
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., 56, 500--511, 2006.

\bibitem{}
Fuller, W. A.: Measurement error models, John Wiley\,\&\,Sons, New York, USA, 1987.

\bibitem{}
Gilardoni, S., Russell, L. M., Sotooshian, A., et al.: Regional variation of organic functional groups in aerosol particles on four US east coast platforms during the International Consortium for Atmospheric Research on Transport and Transformation~2004 campaign, J. Geophys. Res.-Atmos., 112, D10S27, doi:10.1029/2006jd007737, 2007.

\bibitem{}
Hand, J. L. and Malm, W. C.: Review of the IMPROVE equation for estimating

ambient light extinction coefficients -- final report, Colorado State University, CIRA, 146, 2006.

\bibitem{}
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., 49, 725--733, 1999.
\bibitem{}
Huang, X.-F., He, L.-Y., Hu, M., Canagaratna, M.R., Sun, Y., Zhang, Q., Zhu, T., Xue, L., Zeng, L.-W., Liu, X.-G., Zhang, Y.-H., Jayne, J.T., Ng, N.L., Worsnop, D.R.: Highly time-resolved chemical characterization of atmospheric submicron particles during 2008 Beijing Olympic Games using the Aerodyne High-Resolution Aerosol Mass Spectrometer, Atmost. Chem. Phys., 10, 8933--8945, 2010.

\bibitem{}
Hyslop, N. P. and White, W. H.: An evaluation of interagency monitoring of protected visual environments (IMPROVE) collocated precision and uncertainty estimates, Atmos. Environ., 42, 2691--2705, doi:10.1016/j.atmosenv.2007.06.053, 2008.

\bibitem{}
Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., et al.: Evolution of organic aerosols in the atmosphere, Science, 326, 1525--1529, doi:10.1126/science.1180353, 2009.

\bibitem{}
Kiss, G., Varga, B., Galambos, I., et al.: Characterization of water-soluble organic matter isolated from atmospheric fine aerosol, J. Geophys. Res.-Atmos., 107, 8339, doi:10.1029/2001jd000603, 2002.

\bibitem{}
Kleindienst, T. E., Jaoui, M., Lewandowski, M., et al.: Estimates of the contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol at a southeastern~US location, Atmos. Environ., 41, 8288--8300, doi:10.1016/j.atmosenv.2007.06.045, 2007.

\bibitem{}
Liu, S., Takahama, S., Russell, L. M., Gilardoni, S., and Baumgardner, D.: Oxygenated organic functional groups and their sources in single and submicron

organic particles in MILAGRO~2006 campaign, Atmos. Chem. Phys., 9, 6849--6863, doi:10.5194/acp-9-6849-2009, 2009.

\bibitem{}
Lowenthal, D., Zielinska, B., Mason, B., Samy, S., Sambourova, V., Collins, D., Spencer, C., Taylor, N., Allen, J., Kumar, N.: Aerosol characterization studies at Great Smoky Mountains National Park, summer 2006, J. Geophys. Res.-Atmos., 114, doi:10.1029/2008jd011274, 2009.

\bibitem{}
Lowenthal, D. H. and Kumar, N.: PM$_{2.5}$ mass and light extinction reconstruction in IMPROVE, J. Air Waste Manage., 53, 1109--1120, 2003.

\bibitem{}
Malm, W. C., Sisler, J. F., Huffman, D., et al.: Spatial and seasonal trends in particle concentration and optical extinction in the United States, J. Geophys. Res.-Atmos., 99, 1347--1370, 1994.

\bibitem{}
Malm, W. C., Day, D. E., Carrico, C., et al.: Intercomparison and closure calculations using measurements of aerosol species and optical properties during the Yosemite Aerosol Characterization Study, J. Geophys. Res.-Atmos., 110, D14302, doi:10.1029/2004jd005494, 2005.

\bibitem{}
Malm, W. C. and Hand, J. L.: An examination of the physical and optical properties of aerosols collected in the IMPROVE program, Atmos. Environ., 41, 3407--3427, doi:10.1016/j.atmosenv.2006.12.012, 2007.

\bibitem{}
McDade, C. E.: IMPROVE standard operating procedure, Crocker Nuclear Laboratory, University of California, Davis, CASOP~351-2, 258, 2008.

\bibitem{}
McDow, S. R. and Huntzicker, J. J.: Vapor adsorption artifact in the sampling of organic aerosol -- face velocity effects, Atmos. Environ., 24A, 2563--2571, 1990.

\bibitem{}
Murphy, D. M., Cziczo, D. J., Froyd, K. D., et al.: Single-particle mass spectrometry of tropospheric aerosol particles, J. Geophys. Res.-Atmos., 111(15), D23S32, doi:10.1029/2006jd007340, 2006.

\bibitem{}
Pang, Y., Turpin, B. J., and Gundel, L. A.: On the importance of organic oxygen for understanding organic aerosol particles, Aerosol Sci. Tech., 40, 128--133, doi:10.1080/02786820500423790, 2006.

\bibitem{}
Pitchford, M., Malm, W. C., Schichtel, B., et al.: Revised algorithm for estimating light extinction from IMPROVE particle speciation data, J. Air Waste Manage., 57, 1326--1336, doi:10.3155/1047-3289.57.11.1326, 2007.

\bibitem{}
Polidori, A., Turpin, B.J., Davidson, C.I., et al.: Organic PM2.5 Fractionation by polarity, FTIR spectroscopy, and OM/OC ratio for the Pittsburgh aerosol, Aerosol Sci. Tech., 42, 233--246, doi:10.1080/02786820801958767, 2008.

\bibitem{}
R~Development Core Team: R: A language and environment for statistical computing, R~Foundation for Statistical Computing, \url{http://www.R-project.org}, last access: 14~October~2010, ISBN 3-900051-07-0, Vienna, Austria, 2010.

\bibitem{}
Reff, A., Turpin, B. J., Offenberg, J. H., et al.: A functional group characterization of organic PM$_{2.5}$ exposure: Results from the RIOPA study, Atmos. Environ., 41, 4585--4598, doi:10.1016/j.atmosenv.2007.03.054, 2007.

\bibitem{}
Reff, A., Bhave, P. V., Simon, H., et al.: Emissions inventory of PM$_{2.5}$ trace elements across the United States, Environ. Sci. Technol., 43, 5790--5796, doi:10.1021/es802930x, 2009.

\bibitem{}
Russell, L. M.: Aerosol organic-mass-to-organic-carbon ratio measurements, Environ. Sci. Technol., 37, 2982--2987, doi:10.1021/es026123w, 2003.

\bibitem{}
Russell, L. M., Takahama, S., Liu, S., et al.: Oxygenated fraction and mass of organic aerosol from direct emission and atmospheric processing measured on the R/V~Ronald Brown during TEXAQS/GoMACCS~2006, J. Geophys. Res.-Atmos., 114, D00F05, doi:10.1029/2008jd011275, 2009.

\bibitem{}
Saylor, R. D., Edgerton, E. S., and Hartsell, B. E.: Linear regression techniques for use in the EC tracer method of secondary organic aerosol estimation, Atmos. Environ., 40, 7546--7556, doi:10.1016/j.atmosenv.2006.07.018, 2006.

\bibitem{}
Simon, H., Beck, L., Bhave, P. V., et al.: The development and use of EPA's SPECIATE database, Atmos. Pollut. Res., 1, 196--206, 2010.

\bibitem{}
Sun, Y., Zhang, Q., Macdonald, A. M., Hayden, K., Li, S. M., Liggio, J., Liu, P. S. K., Anlauf, K. G., Leaitch, W. R., Steffen, A., Cubison, M., Worsnop, D. R., van Donkelaar, A., and Martin, R. V.: Size-resolved aerosol chemistry on Whistler Mountain, Canada with a high-resolution aerosol mass spectrometer during INTEX-B, Atmos. Chem. Phys., 9, 3095--3111, doi:10.5194/acp-9-3095-2009, 2009.

\bibitem{}
Surratt, J. D., Gomez-Gonzalez, Y., Chan, A. W. H., et al.: Organosulfate formation in biogenic secondary organic aerosol, J. Phys. Chem.~A, 112, 8345--8378, doi:10.1021/jp802310p, 2008.

\bibitem{}
Turpin, B. J., Huntzicker, J. J., and Hering, S. V.: Investigation of organic aerosol sampling artifacts in the Los Angeles basin, Atmos. Environ., 28, 3061--3071, 1994.

\bibitem{}
Turpin, B. J. and Lim, H.-J.: Species contributions to PM$_{2.5}$ mass concentrations: Revisiting common assumptions for estimating organic mass, Aerosol Sci. Tech., 35, 602--610, 2001.

\bibitem{}
Watson, J. G., Chow, J. C., Chen, L.-W. A., et al.: Assessment of carbon sampling artifacts in the IMPROVE, STN/CSN, and SEARCH networks, Desert Research Institute, Reno, NV, 2008.

\bibitem{}
Wexler, A. S. and Clegg, S. L.: Atmospheric aerosol models for systems including the ions H^{+}, NH$_{4}^{+}$, Na^{+}, SO$_{4}^{2-}$, NO^{3-}, Cl^{-}, Br^{-}, and H$_{2}$O, J. Geophys. Res.-Atmos., 107, 4207, doi:10.1029/2001jd000451, 2002.

\bibitem{}
White, W. H. and Roberts, P. T.: On the nature and origins of visibility-reducing aerosols in the Los Angeles air basin, Atmos. Environ., 11, 803--812, 1977.

\bibitem{}
White, W. H.: On the theoretical and empirical-basis for apportioning extinction by aerosols -- a critical-review, Atmos. Environ., 20, 1659--1672, 1986.

\bibitem{}
White, W. H.: Statistical considerations in the interpretation of size-resolved particulate mass data, J. Air Waste Manage., 48, 454--458, 1998.

\bibitem{}
White, W. H.: IMPROVE data advisory: Shift in EC/OC split with 1~January~2005 TOR hardware upgrade, 2007.

\bibitem{}
White, W. H.: Chemical markers for sea salt in IMPROVE aerosol data, Atmos. Environ., 42, 261--274, doi:10.1016/j.atmosenv.2007.09.040, 2008.

\bibitem{}
White, W. H.: IMPROVE data advisory, Inconstant bias in XRF sulfur, 2009a.

\bibitem{}
White, W. H.: IMPROVE data advisory: Under-correction of chloride concentrations for filter blank levels -- historical advisory, Applies to downloads before 23~November~2009, 2009b.

\bibitem{}
Yu, L. E., Shulman, M. L., Kopperud, R., et al.: Fine organic aerosols collected in a humid, rural location (Great Smoky Mountains, Tennessee, USA): Chemical and temporal characteristics, Atmos. Environ., 39, 6037--6050, doi:10.1016/j.atmosenv.2005.06.043, 2005.

\bibitem{}
Yu, S., Bhave, P. V., Dennis, R. L., et al.: Seasonal and regional variations of primary and secondary organic aerosols over the continental United States: Semi-empirical estimates and model evaluation, Environ. Sci. Technol., 41, 4690--4697, doi:10.1021/es061535g, 2007.

```
\bibitem{}
Zhang, Q., Jimenez, J. L., Canagartna, M. R., et al.: Ubiquity and
dominance of oxygenated species in organic
aerosols in anthropogenically-influenced Northern Hemisphere
midlatitudes,
Geophys. Res. Lett., 34, L13801, doi:10.1029/2007gl029979, 2007.

\end{thebibliography}
```

Table 1. Summary of measurement techniques and filter types for each PM component included in the regression analyses. For details, see Malm et al. (2004) and in the IMPROVE data guide (http://vista.cira.colostate.edu/improve/Publications/OtherDocs/IMPROVEDataGuide/IMPROVEdataguide.htm).

Chemical Component	Measurement Technique	Filter Type
$PM_{2.5}$	Gravimetric	Teflon
Nitrate and Chloride	Ion Chromatography	Nylon
Si, S, K, Ca, Ti, and Fe	X-Ray Fluorescence	Teflon
OC and EC	Thermal Optical Reflectance	Quartz

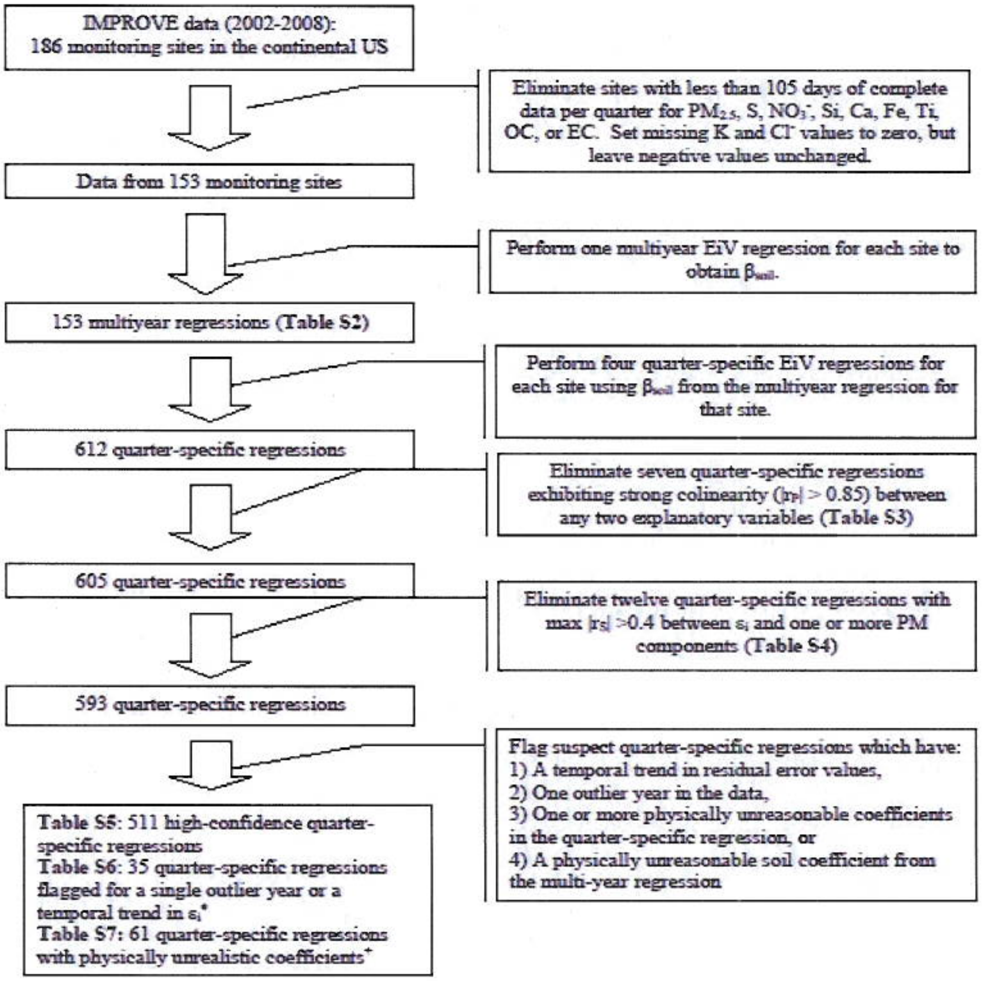
Table 2. Summary of β_{OC} distributions across sites for each quarter and region.

Region	Quarter	B_{OC}					Number of regressions
		5th percentile	25^{th} percentile	50th percentile	75th percentile	95th percentile	
West	1	0.67	1.06	1.29	1.42	1.76	89
West	2	1.36	1.66	1.81	1.90	2.14	86
West	3	1.33	1.66	1.80	1.88	2.04	85
West	4	1.22	1.43	1.57	1.68	1.88	86
Central	1	1.18	1.27	1.32	1.52	1.64	21
Central	2	1.59	1.69	1.78	1.87	2.10	21
Central	3	1.51	1.72	1.81	1.92	2.07	19
Central	4	1.37	1.45	1.53	1.64	1.90	21
Great Lakes	1	1.43	1.44	1.58	1.81	1.98	5
Great Lakes	2	1.83	1.83	1.94	1.95	1.97	5
Great Lakes	3	1.67	1.90	1.93	1.95	2.01	5
Great Lakes	4	1.31	1.31	1.48	1.61	1.61	5
Southeast	1	1.44	1.58	1.64	1.80	1.87	17
Southeast	2	1.50	1.76	1.89	2.00	2.16	16
Southeast	3	1.47	1.75	1.87	2.08	2.25	16
Southeast	4	1.42	1.60	1.67	1.75	1.83	17
Northeast	1	1.29	1.43	1.51	1.60	1.78	20
Northeast	2	1.23	1.74	1.87	2.01	2.09	19
Northeast	3	1.69	1.76	1.81	1.90	2.03	20
Northeast	4	1.07	1.49	1.57	1.67	1.85	16
all	1	0.79	1.20	1.39	1.58	1.84	153
all	2	1.39	1.69	1.83	1.94	2.15	148
all	3	1.44	1.72	1.81	1.91	2.08	146
all	4	1.24	1.44	1.59	1.68	1.87	146
all	all	1.10	1.44	1.66	1.83	2.06	593

Table 3: Differences between our regression methodology and that of Malm and Hand (2007).

Methodological Aspect	**Malm and Hand (2007)**	**This work**
IMPROVE dataset	Download date:3 Dec. 2004 Years analyzed: 1988-2003	Download date:6 Jan 2010 Years analyzed: 2002-2008
Data segregated by	Monitoring site	Monitoring site for β_{soil} Monitoring site and quarter for all other coefficients
Regression type	Ordinary least squares	Errors-in-variables
Response variable	$PM_{2.5}$	$PM_{2.5} - (1.2KNON + 1.8Cl^- + EC)$
Intercept (β_0)	Included	Excluded
Explanatory variables	$(NH_4)_2SO_4$, NH_4NO_3, OC, EC, soil, sea salt*	$(NH_4)_2SO_4$, NH_4NO_3, OC, soil
Calculation of explanatory variables	$(NH_4)_2SO_4 = 1.37\times SO_4^{2-}$ (SO_4^{2-} measured by ion chromatography)	$(NH_4)_2SO_4 = 4.125\times S$ (S measured by XRF)
	SOIL from Eq. (2)	SOIL from Eq. (7)

*Note: Malm and Hand (2007) did not use sea salt as an explanatory variable at sites with very few available Cl^- concentrations: ADPI1, AGTI1, AREN1, BALD1, BOAP1, BRLA1, CACR1, CADI1, CAPI1, CEBL1, CHER1, CHOI1, COHU1, CRES1, CRMO1, DEVA1, DOME1, ELDO1, ELLI1, FOPE1, GAMO1, GRGU1, HALE1, HEGL1, HOOV1, IKBA1, JARI1, JOSH1, LASU1, LIGO1, LIVO1, LOST1, MACA1, MELA1, MING1, MKGO1, MOM1, MONT1, NEBR1, NOCH1, PMRF1, QUCI1, QURE1, QUVA1, SAFO1, SAGA1, SAGU1, SAMA1, SAPE1, SENE1, SHRO1, SIKE1, SIPS1, SPOK1, SWAN1, TALL1, THBA1, THRO1, ULBE1, WHRI1, WICA1, WIMO1, ZION1



* Table S6 includes 10 regressions flagged for an outlier year, yet deemed as high confidence and included in Table S5.
\+ Table S7 includes 4 regressions yielding physically unrealistic coefficients and low-confidence temporal trends, thus also appearing in Table S6.

Figure 1: Flow diagram outlining regression methodology used in this work. Some results appear in multiple tables as indicated by the footnotes.

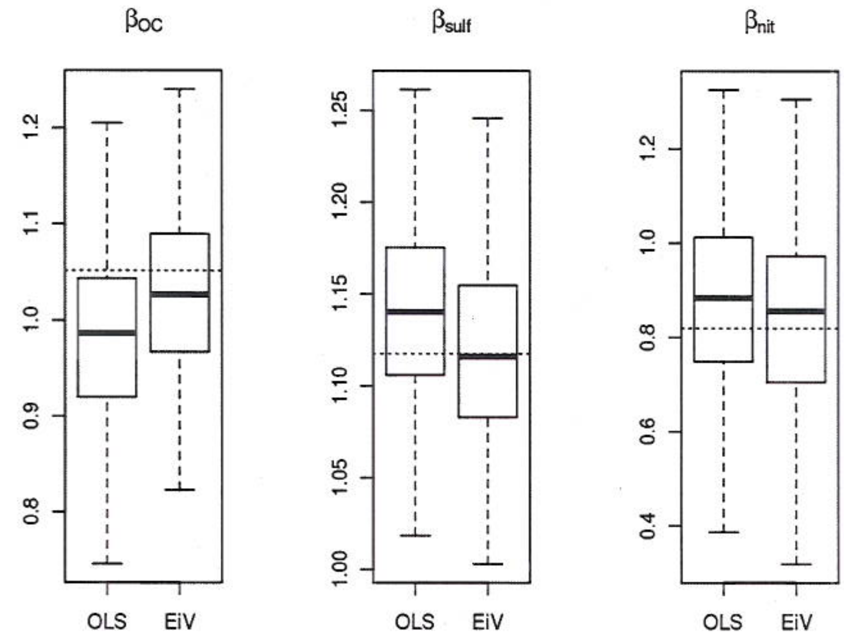


Figure 2: Bias in regression coefficients caused by measurement error in synthetic datasets representative of Gila Wilderness, NM in quarter 1. Horizontal dotted lines represent the "true" value of each coefficient. The left box in each panel illustrates bias for OLS regressions and the right box shows a greatly reduced bias after implementing the errors-in-variables (EiV) regression method.

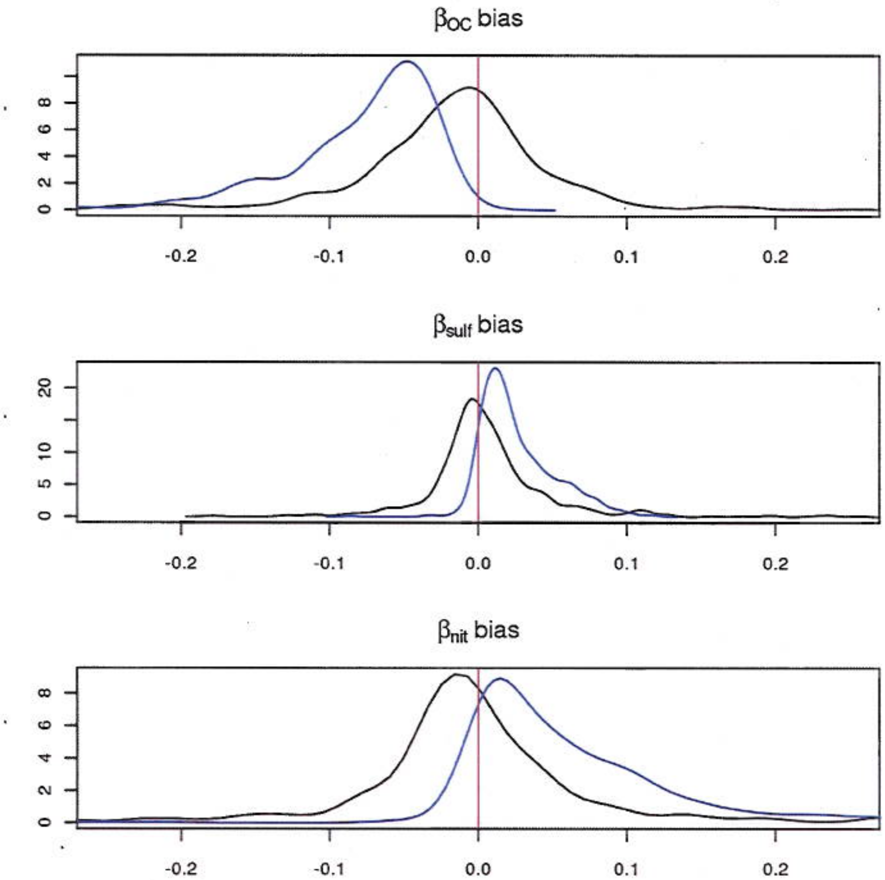


Figure 3: Distribution of bias in regression coefficients for quarter-specific regressions at all IMPROVE sites. For each technique, we compute the median bias from 200 synthetic datasets at each site/quarter using ordinary least squares (blue) and errors-in-variables regression (black) and plot the distribution of those median values across all 612 site- and quarter-specific regressions. The red vertical line shows zero bias.

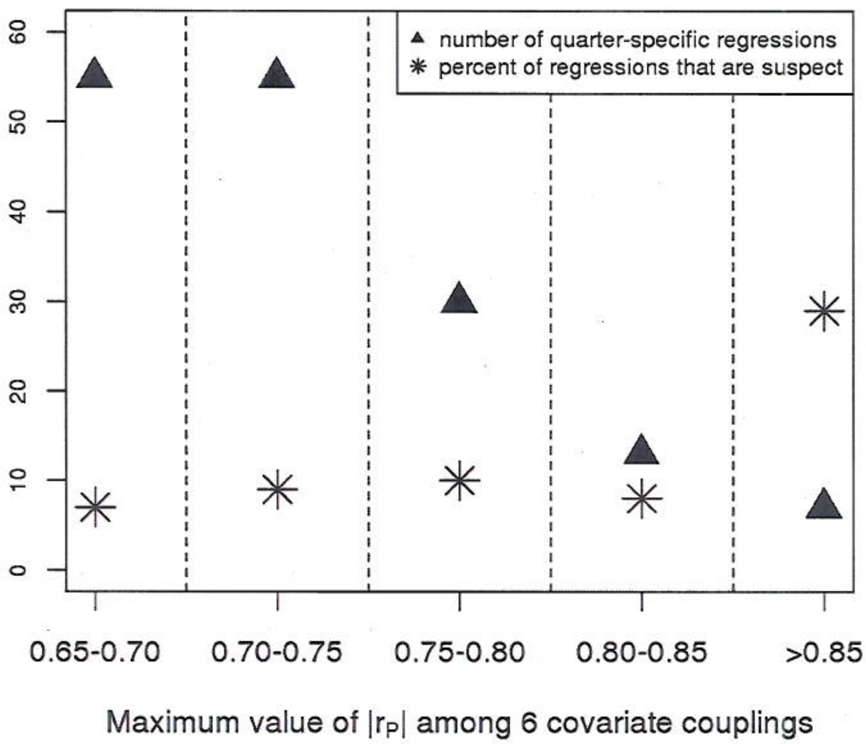


Figure 4: Empirical selection of the 0.85 threshold $|r_p|$ value for identifying site- and quarter-specific regressions which may be biased due to multicolinearity. See Sect. 2.4.1 for an explanation of what constitutes a regression that is "suspect." The 452 EiV regressions with max $|r_P| < 0.65$ were not examined when determining this empirical threshold.

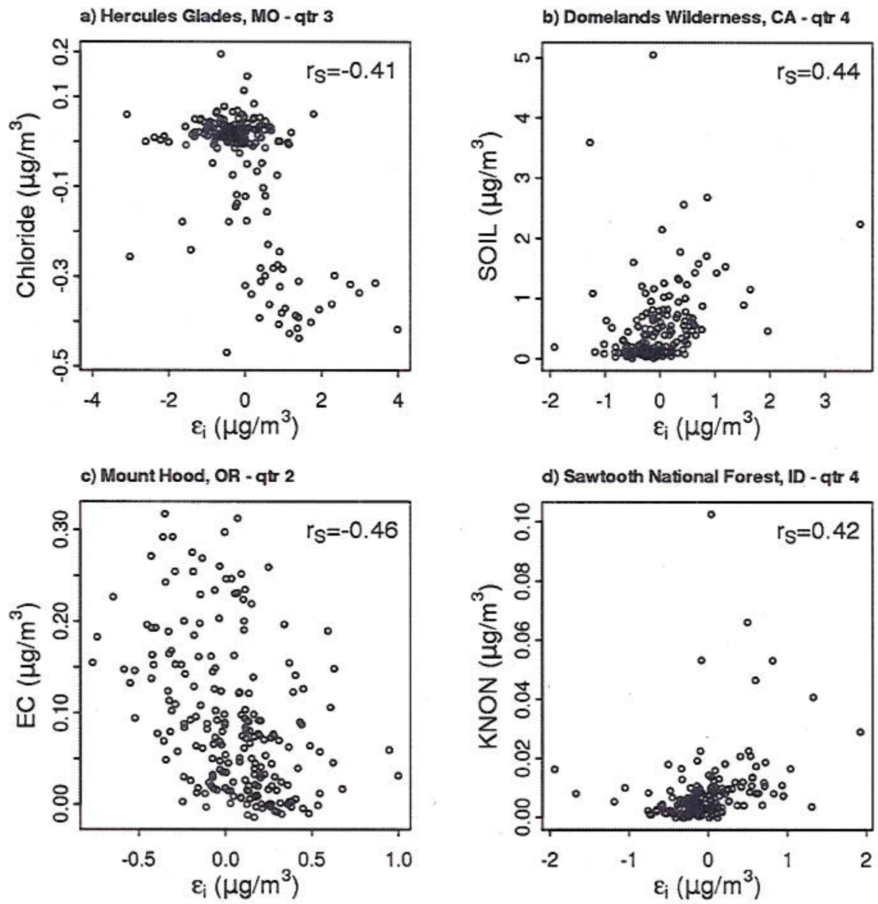


Figure 5. Example datasets in which residual error (ε_i) exhibits a strong correlation with a $PM_{2.5}$ component, indicating that Eq. (5) is an enreliable represe3ntation of $PM_{2.5}$ composition at these sites during these quarters. Twelve regressions are eliminated because $\max|r_S| > 0.4$, inclusing examples shown here. See Sect 2.4.2 for a discussion of the negative Cl^- values in (a).

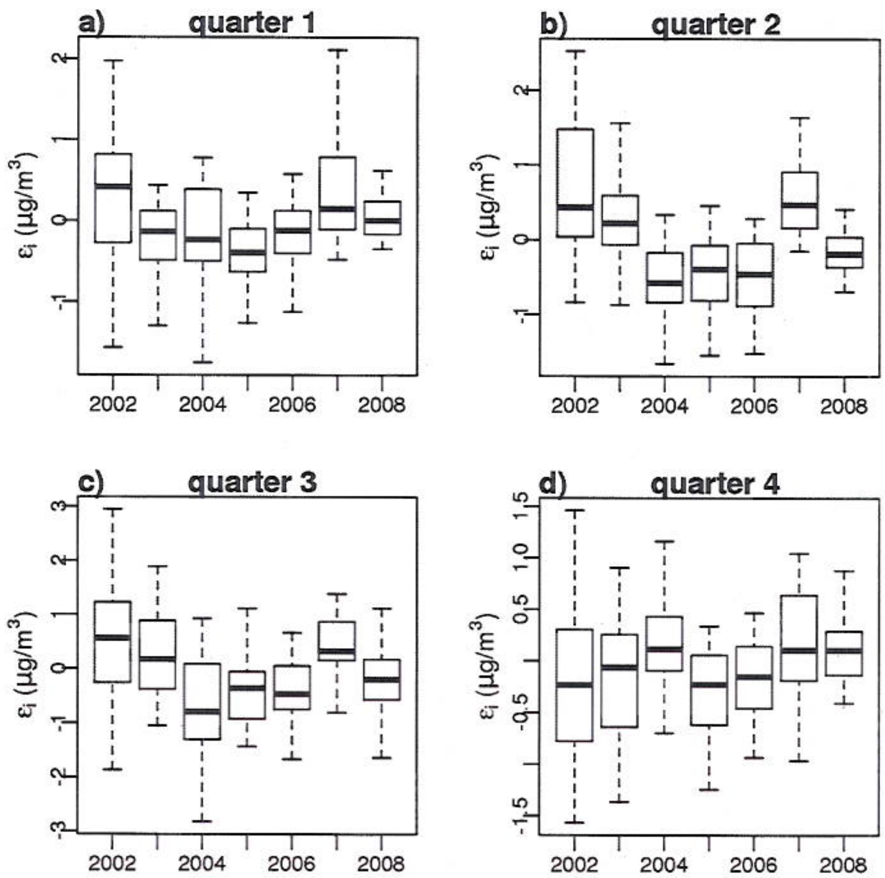


Figure 6. Lack of systematic change in residual error values (ε_i) between 2004 and 2005 at the Sipsy Wilderness in Alabama, a site with one of the highest OC concentrations. The analogous plots from other sites were also inspected, but no abrupt change in ε_i was found.

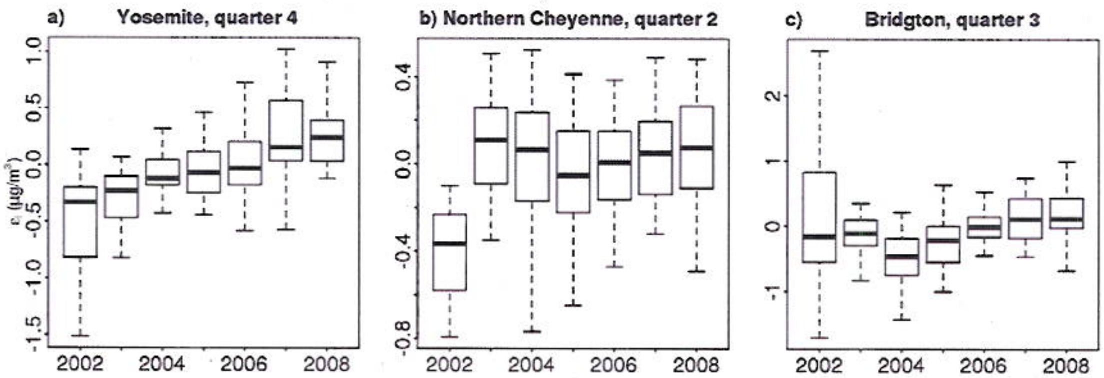


Figure 7. (a) Residual error values (ε_i) from quarter 4 at Yosemite National Park show a monotonically increasing trend between 2002 and 2008 . (b) In the quarter 2 regression of Northern Cheyenne data, the inter-quartile range of ε_i in 2002 does not overlap with other years. (c) There is a substantially larger spread in ε_i during quarter 3 at Bridgton, Maine in 2002 than in all other years.

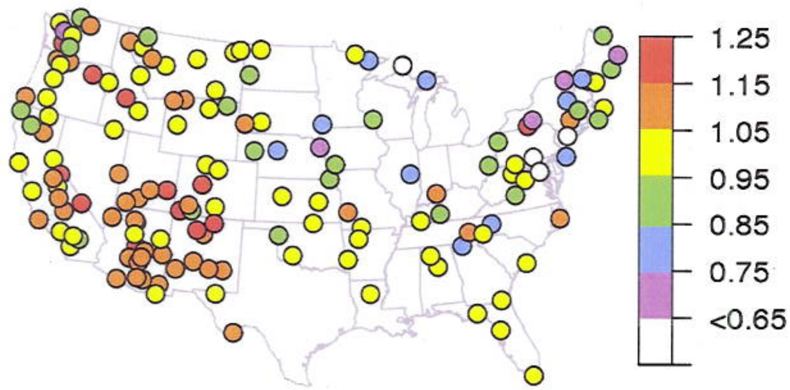
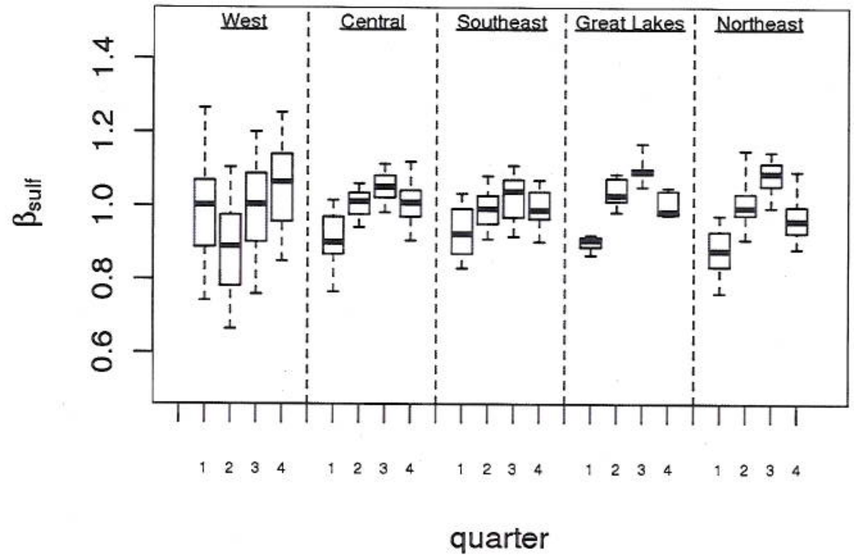


Figure 8. β_{soil} at 153 IMPROVE sites



1

2 Figure 9. Spatial and temporal trends in β_{sulf}.

3

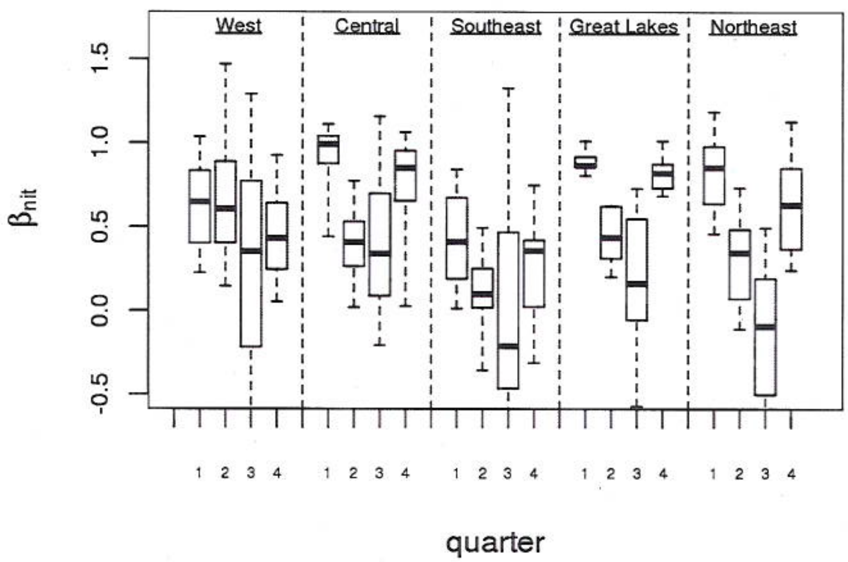


Figure 10. Spatial and temporal trends in β_{nit}.

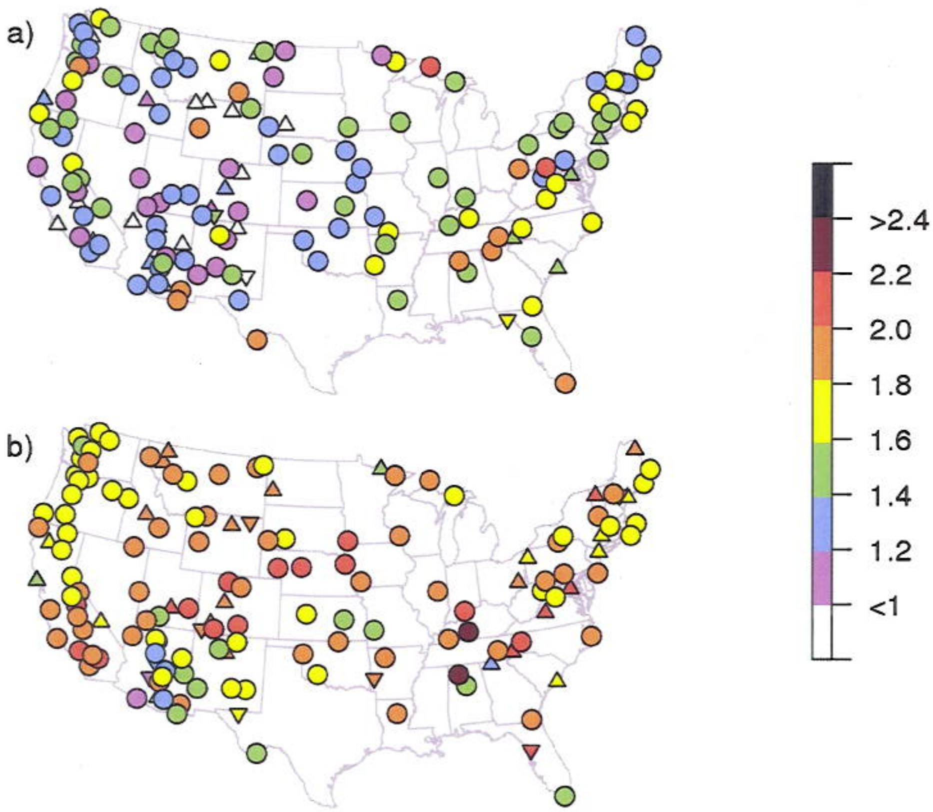


Figure 11. β_{OC} values for quarter 1 (top) and quarter 3 (bottom). High confidence results are depicted by circles, regressions with questionable residual trends are depicted by downward facing triangles, and regressions with a questionable coefficient are depicted by upward facing triangles.

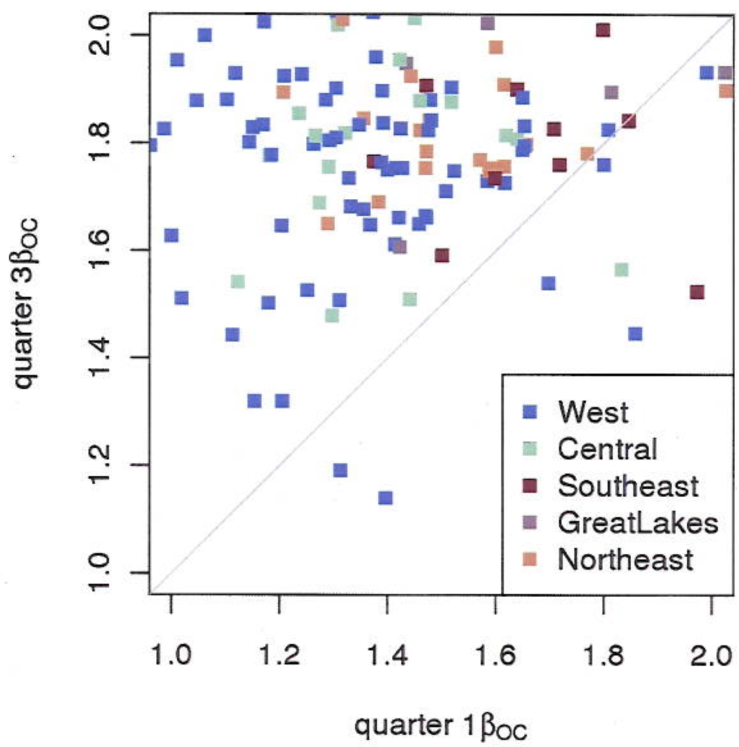


Figure 12. Comparison of β_{OC} values for quarters 1 and 3.

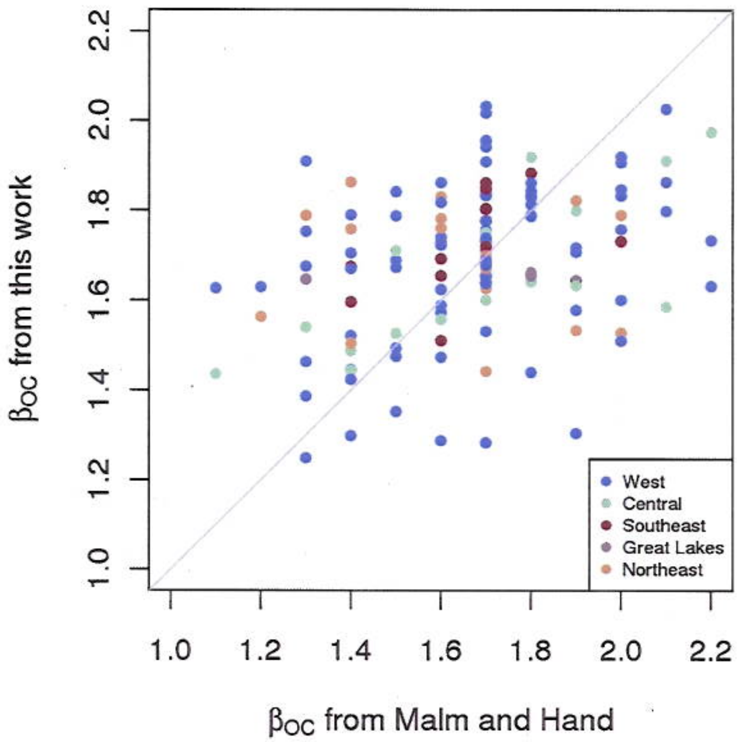


Fig 13. Comparisons of β_{OC} values reported by Malm and Hand (2007) to multiyear β_{OC} values from this work.

Supplementary Information for:

Determining the spatial and seasonal variability in OM/OC ratios across the U.S. using multiple regression

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

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

Table of contents:

S1. Methodology

S1.1. Calculating the variance of regression coefficients

To find the estimated variance associated with the regression coefficients in equation (11), we need to make some additional calculations. This discussion, like that in Sec. 2.3, is based entirely on the work of Fuller (1987) (Sec 3.1.2), conforming to his original notation as much as is feasible. We begin by defining the matrix $\hat{M}_{z\pi z}$ as

$$\hat{M}_{z\pi z} = n^{-1}\sum_{t=1}^{n}[\tilde{\sigma}_{vvtt}^{-1}(Z_t' Z_t - \Sigma_{aatt})] \qquad \text{(S1)}$$

We are most interested in the lower right submatrix of $\hat{M}_{z\pi z}$; i.e., the submatrix which remains when the first row and first column of $\hat{M}_{z\pi z}$ are removed. We call this $k \times k$ submatrix $\hat{M}_{x\pi x}$, where k is the number of explanatory variables in the regression model.

The estimated covariance matrix associated with our regression coefficients is given by

$$\hat{V}(\hat{\beta}) = n^{-2}\hat{M}_{x\pi x}^{-1}\{\sum_{t=1}^{n}[\tilde{\sigma}_{vvtt}^{-1}(X_t' X_t + \tilde{\sigma}_{vvtt}^{-1}\Sigma_{uutt}\tilde{\beta}\tilde{\beta}'\Sigma_{uutt})]\}\hat{M}_{x\pi x}^{-1} \qquad \text{(S2)}$$

As mentioned in Sec 2.3, the diagonal elements of the matrix given by $\hat{V}(\hat{\beta})$ are the estimated variances associated with each of the regression coefficients (each of the elements of $\hat{\beta}$). The square roots of these variances are referred to as the estimated standard errors for the regression coefficients.

S1.2. Sample R Code

The following R code can be used to calculate regression coefficients

```
# load functions necessary for these calculations
source("func_for_beta_est.r")

# calculate response variable
data$response <- data$PM25_Value - (data$EC_Value + 1.2*data$knon_Value +
1.8*data$Cl_Value)

# Set up a data frame with response variable and covariates.  Each entry in the data
# frame (measured sample) includes values and reported uncertainties for the PM
# components
regdata <- data.frame(y = data$response, sulfate = data$ammsulfate, nitrate <-
       data$ammnitrate_Value, OC = data$OC_Value, soil = data$soil_Value, y_Unc =
       data$response_Unc, sulfate_Unc = data$ammsulfate_Unc, nitrate_Unc =
       data$ammnitrate_Unc, OC_Unc = data$OC_Unc, soil_Unc = data$soil_Unc)

# Create a data frame.  Each row contains the name of the covariate value in the
# first column and the name of the column containing the uncertainty values for that
# variable in the second column.
names.covariates.columns.df <- data.frame(value=c("sulfate", "nitrate", "OC", "soil"),
sd=c("sulfate_Unc", "nitrate_Unc", "OC_Unc", "soil_Unc"), stringsAsFactors=F)
       num.covariates = nrow(names.covariates.columns.df)

# Create a data frame containing just one row.  The first column holds the name of
# the column for the response variable value.  The second column holds the name
# of the column for the error associated with the response.
names.response.columns.df <- data.frame(value="y", sd="y_Unc", stringsAsFactors=F)

# obtain preliminary estimate for betas
prelim.beta.est <- find.prelim.beta.est(regdata, names.response.columns.df,
       names.covariates.columns.df)

# Calculate var.qq given preliminary estimate
var.qq <- find.var.qq(regdata, names.response.columns.df, names.covariates.columns.df,
       prelim.beta.est)

# calculate new beta est, G, M.zpiz
beta.est.etc <- find.beta.est.etc(regdata, names.response.columns.df,
       names.covariates.columns.df, var.qq, prelim.beta.est)
beta.est <- beta.est.etc$beta.est
M.zpiz <- beta.est.etc$M.zpiz
var.beta.est <- find.beta.est.var(regdata, names.response.columns.df,
       names.covariates.columns.df, var.qq, M.zpiz)
stdev.beta.est <- sqrt(diag(var.beta.est))
var.beta.est <- find.beta.est.var(regdata, names.response.columns.df,
       names.covariates.columns.df, var.qq, M.zpiz)
stdev.beta.est <- sqrt(diag(var.beta.est))

sulfate_coeff <- beta.est[1]
nitrate_coeff <- beta.est[2]
oc_coeff <- beta.est[3]
soil_coeff <- beta.est[4]
sulfate_stdev <- stdev.beta.est[1]
nitrate_stdev <- stdev.beta.est[2]
oc_stdev <- stdev.beta.est[3]
soil_stdev <- stdev.beta.est[4]
```

Below is the text from a file that defines the functions needed to estimate the regression coefficients and standard deviations: func_for_beta_est.r.

```
##### find.prelim.beta.est function #####
find.prelim.beta.est <- function(data.df, names.response.columns.df,
       names.covariates.columns.df){
```

```
  # The number of observations is equal to the number of rows of data.df
  num.obs <- nrow(data.df)

  # The number of covariates is equal to the number of rows of
  # names.covariates.columns.df.
  num.covariates <- nrow(names.covariates.columns.df)


   # Initialize at 0.
   M.xx <- matrix(0.0, nrow=num.covariates, ncol=num.covariates)
   M.xy <- rep(0.0, num.covariates)

  for (j in 1:num.obs){

      # Calculate beta estimate.
      X.j <- as.vector(as.matrix(data.df[j, names.covariates.columns.df$value]))
      Y.j <- data.df[j, names.response.columns.df$value]

      M.xy <- M.xy + (X.j * Y.j)

      # Covariance matrix of measurement standard deviations among covariates.
      cov.uu <- diag(as.vector(as.matrix(data.df[j, names.covariates.columns.df$sd]^2)),
      ncol=num.covariates, nrow=num.covariates)
      M.xx <-  M.xx + ( X.j %*% t(X.j)  - cov.uu )
   }

  M.xy <- M.xy / num.obs
  M.xx <- M.xx / num.obs

  return(as.vector(solve(M.xx) %*% M.xy))
}

##### find.var.qq function #####
find.var.qq <- function(data.df, names.response.columns.df, names.covariates.columns.df,
prelim.beta.est){

  # The number of observations is equal to the number of rows of
  # data.df
  num.obs <- nrow(data.df)

  # The number of covariates is equal to the number of rows of
  # names.covariates.columns.df.
  num.covariates <- nrow(names.covariates.columns.df)

  # Initialize at 0.
  sig.qq <- 0.0
  A <- matrix(0, num.covariates+1, num.covariates+1)
  M <- matrix(0, num.covariates+1, num.covariates+1)

  # Loop through the observations, adding a contribution from each to sig.qq.
  for (j in 1:num.obs){

      # Identify response, covariates, and combined error matrix for
      # observation j.
      X.j <- as.vector(as.matrix(data.df[j, names.covariates.columns.df$value]))
      Y.j <- data.df[j, names.response.columns.df$value]
      cov.aa <- diag(as.vector(as.matrix(data.df[j, c(names.response.columns.df$sd,
             names.covariates.columns.df$sd)]^2)), ncol=num.covariates+1,
             nrow=num.covariates+1)

      ## Estimate var.qq.
      first.part <- ( (Y.j - (t(X.j) %*% prelim.beta.est))^2 ) / (num.obs -
num.covariates)
      one.and.neg.beta <- c(1.0, -prelim.beta.est)
      second.part <- ( t(one.and.neg.beta) %*% cov.aa %*% one.and.neg.beta ) / num.obs

      sig.qq <- sig.qq + (first.part - second.part)

      ## Calculate generalized eigenvalues.
      A <- A + ( c(Y.j, X.j) %*% t(c(Y.j, X.j)) )
      M <- M + cov.aa
```

```
    }


  # Find the minimum of the generalized eigenvalues det(A - lamda M)
  # = 0.  Since our M is diagonal, we can simplify this to finding
  # the eigenvalues (in the standard fashion) of inv(M) %*% A.  We
  # know that these eigenvalues must be real, so any small imaginary
  # parts are numerical artifacts.
  lambda <- min ( Re( eigen(solve(M) %*% A)$values ) )


  # If lambda is smaller than one, then sig.qq should be 0,
  # instead of the value we calculated in the loop.
  if (lambda < 1)
        return(0)
  else
        return(as.vector(sig.qq))
  }

##### find.var.vv.for.indiv.obs function #####
# Assumes no correlation among covariate measurement errors and no
# correlation between covariate response measurement errors.
find.var.vv.for.indiv.obs <- function(var.qq, response.sd, covariates.sd,
prelim.beta.est){

  var.ww <- response.sd^2
  cov.uu <- diag(covariates.sd^2, ncol=num.covariates, nrow=num.covariates)

  return( as.vector(var.qq + var.ww + ( t(prelim.beta.est) %*% cov.uu %*%
       prelim.beta.est ) ) )
}

##### find.beta.est.etc function #####
find.beta.est.etc <- function(data.df, names.response.columns.df,
       names.covariates.columns.df, var.qq, prelim.beta.est){

  # The number of observations is equal to the number of rows of data.df
  num.obs <- nrow(data.df)

  # The number of covariates is equal to the number of rows of
  # names.covariates.columns.df.
  num.covariates <- nrow(names.covariates.columns.df)

  # Initialize to 0.
  G <- matrix(0.0, nrow=num.covariates, ncol=num.covariates)
  mult1 <- matrix(0.0, nrow=num.covariates, ncol=num.covariates)
  mult2 <- rep(0.0, num.covariates)
  M.zpiz <- matrix(0.0, nrow=num.covariates+1, ncol=num.covariates+1)


  for (j in 1:num.obs){

    # Find var.vv for this observation.
    var.vv <- find.var.vv.for.indiv.obs(var.qq, response.sd=as.vector(data.df[j,
           names.response.columns.df$sd]), covariates.sd=as.vector(as.matrix(data.df[j,
           names.covariates.columns.df$sd])), prelim.beta.est)

    cov.uu <- diag(as.vector(as.matrix(data.df[j, names.covariates.columns.df$sd]^2)),
           ncol=num.covariates, nrow=num.covariates)
    cov.uv <- -cov.uu %*% prelim.beta.est


    # Now, we have enough info to get G.
    X.j <- as.vector(as.matrix(data.df[j, names.covariates.columns.df$value]))
    G <- G + ( ( (X.j %*% t(X.j) ) * var.vv) + (cov.uv %*% t(cov.uv)) )

    # Find final beta estimate.

    # Find Y.j
    Y.j <- data.df[j, names.response.columns.df$value]
```

```
      # Combine with var.vv, X.j, and cov.uu.
      # First multiplier.
      mult1 <- mult1 + ( ( ( X.j %*% t(X.j) ) - cov.uu ) / var.vv )
      mult2 <- mult2 + ( (X.j * Y.j) / var.vv )

      # Find M.zpiz.

      Z.j <- c(Y.j, X.j)
      cov.aa <- diag(as.vector(as.matrix(data.df[j, c(names.response.columns.df$sd,
             names.covariates.columns.df$sd)]^2)), ncol=num.covariates+1,
             nrow=num.covariates+1)
      M.zpiz <- M.zpiz + ( ( ( Z.j %*% t(Z.j) ) - cov.aa ) / var.vv )

    }

    G <- G / num.obs
    beta.est <- solve(mult1) %*% mult2
    M.zpiz <- M.zpiz / num.obs

    return(list(beta.est=as.vector(beta.est), G=G, M.zpiz=M.zpiz))
  }



  ##### find.beta.est.var function #####
  find.beta.est.var <- function(data.df, names.response.columns.df,
         names.covariates.columns.df, var.qq, M.zpiz){

    # The number of observations is equal to the number of rows of
    # data.df
    num.obs <- nrow(data.df)

    # The number of covariates is equal to the number of rows of
    # names.covariates.columns.df.
    num.covariates <- nrow(names.covariates.columns.df)

    # Initialize to 0.
    mid.part <- matrix(0.0, nrow=num.covariates, ncol=num.covariates)

    for (j in 1:num.obs){

      cov.uu <- diag(as.vector(as.matrix(data.df[j, names.covariates.columns.df$sd]^2)),
             ncol=num.covariates, nrow=num.covariates)
      cov.uv <- -cov.uu %*% prelim.beta.est

      X.j <- as.vector(as.matrix(data.df[j, names.covariates.columns.df$value]))

      var.vv <- find.var.vv.for.indiv.obs(var.qq, response.sd=as.vector(data.df[j,
             names.response.columns.df$sd]), covariates.sd=as.vector(as.matrix(data.df[j,
             names.covariates.columns.df$sd])), prelim.beta.est)

      mid.part <- mid.part + ( ( (X.j %*% t(X.j)) + ((cov.uv %*% t(cov.uv))/var.vv) ) /
             var.vv )
    }

    M.xpix <- M.zpiz[-1, -1]

    return( (1.0/(num.obs^2)) * (solve(M.xpix) %*% mid.part %*% solve(M.xpix)) )
  }
```

S2. Sulfate coefficient analysis

Table S1 shows how β_{sulf} should change with laboratory RH and degree of sulfate neutralization (DSN). The DSN is calculated assuming that all nitrate is in the form of ammonium nitrate and that any ammonium not bound to nitrate is bound to sulfate (Pinder et al., 2008). A DSN of 2 means that two moles of ammonium are available to bond with every mole of sulfate, indicating fully neutralized ammonium sulfate.

$$DSN = \frac{NH_4(moles) - NO_3(moles)}{SO_4(moles)} \quad (S3)$$

We used the AIM model (Wexler and Clegg, 2002) to estimate total water mass associated with sulfate aerosols for both the dry hysteresis branch and for supersaturated aerosols. Ammoniated sulfate switches from wet to dry at its efflorescence RH (Colberg et al., 2003).

Table S1. Estimated β_{sulf} values based on laboratory RH and DSN for dry (red) and wet (blue) particles.

		Degree of sulfate neutralization (DSN)										
		1	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**	**2**
Laboratory relative humidity (RH)	**30%**	1.03	1.00	0.99	1.01	1.04	0.94	0.94	0.96	0.97	0.99	1.00
	31%	1.04	1.01	1.00	1.01	1.05	0.94	0.94	0.96	0.97	0.99	1.00
	32%	1.05	1.02	1.01	1.02	1.05	0.94	0.94	0.96	0.97	0.99	1.00
	33%	1.06	1.03	1.02	1.03	1.06	0.94	0.94	0.96	0.97	0.99	1.00
	34%	1.07	1.04	1.03	1.04	1.07	1.12	0.94	0.96	0.97	0.99	1.00
	35%	1.08	1.05	1.04	1.05	1.08	1.13	0.94	0.96	0.97	0.99	1.00
	36%	1.09	1.06	1.05	1.06	1.09	1.13	1.18	0.96	0.97	0.99	1.00
	37%	1.11	1.07	1.06	1.06	1.10	1.14	1.19	0.96	0.97	0.99	1.00
	38%	1.12	1.08	1.07	1.07	1.11	1.15	1.20	0.96	0.97	0.99	1.00
	39%	1.13	1.09	1.08	1.08	1.12	1.16	1.21	0.96	0.97	0.99	1.00
	40%	1.14	1.11	1.09	1.10	1.13	1.18	1.22	1.26	0.97	0.99	1.00
	41%	1.15	1.12	1.10	1.11	1.14	1.19	1.23	1.27	1.30	0.99	1.00
	42%	1.17	1.13	1.11	1.12	1.15	1.20	1.24	1.28	1.31	1.34	1.00

Figure S1 shows seasonal variation in the laboratory RH where the filters were weighed, based on several years of data. Though the variation is modest, laboratory RH values are slightly higher during quarter 3 and lowest during quarter 4. Ninety percent of the samples were weighed between 5 and 29 days after sampling, so samples are generally weighed during the same time of year as they are sampled.

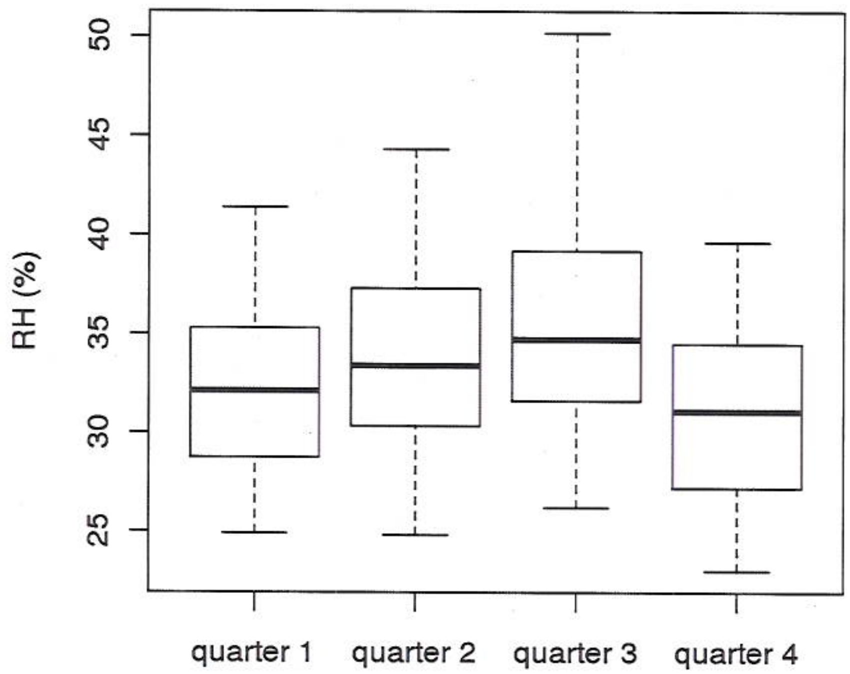


Figure S1. Seasonal RH variation in the IMPROVE gravimetric measurement laboratory

To determine if the seasonal variation in our β_{sulf} estimates is reasonable, we examined measurements collected between 1999 and 2007 across the CSN where ammonium concentrations are routinely measured along with sulfate and nitrate (downloaded September 24, 2009 from http://www.epa.gov/cgi-bin/htmSQL/mxplorer/query_spe.hsql). These calculations show that DSN does indeed vary seasonally in the southeast, great lakes and northeast regions, with less seasonal variation in the central and western regions. The seasonal variations in DSN are consistent with measurements reported from the Pittsburgh supersite which showed that sulfate was fully neutralized in the winter but not in the summer (Khlystov et al., 2005). The calculated DSN values are used to approximate β_{sulf} at CSN sites using Table S1 and assuming laboratory RH values of 35% in q1, 37% in q2, 39% in q3, and 35% in q4 (Fig.

S2b) . Except for the western region, our approximations of CSN β_{sulf} show a seasonal pattern similar to that estimated by our regression analysis of IMPROVE data (copied from Fig. 9 to Fig. S2a to facilitate comparison) with both having higher values in the summer and lower values in the winter. This analysis suggests that the trends predicted by this regression analysis are reasonably explained by known physical phenomena.

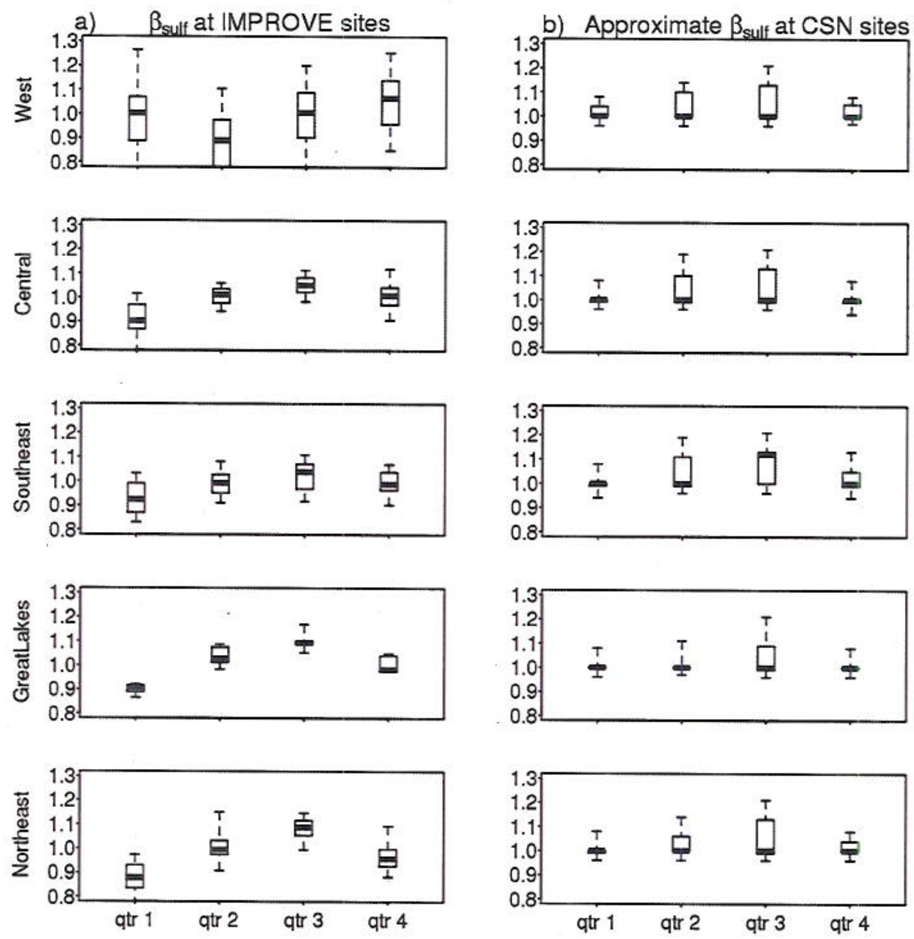


Figure S2. (a) Seasonal and temporal trends in β_{sulf} from regression of IMPROVE data. (b) Calculated β_{sulf} values based on CSN measurements of NH_4^+, SO_4^{2+}, and NO_3^-, and RH in the IMPROVE gravimetric measurement laboratory.

S3. Sensitivity of β_{OC} to the inclusion of β_{EC} and assumptions about OC artifacts

The component of $PM_{2.5}$ reported as EC may not be purely graphitic and therefore may have some non-carbon mass associated with it. In that case, the EC coefficient in Eq. (5) could be greater than one. Also, there is some uncertainty in the measurement method used to split total carbon (TC) into EC and OC which could lead to either a positive or negative EC artifact. Average EC/TC values have been reported to shift by around 15% due to changes in measurement equipment (White, 2007). For these reasons, we investigate the net effect of assuming an EC coefficient of 1. We perform 10 sets of site- and quarter-specific EiV regressions in which we fix the coefficient for EC at various values (0. 0.25, 0.5, 0.75, 1.25, 1.5, 1.75, 2, 2.5, and 3). This analysis shows that when the EC coefficient is fixed between 0.25 and 1.75, most β_{OC} values change by less than 0.2 (see Fig. S3). When the EC coefficient is changed to 0 or 2, β_{OC} is affected substantially.

To explore this further, we repeat all of the site- and quarter-specific regressions using both EC and OC as explanatory variables (Eq. S4).

$$\begin{aligned} PM_{2.5,i} &= \beta_{OC}OC_i + \beta_{sulf}(NH_4)_2SO_{4,i} + \beta_{nit}NH_4NO_{3,i} + \beta_{soil}SOIL_i \\ &\quad + \beta_{EC}EC_i + 1.8 \times Cl_i^- + 1.2 \times KNON_i + \varepsilon_i \end{aligned} \tag{S4}$$

Twenty five percent of the EC coefficients fall below -0.3 and 50% fall below 0.3. Such low coefficients are unrealistic and can cause substantial overestimates of β_{OC}. The results reported by Hand and Malm (2006) show the opposite effect with most EC coefficients exceeding one. About one quarter of their reported EC coefficients are greater than 3 and one is as high as 11. These EC coefficients appear to be unrealistically high and are likely an artifact of co-linear explanatory variables used in their OLS regression. Again, Fig. S3 demonstrates that high EC coefficients like those from Hand and Malm (2006) would cause drastic underestimates of β_{OC}.

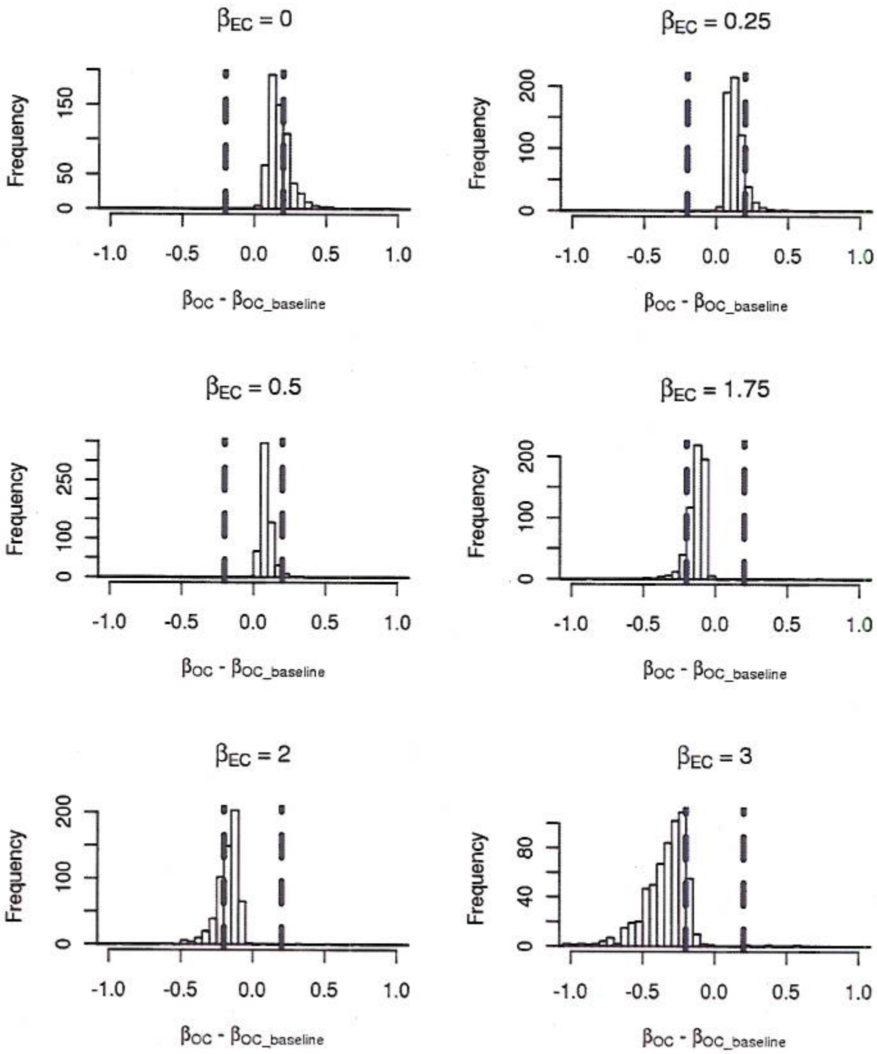


Figure S3. Change in β_{OC} when EC coefficient (β_{EC}) is altered from the baseline value of 1 to other fixed values: 0, 0.25, 0.5, 1.75, 2, and 3. Vertical dashed lines encompass all site- and quarter-specific regressions in which β_{OC} deviates by less than 0.2 from the baseline results presented in Section 3.3.

We conduct a separate analysis to estimate the actual EC coefficients. To accomplish this, we perform a set of regressions in which we use TC instead of OC as our covariate (Eq. S5).

$$\begin{aligned} PM_{2.5,i} &= \beta_{TC} TC_i + \beta_{sulf} (NH_4)_2 SO_{4,i} + \beta_{nit} NH_4 NO_{3,i} + \beta_{soil} SOIL_i \\ &+ 1.8 \times Cl_i^- + 1.2 \times KNON_i + \varepsilon_i \end{aligned} \quad \text{(S5)}$$

We expect the actual coefficient for TC (β_{TC}) to be an intermediate value between our original β_{OC} results and the actual EC coefficient. By applying Eq. (S5), we find that β_{TC} is very close to our original β_{OC} results for most site- and quarter-specific regressions. On average, β_{TC} is slightly lower than β_{OC} (see Fig. S4). Only 3% of the TC coefficients differ from our original β_{OC} values by more than 0.2. It may seem counter-intuitive that results using Eq. (S5) would be so similar to the original regression results, whereas including EC as a separate covariate (as in Eq. (S4)) has a much larger effect.

Assuming a maximum measurement artifact of 15%, we can set a lower bound for β_{EC} around 0.85. From this we can infer that $0.85 < \beta_{EC} < \beta_{TC} < \beta_{OC}$. Over 80% of the estimated TC coefficients from Eq. (S5) fall in the range of 1.2 to 1.9. It follows that the true EC coefficients lie between 0.85 and 1.9. Consequently, EC coefficients in this analysis are much closer to 1 than the EC coefficients estimated by treating EC as a separate explanatory variable. Combining this analysis with the results shown in Fig. S3, we conclude that our assumption of an EC coefficient equal to 1 does not greatly bias our β_{OC} results.

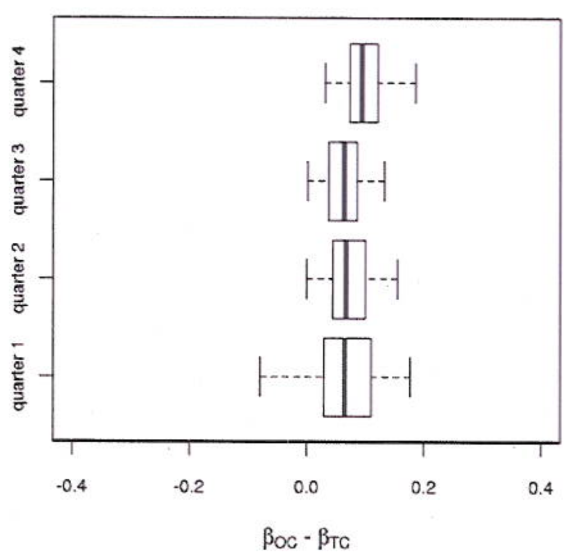


Figure S4: Comparison of our baseline OC coefficients from Eq. (5) to TC coefficients obtained using Eq. (S5).

As mentioned in Sect. 3.3, β_{OC} is influenced by differences in the OC sampling artifacts on quartz versus Teflon filters. Whereas the literature is inconclusive regarding negative artifacts, quartz filters are more prone to positive artifact than Teflon filters. The IMPROVE data include a network-wide and month-specific correction for positive OC artifact on the quartz filter, but no correction for the Teflon filter. Quartz-behind-quartz backup filters are collected at six IMPROVE sites (Chiricahua, Grand Canyon, Mount Rainier, Okefenokee, Shenandoah, and Yosemite). Each month, the median of all quartz-behind-quartz backup filters from these six sites is used as a network-wide average value for positive OC artifact. The reported OC concentrations are calculated by subtracting the median artifact value for that month (μg/filter) from each OC sample at all sites (μg/filter) before converting filter measurements to ambient concentrations of $\mu g/m^3$ (McDade, 2008). Here we evaluate the effect of using a single median artifact at all IMPROVE sites.

Since backup filters are only collected at 6 monitoring sites, it is not possible to determine how much site-to-site variability occurs network-wide. However, we perform a sensitivity study in which we look at site-to-site variability in back-up filter concentrations within the six sites used to create the median OC artifact value. For this analysis, all OC values for these six sites are recalculated using sample-specific backup filter values instead of the network-wide monthly median. We repeat the EiV regression analysis using these new sample-specific-corrected OC values and evaluate changes in β_{OC}. These results are shown in Fig. S5. In all regressions, changes in β_{OC} values are modest, with the average change being 0.05 (3%) and the maximum change being 0.14 (9%). Although it is not known how representative these six sites are of the network as a whole, this analysis suggests that using a single artifact correction network-wide does not substantially affect our estimations of β_{OC}.

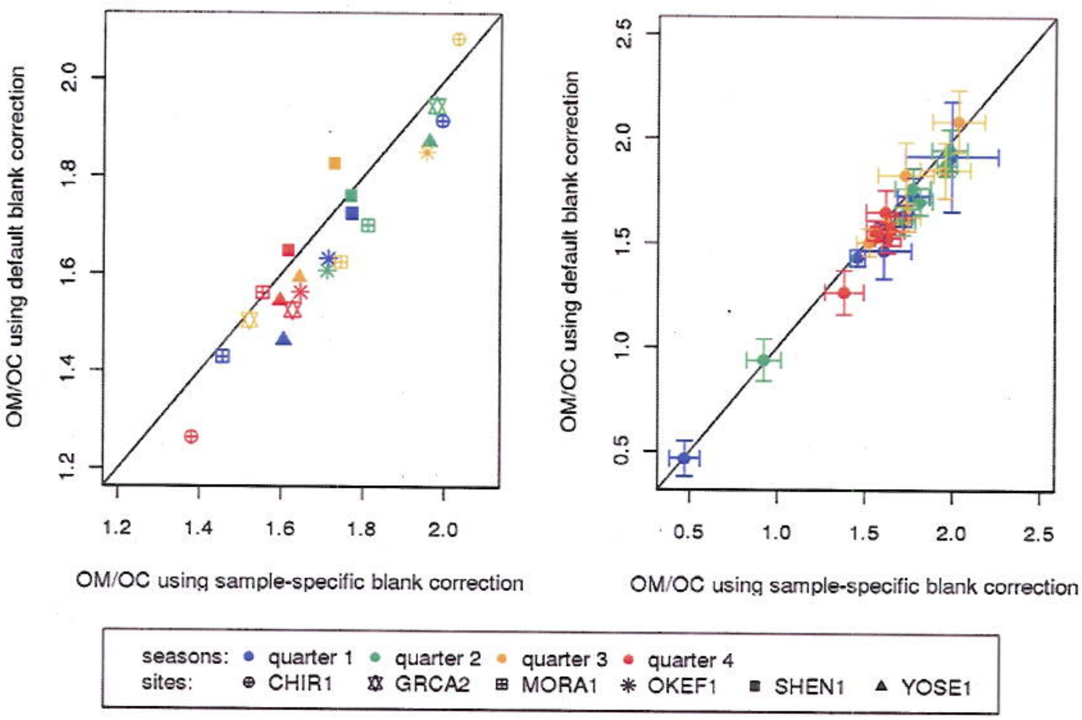
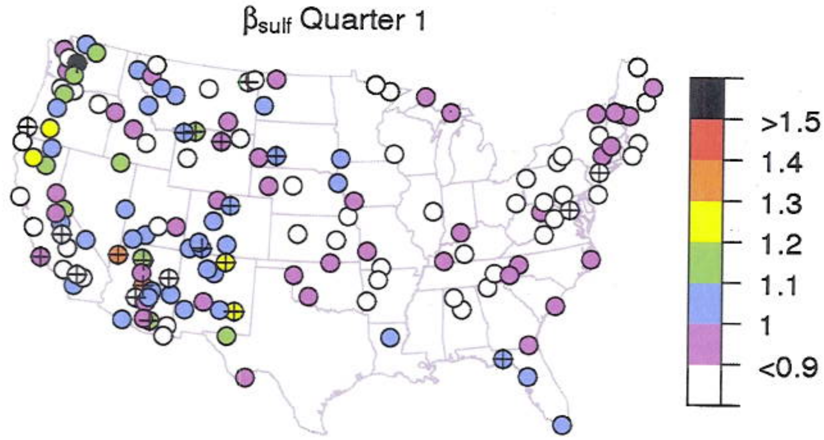
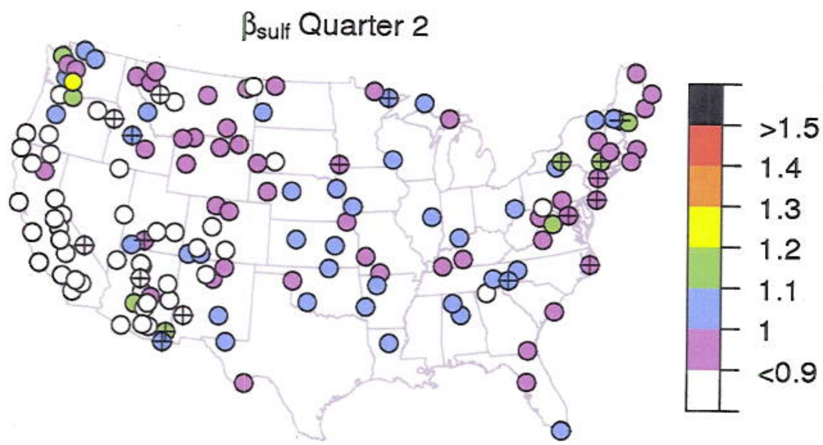
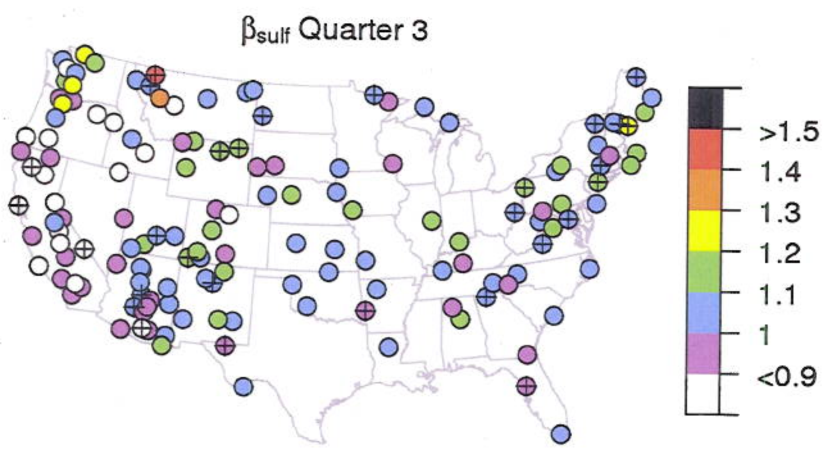


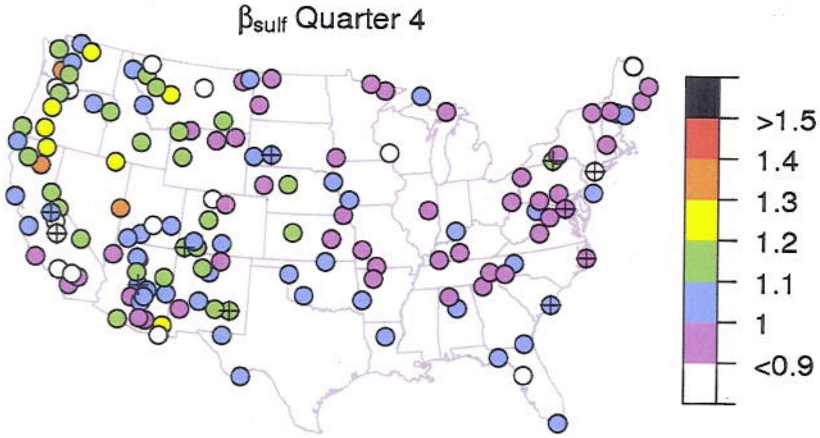
Figure S5. Comparison of β_{OC} values when using default artifact correction versus sample-specific artifact correction for only good regressions (left) and for all quarter-specific regressions (right). Uncertainty bars in the right-hand plot are standard error values for β_{OC} at each site and quarter.

S4. Maps of regression coefficients

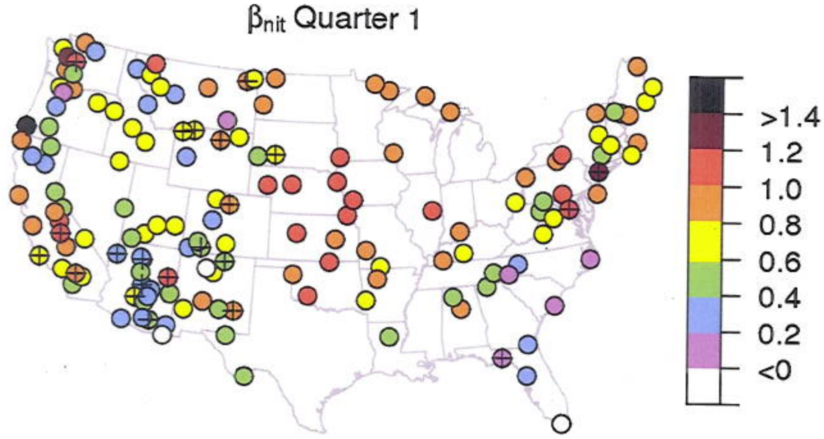


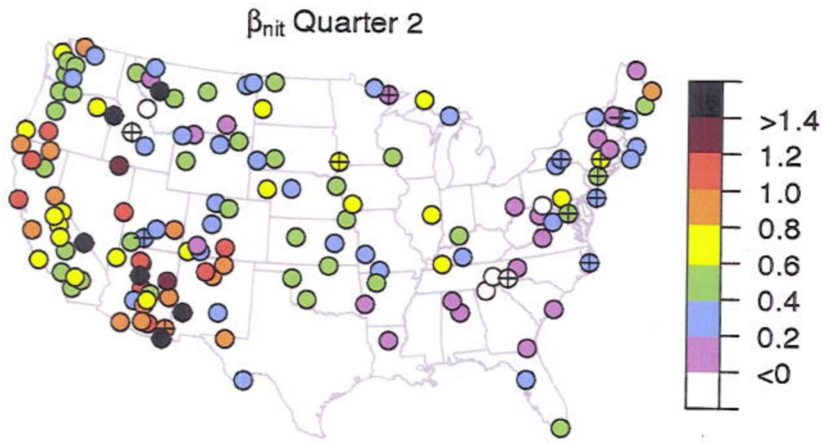


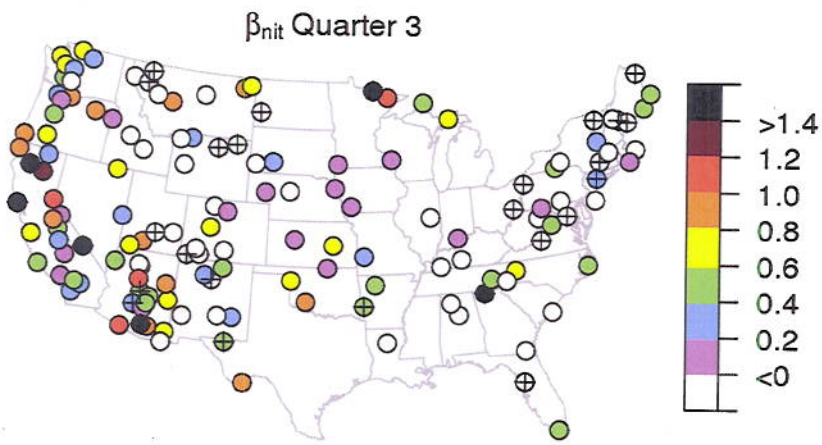


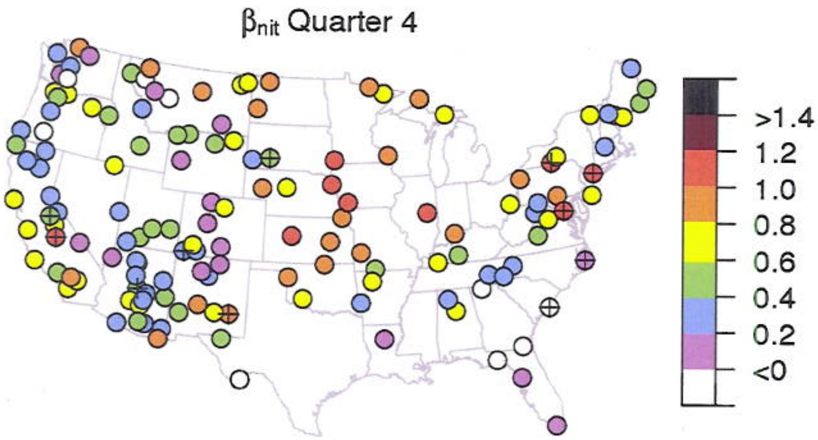


Figures S6-S9: Maps of sulfate coefficients in quarters 1-4. High confidence results are plotted with colored dots. Regressions that were flagged for problematic coefficients or temporal trends in the residual errors are marked with crosses or black dots.

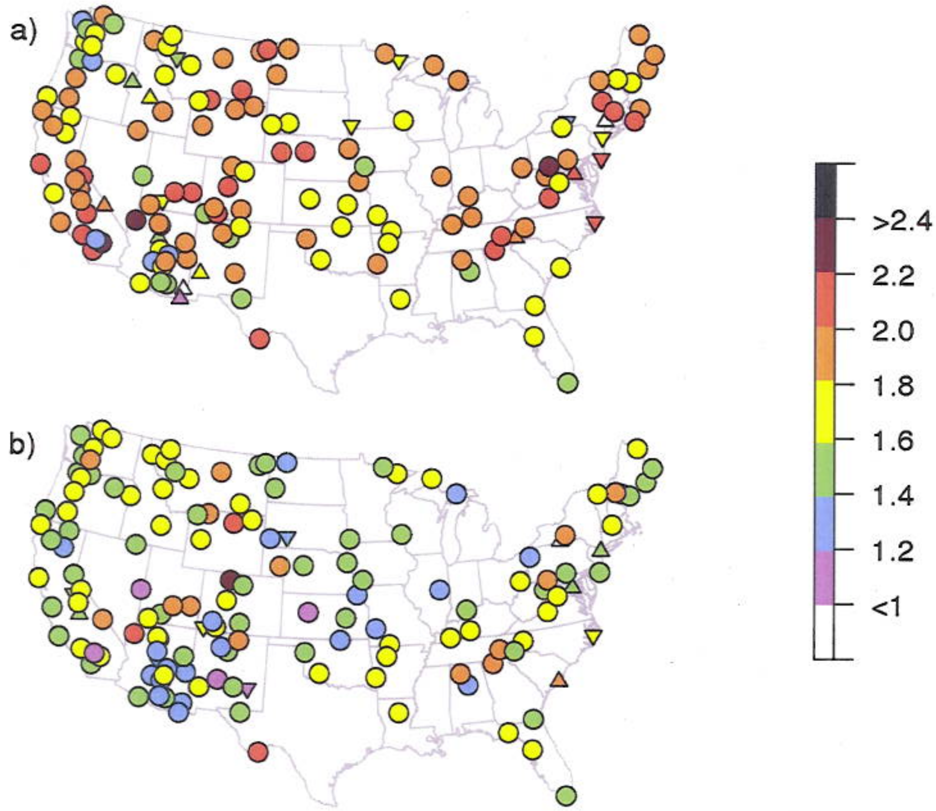

βnit Quarter 1
>1.4
1.2
1.0
0.8
0.6
0.4
0.2
<0


βnit Quarter 2
>1.4
1.2
1.0
0.8
0.6
0.4
0.2
<0


βnit Quarter 3
>1.4
1.2
1.0
0.8
0.6
0.4
0.2
<0



Figures S10-S13: Maps of nitrate coefficients in quarters 1-4. High confidence results are plotted with colored dots. Regressions that were flagged for problematic coefficients or temporal trends in the residuals are marked with crosses or black dots.



Figures S14: β_{OC} values for quarter 2 (top) and quarter 4 (bottom). High confidence results are depicted by circles, regressions with questionable residual trends are depicted by downward facing triangles, and regressions with any physically unreasonable coefficient are depicted by upward facing triangles..

S5. Tabulated regression results

Table S2. Multiyear regression results. Physically unreasonable coefficients are shown in bold.

site	β_{oc}	β_{sulf}	β_{nit}	β_{soil}
Acadia NP	1.82 +/- 0.04	1.05 +/- 0.01	0.13 +/- 0.06	0.95 +/- 0.11
Addison Pinnacle	1.44 +/- 0.07	1.10 +/- 0.02	0.58 +/- 0.04	1.24 +/- 0.18
Agua Tibia	1.74 +/- 0.05	0.97 +/- 0.02	0.52 +/- 0.02	0.98 +/- 0.04
Arendtsville	1.63 +/- 0.06	1.06 +/- 0.02	0.77 +/- 0.02	0.49 +/- 0.13
Badlands NP	1.60 +/- 0.04	0.97 +/- 0.03	0.30 +/- 0.05	0.98 +/- 0.05
Bandelier NM	1.48 +/- 0.04	1.05 +/- 0.03	0.47 +/- 0.04	1.11 +/- 0.02
Big Bend NP	1.92 +/- 0.07	1.00 +/- 0.01	0.32 +/- 0.08	1.07 +/- 0.02
Birmingham	1.39 +/- 0.04	1.08 +/- 0.02	0.64 +/- 0.06	1.02 +/- 0.04
Bliss SP (TRPA)	1.76 +/- 0.02	0.95 +/- 0.03	0.38 +/- 0.05	1.01 +/- 0.03
Blue Mounds	1.70 +/- 0.05	0.93 +/- 0.02	1.05 +/- 0.01	0.79 +/- 0.05
Bondville	1.65 +/- 0.07	1.10 +/- 0.02	0.85 +/- 0.02	0.82 +/- 0.09
Bosque del Apache	1.28 +/- 0.05	0.96 +/- 0.02	0.79 +/- 0.04	1.06 +/- 0.02
Boundary Waters Canoe Area	1.80 +/- 0.04	0.95 +/- 0.02	0.77 +/- 0.02	0.82 +/- 0.10
Bridger Wilderness	1.85 +/- 0.03	0.97 +/- 0.03	0.23 +/- 0.09	1.01 +/- 0.03
Bridgton	1.76 +/- 0.04	1.07 +/- 0.02	-0.05 +/- 0.07	1.09 +/- 0.13
Brigantine NWR	1.79 +/- 0.06	1.03 +/- 0.02	0.52 +/- 0.04	0.79 +/- 0.14
Bryce Canyon NP	1.52 +/- 0.04	1.04 +/- 0.04	0.50 +/- 0.05	1.08 +/- 0.03
Cabinet Mountains	1.76 +/- 0.02	1.02 +/- 0.03	0.17 +/- 0.07	1.11 +/- 0.03
Cadiz	1.68 +/- 0.05	1.04 +/- 0.01	0.71 +/- 0.02	0.96 +/- 0.05
Caney Creek	1.75 +/- 0.04	0.99 +/- 0.01	0.44 +/- 0.03	0.98 +/- 0.02
Canyonlands NP	2.03 +/- 0.05	0.91 +/- 0.03	0.40 +/- 0.04	1.18 +/- 0.02
Cape Cod	1.78 +/- 0.05	1.03 +/- 0.02	0.18 +/- 0.06	0.99 +/- 0.13
Cape Romain NWR	1.69 +/- 0.04	1.02 +/- 0.01	**-0.27 +/- 0.10**	1.00 +/- 0.05
Capitol Reef NP	2.03 +/- 0.05	0.83 +/- 0.04	0.40 +/- 0.04	1.14 +/- 0.02
Casco Bay	1.56 +/- 0.03	1.17 +/- 0.02	0.18 +/- 0.07	1.03 +/- 0.10
Cedar Bluff	1.43 +/- 0.09	1.05 +/- 0.04	0.88 +/- 0.02	1.02 +/- 0.06
Chassahowitzka NWR	1.60 +/- 0.04	1.00 +/- 0.02	0.17 +/- 0.10	1.01 +/- 0.03
Cherokee Nation	1.49 +/- 0.04	1.06 +/- 0.02	0.87 +/- 0.02	0.97 +/- 0.02
Chiricahua NM	1.45 +/- 0.07	1.15 +/- 0.02	0.25 +/- 0.10	1.11 +/- 0.02
Cloud Peak	1.94 +/- 0.03	0.97 +/- 0.03	0.31 +/- 0.08	1.05 +/- 0.03
Cohutta	1.88 +/- 0.06	0.96 +/- 0.01	0.12 +/- 0.05	0.82 +/- 0.05
Columbia Gorge #1	1.57 +/- 0.03	0.85 +/- 0.03	0.57 +/- 0.03	1.13 +/- 0.06
Columbia River Gorge	1.47 +/- 0.03	0.96 +/- 0.04	0.62 +/- 0.02	1.05 +/- 0.02
Connecticut Hill	1.53 +/- 0.08	1.07 +/- 0.02	0.61 +/- 0.04	0.69 +/- 0.15
Crater Lake NP	1.71 +/- 0.02	1.08 +/- 0.03	0.15 +/- 0.10	0.95 +/- 0.03
Craters of the Moon NM	1.85 +/- 0.03	0.88 +/- 0.04	0.50 +/- 0.02	1.04 +/- 0.02

site	β_{oc}	β_{sulf}	β_{nit}	β_{soil}
Crescent Lake	1.97 +/- 0.05	0.92 +/- 0.03	0.92 +/- 0.02	0.86 +/- 0.05
Death Valley NP	1.83 +/- 0.05	0.87 +/- 0.03	0.70 +/- 0.09	1.16 +/- 0.02
Dolly Sods Wilderness	1.51 +/- 0.05	1.06 +/- 0.01	0.28 +/- 0.05	1.03 +/- 0.08
Dome Lands Wilderness	1.79 +/- 0.06	0.77 +/- 0.05	0.70 +/- 0.02	1.12 +/- 0.05
Douglas	1.31 +/- 0.08	1.04 +/- 0.04	0.60 +/- 0.15	1.02 +/- 0.01
El Dorado Springs	1.44 +/- 0.04	1.04 +/- 0.01	0.77 +/- 0.02	1.07 +/- 0.03
Ellis	1.59 +/- 0.05	1.06 +/- 0.02	0.86 +/- 0.02	0.93 +/- 0.03
Everglades NP	1.58 +/- 0.04	1.05 +/- 0.02	0.31 +/- 0.10	1.02 +/- 0.02
Flathead	1.74 +/- 0.02	0.97 +/- 0.03	0.17 +/- 0.06	1.01 +/- 0.04
Fort Peck	1.86 +/- 0.04	0.88 +/- 0.02	0.68 +/- 0.02	0.96 +/- 0.04
Frostberg Reservoir (Big Piney Run)	1.95 +/- 0.06	0.94 +/- 0.01	0.24 +/- 0.04	0.96 +/- 0.08
Gates of the Mountains	1.73 +/- 0.02	1.02 +/- 0.03	0.15 +/- 0.07	1.02 +/- 0.04
Gila Wilderness	1.49 +/- 0.03	0.96 +/- 0.03	0.77 +/- 0.18	1.10 +/- 0.02
Glacier NP	1.72 +/- 0.02	0.88 +/- 0.03	0.79 +/- 0.06	0.91 +/- 0.03
Great Basin NP	1.66 +/- 0.04	1.01 +/- 0.04	0.13 +/- 0.10	1.11 +/- 0.02
Great Gulf Wilderness	1.83 +/- 0.04	1.02 +/- 0.02	0.06 +/- 0.07	0.82 +/- 0.12
Great River Bluffs	1.67 +/- 0.06	0.94 +/- 0.02	0.86 +/- 0.01	0.87 +/- 0.13
Great Sand Dunes NM	1.83 +/- 0.05	0.96 +/- 0.04	0.18 +/- 0.08	1.02 +/- 0.01
Great Smoky Mountains NP	1.86 +/- 0.05	1.05 +/- 0.01	0.18 +/- 0.05	1.06 +/- 0.08
Guadalupe Mountains NP	1.60 +/- 0.09	1.02 +/- 0.03	0.53 +/- 0.05	1.04 +/- 0.01
Hance Camp at Grand Canyon NP	1.54 +/- 0.04	1.15 +/- 0.03	0.57 +/- 0.06	1.11 +/- 0.02
Hells Canyon	1.67 +/- 0.02	0.89 +/- 0.04	0.64 +/- 0.02	1.01 +/- 0.04
Hercules-Glades	1.64 +/- 0.04	0.99 +/- 0.01	0.58 +/- 0.02	1.00 +/- 0.03
Hoover	1.80 +/- 0.03	0.94 +/- 0.04	0.37 +/- 0.08	1.20 +/- 0.03
Ikes Backbone	1.30 +/- 0.05	1.03 +/- 0.04	0.40 +/- 0.05	1.19 +/- 0.02
Indian Gardens	1.71 +/- 0.05	0.99 +/- 0.03	0.36 +/- 0.08	1.11 +/- 0.02
Isle Royale NP	1.95 +/- 0.04	1.03 +/- 0.02	0.81 +/- 0.02	0.61 +/- 0.13
James River Face Wilderness	1.72 +/- 0.04	1.04 +/- 0.01	0.29 +/- 0.05	0.92 +/- 0.08
Jarbidge Wilderness	1.84 +/- 0.04	0.94 +/- 0.03	0.72 +/- 0.05	0.99 +/- 0.02
Joshua Tree NP	1.92 +/- 0.07	0.87 +/- 0.03	0.62 +/- 0.02	0.93 +/- 0.03
Kaiser	1.83 +/- 0.03	0.76 +/- 0.04	0.72 +/- 0.03	0.99 +/- 0.03
Kalmiopsis	1.53 +/- 0.02	0.93 +/- 0.04	0.93 +/- 0.15	1.06 +/- 0.08
Lassen Volcanic NP	1.67 +/- 0.03	1.06 +/- 0.04	0.31 +/- 0.06	1.06 +/- 0.04
Lava Beds NM	1.68 +/- 0.03	1.04 +/- 0.07	0.25 +/- 0.11	1.04 +/- 0.07
Linville Gorge	1.78 +/- 0.04	1.10 +/- 0.01	-0.07 +/- 0.07	0.83 +/- 0.08
Livonia	1.66 +/- 0.07	1.08 +/- 0.02	0.71 +/- 0.02	1.06 +/- 0.07
Lostwood	1.81 +/- 0.04	0.91 +/- 0.02	0.80 +/- 0.02	0.99 +/- 0.05
Lye Brook Wilderness	1.94 +/- 0.06	1.00 +/- 0.02	0.31 +/- 0.04	0.79 +/- 0.12
M.K. Goddard	1.50 +/- 0.05	1.08 +/- 0.02	0.69 +/- 0.03	0.93 +/- 0.15
Mammoth Cave NP	1.88 +/- 0.05	0.93 +/- 0.01	0.42 +/- 0.02	0.91 +/- 0.05
Marthas Vineyard	1.87 +/- 0.06	1.04 +/- 0.01	0.18 +/- 0.06	0.93 +/- 0.12
Meadview	1.84 +/- 0.07	0.99 +/- 0.03	0.32 +/- 0.06	1.10 +/- 0.02

site	β_{oc}	β_{sulf}	β_{nit}	β_{soil}
Medicine Lake	1.84 +/- 0.04	0.91 +/- 0.02	0.69 +/- 0.02	0.99 +/- 0.04
Mesa Verde NP	1.79 +/- 0.06	1.08 +/- 0.05	0.18 +/- 0.09	1.19 +/- 0.02
Mohawk Mt.	1.53 +/- 0.07	1.02 +/- 0.02	0.28 +/- 0.06	1.11 +/- 0.19
Monture	1.67 +/- 0.02	1.01 +/- 0.03	0.22 +/- 0.13	1.08 +/- 0.03
Moosehorn NWR	1.70 +/- 0.04	1.02 +/- 0.02	0.23 +/- 0.07	0.70 +/- 0.14
Mount Baldy	1.44 +/- 0.03	1.05 +/- 0.02	0.55 +/- 0.06	1.09 +/- 0.02
Mount Hood	1.79 +/- 0.03	1.17 +/- 0.03	0.21 +/- 0.06	1.00 +/- 0.06
Mount Rainier NP	1.59 +/- 0.03	1.20 +/- 0.04	0.38 +/- 0.13	1.16 +/- 0.09
Mount Zirkel Wilderness	2.02 +/- 0.04	0.82 +/- 0.03	0.33 +/- 0.06	1.04 +/- 0.03
Nebraska NF	1.91 +/- 0.07	0.98 +/- 0.03	0.77 +/- 0.02	0.84 +/- 0.07
New York City	1.62 +/- 0.09	1.01 +/- 0.03	0.91 +/- 0.04	**0.27 +/- 0.19**
North Absaroka	1.91 +/- 0.03	0.97 +/- 0.03	0.32 +/- 0.06	1.11 +/- 0.03
North Cascades	1.79 +/- 0.03	1.10 +/- 0.03	0.75 +/- 0.14	0.89 +/- 0.07
Northern Cheyenne	1.91 +/- 0.03	1.03 +/- 0.03	0.10 +/- 0.05	0.96 +/- 0.04
Okefenokee NWR	1.65 +/- 0.03	0.98 +/- 0.01	0.04 +/- 0.10	0.97 +/- 0.03
Olympic	1.53 +/- 0.03	1.11 +/- 0.03	0.47 +/- 0.06	0.99 +/- 0.10
Omaha	1.75 +/- 0.06	1.00 +/- 0.02	0.97 +/- 0.02	0.67 +/- 0.07
Organ Pipe	1.45 +/- 0.08	0.99 +/- 0.02	0.44 +/- 0.07	1.09 +/- 0.02
Pasayten	1.69 +/- 0.02	1.08 +/- 0.03	0.19 +/- 0.06	1.09 +/- 0.05
Petrified Forest NP	1.66 +/- 0.05	1.05 +/- 0.03	0.44 +/- 0.09	1.04 +/- 0.02
Phoenix	1.25 +/- 0.02	0.95 +/- 0.03	0.64 +/- 0.03	1.08 +/- 0.02
Pinnacles NM	1.69 +/- 0.06	0.93 +/- 0.05	0.65 +/- 0.05	1.05 +/- 0.13
Point Reyes National Seashore	1.58 +/- 0.07	0.95 +/- 0.03	0.78 +/- 0.03	1.03 +/- 0.17
Presque Isle	1.79 +/- 0.03	0.94 +/- 0.01	-0.01 +/- 0.05	0.94 +/- 0.03
Proctor Maple R. F.	1.86 +/- 0.04	1.01 +/- 0.01	0.46 +/- 0.04	0.69 +/- 0.13
Puget Sound	1.39 +/- 0.03	0.85 +/- 0.03	1.00 +/- 0.04	0.73 +/- 0.09
Quabbin Summit	1.76 +/- 0.04	0.96 +/- 0.01	0.30 +/- 0.04	0.85 +/- 0.11
Quaker City	1.64 +/- 0.06	1.07 +/- 0.01	0.55 +/- 0.03	0.86 +/- 0.09
Queen Valley	1.51 +/- 0.07	1.01 +/- 0.03	0.57 +/- 0.02	1.08 +/- 0.02
Redwood NP	1.73 +/- 0.03	0.95 +/- 0.03	0.80 +/- 0.09	0.88 +/- 0.09
Rocky Mountain NP	1.84 +/- 0.05	0.81 +/- 0.05	0.56 +/- 0.04	1.05 +/- 0.03
Sac and Fox	1.56 +/- 0.05	0.98 +/- 0.02	0.90 +/- 0.01	0.93 +/- 0.05
Saguaro NM	1.35 +/- 0.06	1.00 +/- 0.03	0.38 +/- 0.04	1.14 +/- 0.01
Saguaro West	1.35 +/- 0.11	0.99 +/- 0.04	0.37 +/- 0.05	1.12 +/- 0.02
Salt Creek	1.42 +/- 0.10	1.02 +/- 0.03	0.94 +/- 0.04	1.10 +/- 0.02
San Gabriel	1.86 +/- 0.05	0.81 +/- 0.03	0.50 +/- 0.02	1.02 +/- 0.04
San Gorgonio Wilderness	1.46 +/- 0.06	0.83 +/- 0.04	0.77 +/- 0.01	0.96 +/- 0.04
San Pedro Parks	1.62 +/- 0.05	1.02 +/- 0.03	0.18 +/- 0.11	1.16 +/- 0.02
San Rafael	1.70 +/- 0.05	0.95 +/- 0.03	0.53 +/- 0.03	1.06 +/- 0.05
Sawtooth NF	1.65 +/- 0.03	1.01 +/- 0.07	**-1.79 +/- 0.50**	1.21 +/- 0.05
Seney	1.66 +/- 0.04	0.97 +/- 0.02	0.70 +/- 0.02	0.78 +/- 0.14
Sequoia NP	1.68 +/- 0.04	0.79 +/- 0.05	0.90 +/- 0.01	1.09 +/- 0.06

site	β_{oc}	β_{sulf}	β_{nit}	β_{soil}
Shamrock Mine	1.92 +/- 0.04	1.05 +/- 0.03	0.11 +/- 0.06	0.93 +/- 0.01
Shenandoah NP	1.73 +/- 0.06	1.10 +/- 0.01	0.36 +/- 0.04	1.02 +/- 0.08
Shining Rock Wilderness	1.80 +/- 0.08	1.02 +/- 0.02	**-0.27 +/- 0.12**	0.95 +/- 0.07
Sierra Ancha	1.30 +/- 0.04	1.04 +/- 0.03	0.29 +/- 0.06	1.13 +/- 0.02
Sikes	1.71 +/- 0.03	1.06 +/- 0.01	0.14 +/- 0.05	1.03 +/- 0.02
Sipsy Wilderness	1.85 +/- 0.04	0.98 +/- 0.01	0.23 +/- 0.03	0.97 +/- 0.04
Snoqualmie Pass	1.64 +/- 0.05	1.09 +/- 0.05	0.43 +/- 0.06	0.98 +/- 0.15
St. Marks	1.65 +/- 0.04	1.04 +/- 0.01	0.15 +/- 0.12	1.02 +/- 0.03
Starkey	1.63 +/- 0.02	0.90 +/- 0.04	0.70 +/- 0.02	1.15 +/- 0.03
Sula Peak	1.73 +/- 0.02	0.99 +/- 0.04	-0.07 +/- 0.09	0.99 +/- 0.03
Swanquarter	1.80 +/- 0.05	1.01 +/- 0.01	0.01 +/- 0.06	1.06 +/- 0.05
Sycamore Canyon	1.29 +/- 0.04	1.13 +/- 0.04	0.47 +/- 0.06	1.04 +/- 0.01
Tallgrass	1.44 +/- 0.04	1.04 +/- 0.02	0.80 +/- 0.02	1.00 +/- 0.04
Theodore Roosevelt	1.83 +/- 0.04	0.96 +/- 0.03	0.80 +/- 0.03	0.91 +/- 0.03
Three Sisters Wilderness	1.75 +/- 0.02	1.07 +/- 0.03	0.35 +/- 0.10	0.99 +/- 0.04
Thunder Basin	1.86 +/- 0.03	0.94 +/- 0.02	0.54 +/- 0.03	0.92 +/- 0.02
Tonto NM	1.71 +/- 0.05	1.00 +/- 0.03	0.27 +/- 0.04	1.08 +/- 0.01
Trinity	1.63 +/- 0.03	1.08 +/- 0.05	0.40 +/- 0.06	0.94 +/- 0.06
UL Bend	1.91 +/- 0.03	0.86 +/- 0.02	0.81 +/- 0.03	1.04 +/- 0.03
Upper Buffalo Wilderness	1.63 +/- 0.04	1.04 +/- 0.01	0.68 +/- 0.02	1.01 +/- 0.03
Viking Lake	1.54 +/- 0.05	1.05 +/- 0.02	0.99 +/- 0.01	0.91 +/- 0.06
Voyageurs NP #2	1.70 +/- 0.04	0.94 +/- 0.02	0.87 +/- 0.02	0.97 +/- 0.13
Washington D.C.	1.67 +/- 0.06	1.07 +/- 0.02	0.78 +/- 0.03	**0.21 +/- 0.14**
Weminuche Wilderness	1.78 +/- 0.04	1.00 +/- 0.04	**-0.19 +/- 0.13**	1.09 +/- 0.02
Wheeler Peak	1.68 +/- 0.06	1.08 +/- 0.04	0.17 +/- 0.13	1.24 +/- 0.03
White Mountain	1.63 +/- 0.06	1.08 +/- 0.03	0.55 +/- 0.04	1.13 +/- 0.02
White Pass	1.82 +/- 0.04	1.18 +/- 0.04	0.14 +/- 0.09	0.87 +/- 0.06
White River NF	1.96 +/- 0.05	1.08 +/- 0.04	**-0.18 +/- 0.11**	1.17 +/- 0.02
Wichita Mountains	1.53 +/- 0.05	1.10 +/- 0.02	0.81 +/- 0.02	0.95 +/- 0.03
Wind Cave	1.72 +/- 0.03	0.93 +/- 0.03	0.41 +/- 0.03	1.10 +/- 0.03
Yellowstone NP 2	1.75 +/- 0.02	0.87 +/- 0.03	0.52 +/- 0.04	1.08 +/- 0.03
Yosemite NP	1.64 +/- 0.02	1.01 +/- 0.03	0.77 +/- 0.03	1.09 +/- 0.05
Zion Canyon	1.76 +/- 0.05	1.11 +/- 0.03	0.28 +/- 0.05	1.15 +/- 0.02

Table S3. Quarterly regressions eliminated because of high colinearity among covariates

site	quarter	max $\|r_p\|$ among covariate pairs	correlated covariates
Northern Cheyenne	3	0.86	sulfate:nitrate
Sula Peak	3	0.87	oc:nitrate
Cape Cod	4	0.90	sulfate:soil
Lye Brook Wilderness	4	0.86 : 0.86	oc:sulfate ; soil:sulfate
Marthas Vineyard	4	0.85	soil:sulfate
Mohawk Mt.	4	0.88	soil:nitrate
Puget Sound	4	0.86	sulfate:nitrate

Table S4. Quarterly regressions eliminated because of high correlation between residual error (ε_i) and a $PM_{2.5}$ component ($|r_s| > 0.4$).

site	quarter	Max $\|r_s\|$ between ε_i and $PM_{2.5}$ constituents	$PM_{2.5}$ constituents correlated to ε_i
Bosque del Apache	2	0.45	chloride
M.K. Goddard	2	0.41	chloride
Mount Hood	2	0.44 ; 0.46	OC ; EC
Salt Creek	2	0.49	chloride
St. Marks	2	0.53	chloride
Bosque del Apache	3	0.48	chloride
Hercules-Glades	3	0.41	chloride
Lostwood	3	0.45	chloride
Sac and Fox	3	0.44	chloride
St. Marks	3	0.50	chloride
Dome Lands Wilderness	4	0.44	soil
Sawtooth NF	4	0.42	KNON

Table S5. High-confidence quarter-specific regression results. This table includes 10 regressions flagged for an outlier year, but excluding that year did not change the regression coefficients (see Table S6).

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Acadia NP	1	1.66 +/- 0.14	0.89 +/- 0.04	0.62 +/- 0.11	7.38	-0.21
Addison Pinnacle	1	1.46 +/- 0.11	0.82 +/- 0.03	0.98 +/- 0.05	5.93	-0.12
Agua Tibia	1	1.38 +/- 0.12	1.08 +/- 0.08	0.56 +/- 0.04	9.52	0.07
Arendtsville	1	1.21 +/- 0.09	0.88 +/- 0.03	1.18 +/- 0.04	6.15	-0.38
Bandelier NM	1	1.15 +/- 0.08	1.05 +/- 0.06	0.71 +/- 0.06	10.48	-0.30
Big Bend NP	1	1.83 +/- 0.12	0.97 +/- 0.03	0.40 +/- 0.08	6.88	0.04
Birmingham	1	1.50 +/- 0.05	0.87 +/- 0.04	0.82 +/- 0.07	5.88	-0.14
Bliss SP (TRPA)	1	1.65 +/- 0.09	0.97 +/- 0.08	0.40 +/- 0.08	14.37	-1.83
Blue Mounds	1	1.53 +/- 0.14	1.00 +/- 0.04	1.08 +/- 0.02	5.24	0.12
Bondville	1	1.44 +/- 0.14	0.89 +/- 0.04	1.03 +/- 0.03	6.41	-0.19
Bosque del Apache	1	1.05 +/- 0.09	0.99 +/- 0.06	0.91 +/- 0.06	9.64	-0.08
Boundary Waters Canoe Area	1	1.62 +/- 0.25	0.87 +/- 0.06	0.88 +/- 0.04	9.16	0.59
Bridger Wilderness	1	1.81 +/- 0.19	0.89 +/- 0.06	0.36 +/- 0.12	13.66	-0.46
Bridgton	1	1.48 +/- 0.09	0.98 +/- 0.03	0.59 +/- 0.11	7.59	-0.04
Brigantine NWR	1	1.44 +/- 0.10	0.88 +/- 0.03	0.94 +/- 0.06	5.90	-0.10
Bryce Canyon NP	1	1.11 +/- 0.17	1.05 +/- 0.09	0.65 +/- 0.06	16.44	-1.81
Cabinet Mountains	1	1.48 +/- 0.08	1.03 +/- 0.05	0.38 +/- 0.12	14.49	1.49
Cadiz	1	1.47 +/- 0.09	0.91 +/- 0.03	0.93 +/- 0.03	7.14	-0.33
Caney Creek	1	1.64 +/- 0.07	0.89 +/- 0.03	0.62 +/- 0.04	7.95	-0.80
Canyonlands NP	1	1.31 +/- 0.25	0.95 +/- 0.07	0.65 +/- 0.08	11.49	-0.13
Cape Cod	1	1.62 +/- 0.14	0.85 +/- 0.03	0.80 +/- 0.10	6.38	-0.08
Cape Romain NWR	1	1.60 +/- 0.05	0.99 +/- 0.03	0.04 +/- 0.14	6.68	-0.30
Capitol Reef NP	1	1.37 +/- 0.28	0.82 +/- 0.12	0.62 +/- 0.10	15.98	-1.60
Casco Bay	1	1.38 +/- 0.06	0.95 +/- 0.04	0.95 +/- 0.12	7.48	-0.63
Cedar Bluff	1	1.18 +/- 0.13	0.87 +/- 0.07	1.01 +/- 0.03	9.92	-1.06

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Chassahowitzka NWR	1	1.46 +/- 0.06	1.07 +/- 0.03	0.25 +/- 0.18	6.49	0.12
Cherokee Nation	1	1.32 +/- 0.08	0.94 +/- 0.05	1.00 +/- 0.03	7.72	-0.45
Chiricahua NM	1	1.99 +/- 0.25	0.88 +/- 0.09	0.29 +/- 0.16	10.11	-0.49
Cohutta	1	1.84 +/- 0.06	0.86 +/- 0.02	0.41 +/- 0.05	6.50	-0.10
Columbia Gorge #1	1	1.41 +/- 0.06	0.79 +/- 0.06	0.62 +/- 0.04	9.11	-0.89
Columbia River Gorge	1	1.00 +/- 0.08	0.79 +/- 0.09	0.82 +/- 0.04	12.37	0.66
Connecticut Hill	1	1.59 +/- 0.18	0.75 +/- 0.04	1.04 +/- 0.05	6.13	-0.25
Crater Lake NP	1	1.18 +/- 0.11	1.23 +/- 0.08	0.45 +/- 0.17	17.40	-2.30
Craters of the Moon NM	1	1.24 +/- 0.24	0.82 +/- 0.10	0.67 +/- 0.04	17.40	-0.57
Crescent Lake	1	1.31 +/- 0.20	0.95 +/- 0.07	1.04 +/- 0.03	9.64	-0.61
Death Valley NP	1	1.42 +/- 0.23	1.02 +/- 0.11	0.76 +/- 0.14	11.50	1.20
Dolly Sods Wilderness	1	1.38 +/- 0.07	0.99 +/- 0.03	0.57 +/- 0.06	7.56	-0.05
Dome Lands Wilderness	1	1.21 +/- 0.12	0.69 +/- 0.10	0.86 +/- 0.03	10.14	-0.11
Douglas	1	1.86 +/- 0.14	0.62 +/- 0.14	-0.06 +/- 0.25	6.33	0.27
El Dorado Springs	1	1.30 +/- 0.06	0.90 +/- 0.04	0.91 +/- 0.03	7.32	-0.20
Ellis	1	1.24 +/- 0.09	0.97 +/- 0.05	0.99 +/- 0.03	8.14	-1.08
Everglades NP	1	1.97 +/- 0.13	1.02 +/- 0.03	-0.10 +/- 0.20	7.86	0.23
Flathead	1	1.48 +/- 0.09	0.91 +/- 0.05	0.62 +/- 0.10	11.99	0.36
Frostberg Reservoir (Big Piney Run)	1	2.03 +/- 0.12	0.83 +/- 0.03	0.46 +/- 0.06	5.77	-0.12
Gates of the Mountains	1	1.26 +/- 0.11	1.08 +/- 0.05	0.32 +/- 0.10	15.05	0.20
Gila Wilderness	1	1.18 +/- 0.09	1.06 +/- 0.06	0.75 +/- 0.21	9.18	-0.02
Glacier NP	1	1.52 +/- 0.03	0.79 +/- 0.04	1.05 +/- 0.06	7.46	-0.68
Great Basin NP	1	1.05 +/- 0.10	1.06 +/- 0.13	0.49 +/- 0.15	13.34	-1.00
Great Gulf Wilderness	1	1.62 +/- 0.18	0.97 +/- 0.04	0.46 +/- 0.12	8.38	-0.31
Great River Bluffs	1	1.42 +/- 0.20	0.86 +/- 0.07	1.00 +/- 0.03	8.40	-0.28
Great Sand Dunes NM	1	1.15 +/- 0.12	1.01 +/- 0.07	0.74 +/- 0.09	10.27	-0.41
Great Smoky Mountains NP	1	1.85 +/- 0.06	0.87 +/- 0.02	0.54 +/- 0.05	6.33	0.02

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Guadalupe Mountains NP	1	1.29 +/- 0.28	1.14 +/- 0.13	0.53 +/- 0.10	11.30	1.00
Hells Canyon	1	1.33 +/- 0.10	0.97 +/- 0.09	0.72 +/- 0.03	13.60	-0.64
Hercules-Glades	1	1.61 +/- 0.07	0.87 +/- 0.04	0.74 +/- 0.04	8.07	-0.23
Hoover	1	1.42 +/- 0.17	1.17 +/- 0.08	0.43 +/- 0.09	15.29	0.01
Indian Gardens	1	1.20 +/- 0.16	1.15 +/- 0.10	0.32 +/- 0.08	10.96	1.37
Isle Royale NP	1	2.02 +/- 0.28	0.92 +/- 0.06	0.87 +/- 0.03	9.10	0.79
James River Face Wilderness	1	1.64 +/- 0.06	0.90 +/- 0.02	0.69 +/- 0.05	6.06	-0.10
Jarbidge Wilderness	1	1.12 +/- 0.18	1.17 +/- 0.09	0.69 +/- 0.08	16.68	0.89
Joshua Tree NP	1	1.26 +/- 0.16	0.79 +/- 0.10	0.78 +/- 0.03	10.58	0.46
Kaiser	1	1.06 +/- 0.13	1.02 +/- 0.09	1.00 +/- 0.05	12.89	-1.33
Lassen Volcanic NP	1	1.39 +/- 0.09	1.16 +/- 0.07	0.35 +/- 0.08	13.82	-1.05
Lava Beds NM	1	1.43 +/- 0.06	1.07 +/- 0.07	0.41 +/- 0.15	14.09	-1.13
Linville Gorge	1	1.80 +/- 0.05	0.92 +/- 0.02	0.33 +/- 0.08	6.82	-0.23
Livonia	1	1.58 +/- 0.09	0.92 +/- 0.03	0.92 +/- 0.02	6.01	-0.17
Lostwood	1	1.11 +/- 0.16	0.95 +/- 0.04	0.96 +/- 0.04	9.27	-0.92
Lye Brook Wilderness	1	1.60 +/- 0.23	0.87 +/- 0.05	0.77 +/- 0.07	9.20	-0.33
M.K. Goddard	1	1.57 +/- 0.09	0.79 +/- 0.03	0.98 +/- 0.04	6.55	-0.26
Mammoth Cave NP	1	1.79 +/- 0.07	0.77 +/- 0.02	0.67 +/- 0.03	6.77	-0.22
Marthas Vineyard	1	1.77 +/- 0.16	0.84 +/- 0.04	0.73 +/- 0.12	5.80	-0.29
Medicine Lake	1	1.59 +/- 0.19	0.89 +/- 0.04	0.76 +/- 0.04	10.64	-0.44
Mesa Verde NP	1	1.39 +/- 0.17	1.10 +/- 0.09	0.26 +/- 0.14	12.42	0.85
Mohawk Mt.	1	1.47 +/- 0.15	0.94 +/- 0.04	0.58 +/- 0.09	7.32	0.10
Monture	1	1.29 +/- 0.06	1.05 +/- 0.05	0.73 +/- 0.18	14.05	0.68
Moosehorn NWR	1	1.29 +/- 0.10	0.97 +/- 0.03	0.76 +/- 0.11	7.11	-0.07
Mount Baldy	1	1.31 +/- 0.05	1.02 +/- 0.04	0.60 +/- 0.07	9.14	0.06
Mount Hood	1	1.80 +/- 0.14	1.17 +/- 0.06	0.02 +/- 0.13	16.56	0.39
Mount Rainier NP	1	1.46 +/- 0.04	0.97 +/- 0.06	0.90 +/- 0.17	9.86	0.44
Mount Zirkel Wilderness	1	1.17 +/- 0.26	0.91 +/- 0.09	0.68 +/- 0.09	13.37	0.06

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Nebraska NF	1	1.45 +/- 0.22	0.70 +/- 0.08	1.11 +/- 0.05	10.90	-0.15
North Cascades	1	1.62 +/- 0.12	1.04 +/- 0.07	0.93 +/- 0.27	15.51	0.36
Northern Cheyenne	1	1.86 +/- 0.11	0.95 +/- 0.04	0.19 +/- 0.06	10.85	-0.11
Okefenokee NWR	1	1.64 +/- 0.05	0.94 +/- 0.02	0.25 +/- 0.14	6.33	-0.25
Olympic	1	1.36 +/- 0.06	1.00 +/- 0.07	0.67 +/- 0.09	8.86	0.02
Omaha	1	1.23 +/- 0.12	1.02 +/- 0.03	1.09 +/- 0.03	4.82	0.01
Organ Pipe	1	1.40 +/- 0.13	1.01 +/- 0.06	0.35 +/- 0.10	8.48	0.83
Pasayten	1	1.47 +/- 0.11	1.11 +/- 0.06	0.27 +/- 0.08	16.62	0.64
Pinnacles NM	1	1.39 +/- 0.23	0.82 +/- 0.24	0.84 +/- 0.13	13.03	1.07
Point Reyes National Seashore	1	1.02 +/- 0.13	0.88 +/- 0.05	0.99 +/- 0.04	7.92	-0.93
Presque Isle	1	1.36 +/- 0.07	0.87 +/- 0.03	0.91 +/- 0.11	5.49	0.00
Proctor Maple R. F.	1	1.32 +/- 0.09	0.91 +/- 0.02	0.96 +/- 0.05	6.22	-0.04
Puget Sound	1	1.25 +/- 0.05	0.66 +/- 0.09	1.23 +/- 0.08	6.18	-0.05
Quabbin Summit	1	1.47 +/- 0.11	0.93 +/- 0.03	0.65 +/- 0.07	6.14	-0.07
Quaker City	1	1.81 +/- 0.10	0.86 +/- 0.02	0.79 +/- 0.03	6.66	-0.09
Queen Valley	1	1.31 +/- 0.16	0.94 +/- 0.09	0.54 +/- 0.03	7.88	0.34
Redwood NP	1	1.65 +/- 0.07	0.76 +/- 0.05	0.97 +/- 0.13	7.65	-0.85
Sac and Fox	1	1.27 +/- 0.08	0.84 +/- 0.05	1.07 +/- 0.03	7.23	-0.34
Saguaro West	1	1.33 +/- 0.20	0.96 +/- 0.13	0.28 +/- 0.08	8.45	0.73
San Gabriel	1	1.10 +/- 0.18	0.74 +/- 0.10	0.76 +/- 0.04	10.47	-0.40
San Pedro Parks	1	1.70 +/- 0.28	1.04 +/- 0.09	-0.11 +/- 0.18	12.96	0.03
Sawtooth NF	1	1.10 +/- 0.04	0.94 +/- 0.12	0.68 +/- 0.47	14.36	-1.80
Seney	1	1.42 +/- 0.16	0.91 +/- 0.03	0.86 +/- 0.03	7.24	0.87
Shenandoah NP	1	1.72 +/- 0.08	0.85 +/- 0.02	0.78 +/- 0.04	7.21	-0.43
Shining Rock Wilderness	1	1.58 +/- 0.08	0.97 +/- 0.03	0.13 +/- 0.11	8.78	-0.33
Sierra Ancha	1	1.15 +/- 0.09	1.07 +/- 0.09	0.29 +/- 0.10	11.57	-0.45
Sikes	1	1.52 +/- 0.05	1.01 +/- 0.02	0.44 +/- 0.05	5.92	-0.07
Sipsy Wilderness	1	1.81 +/- 0.05	0.84 +/- 0.02	0.44 +/- 0.04	6.30	-0.04

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Starkey	1	1.40 +/- 0.05	0.77 +/- 0.05	0.78 +/- 0.03	10.02	-0.53
Sula Peak	1	1.30 +/- 0.10	1.05 +/- 0.07	0.36 +/- 0.16	15.56	-2.26
Swanquarter	1	1.71 +/- 0.09	0.99 +/- 0.02	0.19 +/- 0.09	6.60	0.02
Sycamore Canyon	1	1.21 +/- 0.05	0.95 +/- 0.06	0.52 +/- 0.05	8.21	0.25
Tallgrass	1	1.44 +/- 0.08	0.77 +/- 0.05	0.97 +/- 0.03	8.75	-0.50
Theodore Roosevelt	1	1.01 +/- 0.14	1.06 +/- 0.04	0.90 +/- 0.04	7.70	-0.31
Three Sisters Wilderness	1	1.65 +/- 0.09	1.05 +/- 0.06	0.33 +/- 0.20	14.39	-1.15
Thunder Basin	1	1.48 +/- 0.10	0.88 +/- 0.05	0.74 +/- 0.05	7.62	0.08
Tonto NM	1	1.52 +/- 0.14	1.05 +/- 0.08	0.33 +/- 0.06	8.99	0.16
Trinity	1	1.42 +/- 0.07	1.26 +/- 0.08	0.37 +/- 0.11	13.19	-0.85
UL Bend	1	1.65 +/- 0.12	0.81 +/- 0.04	0.90 +/- 0.04	9.92	0.05
Upper Buffalo Wilderness	1	1.46 +/- 0.06	0.87 +/- 0.03	0.89 +/- 0.03	7.83	-0.55
Viking Lake	1	1.27 +/- 0.09	0.92 +/- 0.04	1.12 +/- 0.02	5.27	0.01
Voyageurs NP #2	1	1.12 +/- 0.12	0.89 +/- 0.04	0.98 +/- 0.02	7.15	0.14
Weminuche Wilderness	1	1.14 +/- 0.10	1.01 +/- 0.07	0.45 +/- 0.14	10.27	-1.08
White Mountain	1	1.51 +/- 0.17	1.05 +/- 0.07	0.60 +/- 0.06	9.25	-0.14
White Pass	1	1.30 +/- 0.20	1.15 +/- 0.08	0.59 +/- 0.16	22.56	-1.95
White River NF	1	1.29 +/- 0.27	1.02 +/- 0.11	0.20 +/- 0.15	14.38	-0.04
Wichita Mountains	1	1.27 +/- 0.07	0.92 +/- 0.04	1.01 +/- 0.03	8.22	-0.84
Wind Cave	1	1.35 +/- 0.08	0.95 +/- 0.04	0.55 +/- 0.04	9.80	0.15
Yosemite NP	1	1.47 +/- 0.13	1.00 +/- 0.10	0.88 +/- 0.04	13.25	-2.28
Zion Canyon	1	1.17 +/- 0.12	1.07 +/- 0.08	0.55 +/- 0.07	11.12	-0.95
Acadia NP	2	1.93 +/- 0.06	0.97 +/- 0.02	0.45 +/- 0.11	6.54	0.38
Addison Pinnacle	2	1.63 +/- 0.13	1.08 +/- 0.03	0.35 +/- 0.10	7.37	-1.00
Agua Tibia	2	2.10 +/- 0.09	0.79 +/- 0.05	0.58 +/- 0.05	6.55	0.04
Arendtsville	2	1.84 +/- 0.12	0.99 +/- 0.03	0.69 +/- 0.05	7.31	-0.45
Badlands NP	2	1.69 +/- 0.09	0.89 +/- 0.06	0.45 +/- 0.14	11.19	0.13
Bandelier NM	2	1.43 +/- 0.09	0.98 +/- 0.07	0.82 +/- 0.27	7.86	0.07

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Big Bend NP	2	2.16 +/- 0.12	0.96 +/- 0.03	0.24 +/- 0.28	6.72	0.19
Birmingham	2	1.41 +/- 0.08	1.07 +/- 0.04	0.18 +/- 0.34	5.54	0.21
Bliss SP (TRPA)	2	1.94 +/- 0.05	0.67 +/- 0.05	0.91 +/- 0.13	7.31	-0.48
Bondville	2	1.95 +/- 0.12	1.03 +/- 0.04	0.63 +/- 0.05	8.55	-0.67
Bridger Wilderness	2	1.83 +/- 0.07	0.93 +/- 0.05	0.45 +/- 0.18	7.90	-0.13
Cabinet Mountains	2	1.82 +/- 0.05	0.98 +/- 0.05	0.45 +/- 0.18	6.96	0.09
Cadiz	2	1.90 +/- 0.07	0.96 +/- 0.02	0.67 +/- 0.06	6.17	-0.07
Caney Creek	2	1.87 +/- 0.10	1.03 +/- 0.03	0.01 +/- 0.16	6.51	-0.15
Canyonlands NP	2	2.14 +/- 0.12	0.76 +/- 0.06	0.89 +/- 0.24	7.51	0.21
Cape Cod	2	1.98 +/- 0.12	0.98 +/- 0.03	0.35 +/- 0.14	7.03	-0.99
Cape Romain NWR	2	1.76 +/- 0.08	0.98 +/- 0.02	0.04 +/- 0.21	6.76	-0.01
Capitol Reef NP	2	2.20 +/- 0.12	0.75 +/- 0.08	0.33 +/- 0.24	7.39	-0.29
Casco Bay	2	1.72 +/- 0.07	1.10 +/- 0.03	0.25 +/- 0.14	7.36	-0.67
Cedar Bluff	2	1.69 +/- 0.10	1.03 +/- 0.04	0.40 +/- 0.05	6.96	-0.06
Chassahowitzka NWR	2	1.79 +/- 0.10	0.96 +/- 0.03	0.24 +/- 0.21	6.90	-0.02
Cherokee Nation	2	1.72 +/- 0.05	1.00 +/- 0.02	0.45 +/- 0.05	5.38	-0.07
Cloud Peak	2	1.92 +/- 0.07	0.99 +/- 0.06	0.30 +/- 0.16	7.81	-0.07
Cohutta	2	2.15 +/- 0.10	0.90 +/- 0.03	-0.26 +/- 0.19	5.76	-0.18
Columbia Gorge #1	2	1.57 +/- 0.06	0.85 +/- 0.06	0.52 +/- 0.09	7.03	-0.08
Columbia River Gorge	2	1.34 +/- 0.08	1.14 +/- 0.08	0.40 +/- 0.18	8.28	-0.61
Crater Lake NP	2	1.81 +/- 0.06	0.90 +/- 0.05	1.08 +/- 0.26	7.85	-0.14
Craters of the Moon NM	2	1.97 +/- 0.06	0.95 +/- 0.05	0.24 +/- 0.12	8.30	-0.26
Crescent Lake	2	2.02 +/- 0.13	0.94 +/- 0.06	0.73 +/- 0.06	9.17	0.14
Dolly Sods Wilderness	2	1.94 +/- 0.14	0.99 +/- 0.03	0.07 +/- 0.14	6.90	-0.26
Dome Lands Wilderness	2	2.05 +/- 0.14	0.64 +/- 0.11	0.56 +/- 0.06	7.08	-0.23
El Dorado Springs	2	1.78 +/- 0.07	0.98 +/- 0.03	0.27 +/- 0.06	6.24	-0.24
Ellis	2	1.88 +/- 0.08	0.98 +/- 0.03	0.51 +/- 0.06	6.52	-0.39
Everglades NP	2	1.53 +/- 0.04	1.04 +/- 0.03	0.44 +/- 0.21	7.28	0.10

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Flathead	2	1.76 +/- 0.05	0.97 +/- 0.06	0.02 +/- 0.22	8.40	-0.52
Fort Peck	2	1.95 +/- 0.08	0.90 +/- 0.04	0.28 +/- 0.06	7.60	0.41
Frostberg Reservoir (Big Piney Run)	2	2.25 +/- 0.14	0.89 +/- 0.03	-0.07 +/- 0.10	5.34	0.34
Gates of the Mountains	2	1.74 +/- 0.08	0.89 +/- 0.08	0.45 +/- 0.20	9.98	0.30
Glacier NP	2	1.73 +/- 0.05	0.98 +/- 0.08	0.27 +/- 0.55	7.39	-0.73
Great Basin NP	2	1.56 +/- 0.14	0.88 +/- 0.10	1.10 +/- 0.37	9.56	0.10
Great Gulf Wilderness	2	1.70 +/- 0.07	1.06 +/- 0.03	0.04 +/- 0.33	7.36	-0.15
Great River Bluffs	2	1.69 +/- 0.07	1.01 +/- 0.03	0.41 +/- 0.05	8.14	0.01
Great Sand Dunes NM	2	1.89 +/- 0.11	0.75 +/- 0.09	1.07 +/- 0.27	7.76	-0.20
Great Smoky Mountains NP	2	2.15 +/- 0.10	1.01 +/- 0.03	-0.15 +/- 0.18	6.27	-0.39
Guadalupe Mountains NP	2	1.52 +/- 0.15	1.01 +/- 0.05	0.81 +/- 0.24	6.62	0.27
Hance Camp at Grand Canyon NP	2	1.89 +/- 0.10	0.93 +/- 0.06	0.81 +/- 0.15	7.05	-0.22
Hercules-Glades	2	1.72 +/- 0.07	0.98 +/- 0.03	0.39 +/- 0.10	6.94	-0.27
Hoover	2	2.04 +/- 0.07	0.70 +/- 0.06	0.71 +/- 0.20	8.25	0.38
Ikes Backbone	2	1.71 +/- 0.14	0.63 +/- 0.09	1.05 +/- 0.24	8.69	0.28
Indian Gardens	2	1.90 +/- 0.11	0.74 +/- 0.07	1.01 +/- 0.22	7.07	0.17
Isle Royale NP	2	1.94 +/- 0.07	1.09 +/- 0.03	0.62 +/- 0.08	8.44	-0.01
James River Face Wilderness	2	2.19 +/- 0.09	0.91 +/- 0.03	0.02 +/- 0.12	6.36	-0.17
Jarbidge Wilderness	2	1.83 +/- 0.08	0.86 +/- 0.07	1.21 +/- 0.36	8.64	-0.56
Joshua Tree NP	2	2.39 +/- 0.14	0.62 +/- 0.08	0.56 +/- 0.04	7.39	-0.18
Kaiser	2	1.87 +/- 0.08	0.78 +/- 0.07	0.61 +/- 0.08	7.88	0.22
Kalmiopsis	2	1.70 +/- 0.07	0.87 +/- 0.06	0.76 +/- 0.22	8.14	-0.33
Lassen Volcanic NP	2	1.72 +/- 0.05	0.96 +/- 0.05	0.47 +/- 0.09	8.70	1.16
Lava Beds NM	2	1.73 +/- 0.06	0.88 +/- 0.06	0.85 +/- 0.19	9.34	0.22
Linville Gorge	2	1.99 +/- 0.08	1.02 +/- 0.02	0.01 +/- 0.16	5.82	0.08
Livonia	2	1.97 +/- 0.13	1.01 +/- 0.03	0.44 +/- 0.05	7.43	-0.38

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Lostwood	2	1.96 +/- 0.07	0.92 +/- 0.04	0.41 +/- 0.04	8.26	-0.07
Lye Brook Wilderness	2	2.00 +/- 0.10	0.99 +/- 0.03	0.10 +/- 0.15	6.81	-0.20
Mammoth Cave NP	2	1.99 +/- 0.08	0.93 +/- 0.02	0.27 +/- 0.08	6.04	-0.10
Marthas Vineyard	2	2.04 +/- 0.12	1.00 +/- 0.03	0.27 +/- 0.13	6.57	-0.18
Meadview	2	2.23 +/- 0.17	0.74 +/- 0.07	0.76 +/- 0.22	7.15	-0.03
Medicine Lake	2	2.17 +/- 0.07	0.83 +/- 0.04	0.39 +/- 0.07	8.83	-0.10
Mesa Verde NP	2	1.42 +/- 0.15	1.09 +/- 0.12	0.66 +/- 0.37	10.12	0.89
Moosehorn NWR	2	1.87 +/- 0.07	0.91 +/- 0.03	0.89 +/- 0.15	8.04	-0.44
Mount Baldy	2	1.83 +/- 0.10	0.79 +/- 0.07	0.94 +/- 0.17	6.55	-0.05
Mount Rainier NP	2	1.70 +/- 0.06	1.07 +/- 0.07	0.45 +/- 0.21	8.46	0.34
Mount Zirkel Wilderness	2	1.89 +/- 0.12	0.94 +/- 0.08	0.30 +/- 0.18	9.21	-0.46
Nebraska NF	2	2.10 +/- 0.19	1.03 +/- 0.08	0.37 +/- 0.07	10.49	0.01
North Absaroka	2	2.05 +/- 0.08	0.94 +/- 0.07	0.18 +/- 0.14	8.55	-0.14
North Cascades	2	1.90 +/- 0.07	1.02 +/- 0.06	0.81 +/- 0.24	8.42	0.06
Northern Cheyenne	2	2.06 +/- 0.07	0.96 +/- 0.04	0.13 +/- 0.09	7.45	-0.09
Okefenokee NWR	2	1.61 +/- 0.07	1.00 +/- 0.03	0.13 +/- 0.22	6.98	0.01
Olympic	2	1.39 +/- 0.08	1.13 +/- 0.07	0.74 +/- 0.17	7.30	-0.24
Omaha	2	1.91 +/- 0.09	1.03 +/- 0.04	0.53 +/- 0.05	8.00	-0.12
Organ Pipe	2	1.74 +/- 0.16	0.77 +/- 0.07	0.86 +/- 0.23	6.20	0.01
Pasayten	2	1.58 +/- 0.07	1.09 +/- 0.07	0.38 +/- 0.24	10.47	-0.24
Petrified Forest NP	2	1.94 +/- 0.11	0.72 +/- 0.07	1.24 +/- 0.18	6.54	0.20
Phoenix	2	1.23 +/- 0.08	1.11 +/- 0.08	0.21 +/- 0.20	5.95	-0.12
Pinnacles NM	2	1.65 +/- 0.07	0.78 +/- 0.04	1.00 +/- 0.07	6.64	0.36
Point Reyes National Seashore	2	2.06 +/- 0.24	0.69 +/- 0.08	1.12 +/- 0.25	8.31	0.01
Presque Isle	2	1.86 +/- 0.05	0.91 +/- 0.03	0.03 +/- 0.20	5.69	0.03
Proctor Maple R. F.	2	1.95 +/- 0.09	1.01 +/- 0.03	0.25 +/- 0.13	6.81	-0.46
Puget Sound	2	1.48 +/- 0.07	0.98 +/- 0.06	0.44 +/- 0.09	6.25	-0.18
Quabbin Summit	2	2.02 +/- 0.10	0.91 +/- 0.03	0.01 +/- 0.14	6.58	-0.29

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Quaker City	2	1.83 +/- 0.14	1.07 +/- 0.03	0.17 +/- 0.11	7.39	-0.37
Queen Valley	2	1.73 +/- 0.16	0.85 +/- 0.07	0.88 +/- 0.18	6.73	-0.33
Redwood NP	2	1.81 +/- 0.12	0.88 +/- 0.06	0.95 +/- 0.23	7.81	-0.70
Rocky Mountain NP	2	1.67 +/- 0.10	0.94 +/- 0.08	0.47 +/- 0.07	8.43	-0.43
Sac and Fox	2	1.86 +/- 0.07	0.92 +/- 0.03	0.50 +/- 0.04	6.65	-0.43
Saguaro NM	2	1.44 +/- 0.13	0.80 +/- 0.09	1.02 +/- 0.30	7.39	-0.20
Saguaro West	2	1.59 +/- 0.19	0.76 +/- 0.11	0.85 +/- 0.29	6.93	-0.15
San Gabriel	2	2.10 +/- 0.10	0.71 +/- 0.05	0.47 +/- 0.03	7.47	0.05
San Gorgonio Wilderness	2	1.40 +/- 0.12	0.81 +/- 0.09	0.79 +/- 0.03	8.35	0.54
San Pedro Parks	2	1.85 +/- 0.09	0.66 +/- 0.08	1.12 +/- 0.24	8.41	0.03
San Rafael	2	1.81 +/- 0.12	0.81 +/- 0.07	0.71 +/- 0.09	8.42	0.13
Seney	2	1.83 +/- 0.07	0.97 +/- 0.02	0.31 +/- 0.07	6.86	-0.25
Sequoia NP	2	1.84 +/- 0.08	0.75 +/- 0.07	0.69 +/- 0.06	6.59	0.04
Shamrock Mine	2	2.01 +/- 0.11	1.01 +/- 0.07	0.27 +/- 0.23	5.14	-0.22
Shenandoah NP	2	1.76 +/- 0.10	1.13 +/- 0.03	0.21 +/- 0.09	6.69	-0.30
Sierra Ancha	2	1.36 +/- 0.07	0.98 +/- 0.06	0.59 +/- 0.16	6.43	-0.42
Sikes	2	1.79 +/- 0.06	1.06 +/- 0.03	0.02 +/- 0.20	5.60	-0.11
Sipsy Wilderness	2	1.87 +/- 0.07	1.01 +/- 0.02	0.02 +/- 0.11	5.39	0.05
Snoqualmie Pass	2	1.73 +/- 0.09	0.95 +/- 0.09	0.44 +/- 0.17	10.95	-0.88
Starkey	2	1.74 +/- 0.07	0.88 +/- 0.07	0.62 +/- 0.23	9.55	0.22
Sula Peak	2	1.71 +/- 0.07	1.04 +/- 0.11	-0.06 +/- 0.53	10.45	-0.26
Tallgrass	2	1.63 +/- 0.05	1.06 +/- 0.02	0.22 +/- 0.05	5.98	-0.69
Theodore Roosevelt	2	1.83 +/- 0.10	1.03 +/- 0.06	0.60 +/- 0.08	8.73	0.44
Three Sisters Wilderness	2	1.81 +/- 0.06	1.01 +/- 0.06	0.46 +/- 0.20	8.79	-0.06
Thunder Basin	2	1.86 +/- 0.08	0.96 +/- 0.05	0.59 +/- 0.10	6.27	-0.15
Tonto NM	2	1.94 +/- 0.12	0.75 +/- 0.07	0.78 +/- 0.16	5.73	-0.28
Trinity	2	1.80 +/- 0.07	0.81 +/- 0.06	1.08 +/- 0.21	8.68	-0.19
UL Bend	2	1.88 +/- 0.06	0.94 +/- 0.04	0.59 +/- 0.15	7.37	0.03

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Upper Buffalo Wilderness	2	1.80 +/- 0.08	1.01 +/- 0.03	0.54 +/- 0.08	6.49	-0.22
Viking Lake	2	1.59 +/- 0.07	1.05 +/- 0.03	0.77 +/- 0.04	7.92	-0.63
Voyageurs NP #2	2	1.84 +/- 0.05	0.97 +/- 0.03	0.33 +/- 0.09	7.54	-0.34
Weminuche Wilderness	2	1.94 +/- 0.10	0.88 +/- 0.08	0.06 +/- 0.25	7.88	-0.21
Wheeler Peak	2	1.67 +/- 0.13	0.93 +/- 0.11	0.86 +/- 0.34	9.25	0.21
White Mountain	2	1.85 +/- 0.13	1.07 +/- 0.07	0.23 +/- 0.28	7.22	-0.02
White Pass	2	1.78 +/- 0.08	1.21 +/- 0.06	0.33 +/- 0.20	9.65	0.38
White River NF	2	2.14 +/- 0.10	0.87 +/- 0.07	0.28 +/- 0.19	8.32	0.12
Wichita Mountains	2	1.74 +/- 0.10	1.09 +/- 0.04	0.47 +/- 0.09	6.69	-0.54
Wind Cave	2	1.72 +/- 0.06	0.97 +/- 0.03	0.27 +/- 0.05	6.39	-1.02
Yellowstone NP 2	2	1.78 +/- 0.06	0.94 +/- 0.05	0.32 +/- 0.11	7.11	-0.70
Yosemite NP	2	1.87 +/- 0.05	0.86 +/- 0.04	0.72 +/- 0.07	5.53	-0.18
Zion Canyon	2	1.83 +/- 0.11	1.10 +/- 0.07	0.40 +/- 0.31	6.67	-0.31
Acadia NP	3	1.80 +/- 0.06	1.12 +/- 0.02	0.59 +/- 0.31	6.50	-0.15
Addison Pinnacle	3	1.82 +/- 0.11	1.08 +/- 0.02	0.49 +/- 0.46	6.85	-0.30
Agua Tibia	3	1.96 +/- 0.08	0.96 +/- 0.03	0.38 +/- 0.06	5.41	-0.30
Arendtsville	3	1.89 +/- 0.11	1.12 +/- 0.02	-0.06 +/- 0.11	6.22	-0.16
Badlands NP	3	1.80 +/- 0.05	0.94 +/- 0.05	0.22 +/- 0.17	9.80	-0.52
Big Bend NP	3	1.57 +/- 0.17	1.06 +/- 0.03	0.80 +/- 0.25	6.67	0.16
Birmingham	3	1.59 +/- 0.13	1.11 +/- 0.04	-0.48 +/- 0.55	6.10	-0.13
Bliss SP (TRPA)	3	1.79 +/- 0.04	0.88 +/- 0.06	1.18 +/- 0.26	6.44	-0.37
Blue Mounds	3	2.07 +/- 0.05	1.06 +/- 0.03	0.05 +/- 0.09	5.83	-0.42
Bondville	3	1.95 +/- 0.14	1.19 +/- 0.03	-0.05 +/- 0.17	8.03	-0.10
Boundary Waters Canoe Area	3	1.81 +/- 0.04	0.99 +/- 0.04	1.08 +/- 0.61	7.84	-0.64
Bridger Wilderness	3	1.83 +/- 0.05	1.10 +/- 0.09	-0.15 +/- 0.53	9.07	-0.86
Brigantine NWR	3	1.92 +/- 0.12	1.06 +/- 0.03	-0.06 +/- 0.29	8.06	-0.25
Bryce Canyon NP	3	1.44 +/- 0.07	1.13 +/- 0.08	0.88 +/- 0.53	10.89	-0.42
Cabinet Mountains	3	1.84 +/- 0.04	1.02 +/- 0.10	-0.27 +/- 0.41	6.38	-0.45

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Cadiz	3	1.91 +/- 0.17	1.07 +/- 0.03	-0.05 +/- 0.55	7.81	-0.50
Canyonlands NP	3	2.04 +/- 0.08	1.02 +/- 0.07	-0.28 +/- 0.61	8.07	-0.16
Cape Cod	3	1.76 +/- 0.09	1.11 +/- 0.03	-0.09 +/- 0.33	8.50	-0.45
Cape Romain NWR	3	1.74 +/- 0.15	1.03 +/- 0.03	-0.33 +/- 0.40	8.17	-0.16
Cedar Bluff	3	1.78 +/- 0.08	1.05 +/- 0.03	0.14 +/- 0.14	6.06	-0.12
Cherokee Nation	3	1.82 +/- 0.10	1.05 +/- 0.03	0.07 +/- 0.24	5.13	-0.01
Chiricahua NM	3	1.93 +/- 0.13	1.02 +/- 0.04	0.78 +/- 0.43	7.69	-0.09
Columbia Gorge #1	3	1.61 +/- 0.05	0.98 +/- 0.07	0.35 +/- 0.15	7.36	-0.07
Columbia River Gorge	3	1.63 +/- 0.06	0.92 +/- 0.11	0.86 +/- 0.29	7.57	-0.09
Connecticut Hill	3	1.75 +/- 0.11	1.11 +/- 0.02	-0.33 +/- 0.36	6.04	-0.07
Crater Lake NP	3	1.78 +/- 0.04	0.90 +/- 0.11	0.61 +/- 0.77	8.54	-2.03
Craters of the Moon NM	3	1.93 +/- 0.05	0.79 +/- 0.09	-0.13 +/- 0.24	7.42	-0.60
Crescent Lake	3	2.02 +/- 0.05	1.09 +/- 0.04	0.09 +/- 0.08	6.17	-0.87
Dolly Sods Wilderness	3	1.77 +/- 0.12	1.05 +/- 0.02	-0.14 +/- 0.46	6.77	-0.25
Dome Lands Wilderness	3	1.92 +/- 0.07	0.93 +/- 0.07	0.13 +/- 0.10	5.97	-0.46
Douglas	3	1.45 +/- 0.12	1.10 +/- 0.05	-0.42 +/- 0.39	6.26	0.16
El Dorado Springs	3	1.48 +/- 0.09	1.10 +/- 0.03	0.34 +/- 0.23	5.99	-0.03
Ellis	3	1.85 +/- 0.11	1.04 +/- 0.04	0.71 +/- 0.24	6.34	-0.49
Everglades NP	3	1.52 +/- 0.10	1.05 +/- 0.03	0.58 +/- 0.25	7.12	0.05
Fort Peck	3	1.82 +/- 0.04	1.01 +/- 0.05	0.98 +/- 0.34	7.19	-0.13
Frostberg Reservoir (Big Piney Run)	3	1.90 +/- 0.12	0.98 +/- 0.02	0.02 +/- 0.56	4.83	0.10
Gates of the Mountains	3	1.80 +/- 0.04	0.75 +/- 0.10	0.95 +/- 0.34	7.23	-0.68
Gila Wilderness	3	1.50 +/- 0.06	1.06 +/- 0.05	-0.22 +/- 0.84	9.11	-0.28
Great Basin NP	3	1.88 +/- 0.05	0.90 +/- 0.05	0.32 +/- 0.36	7.17	-0.36
Great Gulf Wilderness	3	1.91 +/- 0.08	1.09 +/- 0.02	-1.69 +/- 1.66	6.85	-0.12
Great River Bluffs	3	1.96 +/- 0.05	0.98 +/- 0.02	0.09 +/- 0.09	5.56	-0.50
Great Sand Dunes NM	3	2.12 +/- 0.13	0.90 +/- 0.11	-0.35 +/- 1.30	8.90	-0.40

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Great Smoky Mountains NP	3	1.84 +/- 0.15	1.07 +/- 0.03	0.43 +/- 0.82	6.94	-0.32
Hance Camp at Grand Canyon NP	3	1.52 +/- 0.06	1.13 +/- 0.06	0.93 +/- 0.40	8.28	-0.47
Hells Canyon	3	1.74 +/- 0.03	0.84 +/- 0.11	0.10 +/- 0.53	6.94	-0.25
Hoover	3	1.83 +/- 0.06	0.97 +/- 0.07	0.02 +/- 0.39	8.10	-0.42
Indian Gardens	3	1.65 +/- 0.10	1.09 +/- 0.08	-0.35 +/- 0.51	8.41	-0.49
Isle Royale NP	3	1.93 +/- 0.06	1.09 +/- 0.03	0.55 +/- 0.85	7.46	-0.63
Jarbidge Wilderness	3	1.93 +/- 0.06	0.81 +/- 0.09	0.77 +/- 0.69	6.22	-0.59
Joshua Tree NP	3	2.19 +/- 0.13	0.93 +/- 0.06	0.30 +/- 0.13	7.37	-0.17
Kaiser	3	2.00 +/- 0.05	0.73 +/- 0.06	0.46 +/- 0.07	6.30	-0.56
Kalmiopsis	3	1.65 +/- 0.05	0.83 +/- 0.08	0.96 +/- 0.28	8.59	-0.42
Lassen Volcanic NP	3	1.76 +/- 0.04	0.85 +/- 0.07	1.32 +/- 0.25	7.76	0.88
Lava Beds NM	3	1.75 +/- 0.07	0.98 +/- 0.21	0.35 +/- 1.20	8.65	-0.62
Linville Gorge	3	2.01 +/- 0.15	1.09 +/- 0.03	0.80 +/- 1.24	6.27	-0.32
Livonia	3	2.02 +/- 0.19	1.10 +/- 0.03	0.16 +/- 0.22	8.11	-0.41
Lye Brook Wilderness	3	1.98 +/- 0.08	1.03 +/- 0.02	0.28 +/- 0.38	6.59	0.09
Mammoth Cave NP	3	2.24 +/- 0.15	0.91 +/- 0.03	-0.46 +/- 0.40	5.78	-0.18
Marthas Vineyard	3	1.78 +/- 0.08	1.10 +/- 0.02	0.17 +/- 0.22	6.35	-0.07
Meadview	3	1.80 +/- 0.10	0.98 +/- 0.05	0.57 +/- 0.26	7.45	-0.01
Medicine Lake	3	1.73 +/- 0.04	1.03 +/- 0.05	0.63 +/- 0.29	7.58	0.28
Monture	3	1.88 +/- 0.12	1.37 +/- 0.55	-6.81 +/- 8.29	7.53	-0.05
Moosehorn NWR	3	1.65 +/- 0.06	1.09 +/- 0.03	0.48 +/- 0.46	7.90	-0.44
Mount Baldy	3	1.51 +/- 0.05	1.05 +/- 0.04	0.79 +/- 0.37	6.89	-0.81
Mount Hood	3	1.76 +/- 0.04	1.29 +/- 0.08	0.08 +/- 0.16	8.69	0.63
Mount Rainier NP	3	1.65 +/- 0.06	1.19 +/- 0.08	0.52 +/- 0.30	8.52	0.37
Mount Zirkel Wilderness	3	2.02 +/- 0.07	0.97 +/- 0.07	-0.62 +/- 0.44	6.67	-1.14
Nebraska NF	3	2.03 +/- 0.06	1.11 +/- 0.04	-0.19 +/- 0.14	7.31	-0.51
North Absaroka	3	1.86 +/- 0.05	1.11 +/- 0.09	0.31 +/- 0.42	7.02	-0.96

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
North Cascades	3	1.73 +/- 0.05	1.20 +/- 0.07	0.67 +/- 0.31	7.16	-0.16
Okefenokee NWR	3	1.90 +/- 0.12	0.92 +/- 0.03	-0.28 +/- 0.29	7.30	0.00
Olympic	3	1.68 +/- 0.06	1.00 +/- 0.05	0.71 +/- 0.14	7.76	-0.51
Omaha	3	2.12 +/- 0.08	1.04 +/- 0.03	0.09 +/- 0.14	6.09	-0.46
Organ Pipe	3	1.14 +/- 0.17	0.93 +/- 0.05	1.13 +/- 0.23	7.51	0.25
Pasayten	3	1.66 +/- 0.03	1.17 +/- 0.08	0.20 +/- 0.55	7.68	0.45
Petrified Forest NP	3	1.79 +/- 0.07	1.01 +/- 0.04	0.87 +/- 0.29	7.28	-0.66
Pinnacles NM	3	1.84 +/- 0.04	0.90 +/- 0.03	0.68 +/- 0.08	6.86	0.19
Puget Sound	3	1.53 +/- 0.05	0.87 +/- 0.04	0.71 +/- 0.10	6.04	-0.15
Quabbin Summit	3	1.78 +/- 0.05	1.00 +/- 0.02	-0.15 +/- 0.17	5.62	0.41
Queen Valley	3	1.90 +/- 0.13	0.98 +/- 0.05	0.09 +/- 0.16	7.39	-0.23
Redwood NP	3	1.83 +/- 0.05	0.94 +/- 0.05	0.95 +/- 0.18	7.94	-0.54
Rocky Mountain NP	3	1.97 +/- 0.08	0.84 +/- 0.11	0.18 +/- 0.18	8.26	-0.51
Saguaro NM	3	1.32 +/- 0.12	0.97 +/- 0.06	0.97 +/- 0.28	8.01	0.07
Salt Creek	3	1.78 +/- 0.14	1.00 +/- 0.05	0.35 +/- 0.24	8.47	0.30
San Gabriel	3	2.06 +/- 0.08	0.90 +/- 0.04	0.10 +/- 0.07	7.00	-0.32
San Gorgonio Wilderness	3	1.84 +/- 0.08	0.85 +/- 0.06	0.44 +/- 0.04	6.76	-0.20
San Pedro Parks	3	1.54 +/- 0.08	1.10 +/- 0.06	0.23 +/- 0.49	8.28	-0.08
San Rafael	3	2.00 +/- 0.06	0.85 +/- 0.04	0.57 +/- 0.09	7.33	-0.43
Sawtooth NF	3	1.88 +/- 0.07	1.02 +/- 0.24	-4.18 +/- 2.96	7.19	-0.15
Seney	3	1.61 +/- 0.06	1.04 +/- 0.03	0.77 +/- 0.45	7.71	-0.67
Sequoia NP	3	1.87 +/- 0.06	0.87 +/- 0.07	0.28 +/- 0.12	5.92	0.11
Shamrock Mine	3	2.02 +/- 0.07	1.06 +/- 0.07	-0.25 +/- 0.56	6.01	-0.59
Shenandoah NP	3	1.76 +/- 0.13	1.11 +/- 0.02	0.53 +/- 0.40	7.50	-0.17
Shining Rock Wilderness	3	2.08 +/- 0.20	0.98 +/- 0.03	-0.82 +/- 0.80	6.68	-0.30
Sierra Ancha	3	1.32 +/- 0.07	0.99 +/- 0.06	0.41 +/- 0.38	8.03	-0.21
Sikes	3	1.88 +/- 0.09	1.03 +/- 0.03	-0.37 +/- 0.36	5.79	-0.14
Sipsy Wilderness	3	2.27 +/- 0.11	0.92 +/- 0.03	-0.53 +/- 0.41	5.54	-0.20

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Snoqualmie Pass	3	1.79 +/- 0.06	1.04 +/- 0.08	0.39 +/- 0.20	8.20	-0.60
Starkey	3	1.75 +/- 0.04	0.75 +/- 0.11	0.97 +/- 0.43	8.04	0.15
Swanquarter	3	1.83 +/- 0.09	1.02 +/- 0.02	0.45 +/- 0.27	5.46	0.01
Sycamore Canyon	3	1.32 +/- 0.08	1.03 +/- 0.11	1.05 +/- 0.58	7.34	0.26
Tallgrass	3	1.51 +/- 0.09	1.06 +/- 0.03	0.69 +/- 0.34	6.37	-0.10
Three Sisters Wilderness	3	1.80 +/- 0.04	1.05 +/- 0.08	0.44 +/- 0.33	5.43	-0.64
Tonto NM	3	1.75 +/- 0.10	0.98 +/- 0.05	0.45 +/- 0.28	7.33	-0.54
UL Bend	3	1.89 +/- 0.05	1.09 +/- 0.07	-1.04 +/- 0.94	7.52	-0.58
Upper Buffalo Wilderness	3	1.88 +/- 0.08	1.01 +/- 0.02	0.45 +/- 0.27	6.04	-0.35
Viking Lake	3	1.81 +/- 0.07	1.12 +/- 0.03	0.13 +/- 0.12	5.95	-0.10
Weminuche Wilderness	3	1.80 +/- 0.08	1.15 +/- 0.09	-0.92 +/- 0.94	9.46	-0.82
Wheeler Peak	3	1.63 +/- 0.11	1.15 +/- 0.08	0.49 +/- 0.45	9.63	0.18
White Mountain	3	1.71 +/- 0.11	1.10 +/- 0.04	-0.04 +/- 0.32	7.19	-0.47
White Pass	3	1.81 +/- 0.06	1.27 +/- 0.10	-0.22 +/- 0.35	10.27	0.11
White River NF	3	1.80 +/- 0.08	1.17 +/- 0.08	0.74 +/- 0.64	8.89	-0.42
Wichita Mountains	3	1.69 +/- 0.09	1.08 +/- 0.03	0.83 +/- 0.18	5.32	-0.06
Wind Cave	3	1.83 +/- 0.04	0.98 +/- 0.05	-0.48 +/- 0.42	6.44	-0.66
Yellowstone NP 2	3	1.80 +/- 0.04	0.98 +/- 0.11	-0.46 +/- 0.51	8.58	-0.81
Yosemite NP	3	1.67 +/- 0.04	1.01 +/- 0.07	0.96 +/- 0.28	7.49	0.11
Zion Canyon	3	1.83 +/- 0.07	1.08 +/- 0.04	0.73 +/- 0.34	6.28	-0.42
Acadia NP	4	1.56 +/- 0.09	0.99 +/- 0.04	0.55 +/- 0.11	8.50	-0.37
Agua Tibia	4	1.54 +/- 0.09	0.95 +/- 0.06	0.63 +/- 0.04	9.37	-0.70
Arendtsville	4	1.49 +/- 0.11	0.96 +/- 0.03	0.89 +/- 0.05	7.98	-0.74
Bandelier NM	4	1.43 +/- 0.05	1.06 +/- 0.04	0.39 +/- 0.08	9.10	-0.46
Big Bend NP	4	2.00 +/- 0.14	1.01 +/- 0.03	-0.22 +/- 0.21	8.13	-0.26
Birmingham	4	1.38 +/- 0.05	1.07 +/- 0.04	0.75 +/- 0.08	5.11	-0.10
Bliss SP (TRPA)	4	1.59 +/- 0.04	1.11 +/- 0.06	0.38 +/- 0.14	9.77	-0.70
Blue Mounds	4	1.57 +/- 0.09	0.96 +/- 0.05	1.11 +/- 0.03	6.56	0.13

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Bondville	4	1.31 +/- 0.10	0.98 +/- 0.04	1.05 +/- 0.03	7.49	-0.29
Bosque del Apache	4	1.15 +/- 0.08	1.01 +/- 0.04	0.83 +/- 0.08	8.70	0.11
Boundary Waters Canoe Area	4	1.65 +/- 0.11	0.97 +/- 0.05	0.80 +/- 0.04	8.87	0.26
Bridger Wilderness	4	1.71 +/- 0.12	1.12 +/- 0.08	0.18 +/- 0.21	14.92	-0.88
Bridgton	4	1.60 +/- 0.08	1.06 +/- 0.04	0.19 +/- 0.14	8.38	-0.42
Brigantine NWR	4	1.50 +/- 0.10	1.01 +/- 0.03	0.71 +/- 0.06	6.71	-0.33
Bryce Canyon NP	4	1.49 +/- 0.10	1.07 +/- 0.07	0.51 +/- 0.08	13.47	-2.55
Cabinet Mountains	4	1.66 +/- 0.04	1.01 +/- 0.09	0.41 +/- 0.18	9.42	-0.31
Cadiz	4	1.75 +/- 0.06	0.90 +/- 0.03	0.75 +/- 0.03	6.50	-0.04
Caney Creek	4	1.74 +/- 0.06	1.01 +/- 0.03	0.32 +/- 0.04	7.02	-0.22
Canyonlands NP	4	1.84 +/- 0.15	1.03 +/- 0.07	0.45 +/- 0.08	11.88	0.45
Capitol Reef NP	4	1.96 +/- 0.13	0.86 +/- 0.08	0.53 +/- 0.06	11.25	-0.63
Casco Bay	4	1.50 +/- 0.05	1.00 +/- 0.03	0.66 +/- 0.10	7.34	-0.59
Cedar Bluff	4	1.03 +/- 0.39	1.12 +/- 0.23	1.01 +/- 0.09	14.58	-0.05
Chassahowitzka NWR	4	1.67 +/- 0.05	0.89 +/- 0.02	0.12 +/- 0.10	4.79	0.01
Cherokee Nation	4	1.39 +/- 0.08	1.03 +/- 0.04	0.91 +/- 0.03	7.89	-0.59
Chiricahua NM	4	1.24 +/- 0.10	1.23 +/- 0.04	0.21 +/- 0.14	9.28	1.55
Cloud Peak	4	2.11 +/- 0.13	0.94 +/- 0.07	0.50 +/- 0.25	17.70	-0.45
Cohutta	4	1.83 +/- 0.08	1.00 +/- 0.03	-0.01 +/- 0.06	6.22	-0.44
Columbia Gorge #1	4	1.60 +/- 0.04	0.68 +/- 0.07	0.72 +/- 0.05	8.73	-0.94
Columbia River Gorge	4	1.46 +/- 0.05	0.84 +/- 0.07	0.73 +/- 0.04	9.09	-0.65
Connecticut Hill	4	1.83 +/- 0.18	0.91 +/- 0.04	0.70 +/- 0.06	7.38	-0.89
Crater Lake NP	4	1.68 +/- 0.06	1.22 +/- 0.08	-0.32 +/- 0.25	12.88	-1.87
Craters of the Moon NM	4	1.68 +/- 0.10	1.10 +/- 0.10	0.47 +/- 0.04	12.35	-0.68
Crescent Lake	4	1.90 +/- 0.19	0.98 +/- 0.11	0.95 +/- 0.04	10.63	0.76
Death Valley NP	4	1.84 +/- 0.15	1.12 +/- 0.07	0.09 +/- 0.11	9.79	-0.52
Dolly Sods Wilderness	4	1.43 +/- 0.07	1.02 +/- 0.02	0.39 +/- 0.08	7.23	-0.24
Douglas	4	1.29 +/- 0.14	0.87 +/- 0.13	0.87 +/- 0.28	6.03	-0.09

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
El Dorado Springs	4	1.37 +/- 0.08	0.97 +/- 0.04	0.85 +/- 0.03	7.91	-0.29
Ellis	4	1.53 +/- 0.10	1.04 +/- 0.05	0.86 +/- 0.03	8.74	-1.10
Everglades NP	4	1.52 +/- 0.11	1.09 +/- 0.03	0.17 +/- 0.19	7.48	-0.24
Flathead	4	1.66 +/- 0.04	1.12 +/- 0.06	-0.05 +/- 0.11	8.03	-0.57
Fort Peck	4	1.44 +/- 0.09	0.96 +/- 0.05	0.76 +/- 0.04	10.16	-0.36
Frostberg Reservoir (Big Piney Run)	4	1.85 +/- 0.09	0.93 +/- 0.02	0.33 +/- 0.06	5.15	-0.27
Gates of the Mountains	4	1.67 +/- 0.06	1.25 +/- 0.07	-0.18 +/- 0.16	11.47	0.06
Gila Wilderness	4	1.61 +/- 0.09	0.97 +/- 0.05	0.43 +/- 0.31	9.70	-0.16
Glacier NP	4	1.65 +/- 0.03	0.87 +/- 0.06	0.94 +/- 0.07	6.97	-0.47
Great Basin NP	4	1.01 +/- 0.09	1.33 +/- 0.07	0.24 +/- 0.13	14.95	-1.22
Great Gulf Wilderness	4	1.88 +/- 0.13	0.91 +/- 0.04	0.26 +/- 0.10	8.58	0.22
Great River Bluffs	4	1.45 +/- 0.15	0.86 +/- 0.07	0.91 +/- 0.03	10.20	-0.11
Great Sand Dunes NM	4	1.43 +/- 0.09	1.06 +/- 0.06	0.09 +/- 0.18	11.34	-1.24
Great Smoky Mountains NP	4	1.83 +/- 0.08	0.96 +/- 0.03	0.36 +/- 0.07	6.90	-0.32
Guadalupe Mountains NP	4	1.55 +/- 0.18	1.04 +/- 0.07	0.49 +/- 0.08	9.50	-0.70
Hance Camp at Grand Canyon NP	4	1.54 +/- 0.07	1.16 +/- 0.05	0.59 +/- 0.09	11.29	0.25
Hells Canyon	4	1.64 +/- 0.05	1.17 +/- 0.10	0.51 +/- 0.04	9.27	-0.97
Hercules-Glades	4	1.62 +/- 0.07	0.96 +/- 0.03	0.55 +/- 0.03	7.54	-0.70
Hoover	4	1.63 +/- 0.06	1.13 +/- 0.07	0.24 +/- 0.15	12.62	0.88
Indian Gardens	4	1.69 +/- 0.06	1.09 +/- 0.04	0.36 +/- 0.11	7.55	0.29
Isle Royale NP	4	1.62 +/- 0.11	1.05 +/- 0.04	0.88 +/- 0.03	8.24	0.70
James River Face Wilderness	4	1.61 +/- 0.05	0.99 +/- 0.02	0.42 +/- 0.07	6.20	-0.18
Jarbidge Wilderness	4	1.55 +/- 0.09	1.26 +/- 0.09	0.77 +/- 0.07	13.67	-0.40
Joshua Tree NP	4	1.64 +/- 0.12	0.95 +/- 0.07	0.64 +/- 0.03	10.63	-0.56
Kaiser	4	1.74 +/- 0.07	0.86 +/- 0.09	0.68 +/- 0.06	13.01	-1.63
Kalmiopsis	4	1.55 +/- 0.03	1.14 +/- 0.11	0.25 +/- 0.40	8.92	-0.61

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Lassen Volcanic NP	4	1.39 +/- 0.05	1.31 +/- 0.08	0.28 +/- 0.09	11.99	-0.61
Lava Beds NM	4	1.53 +/- 0.04	1.25 +/- 0.08	0.25 +/- 0.08	9.43	-1.11
Linville Gorge	4	1.66 +/- 0.06	1.04 +/- 0.02	0.38 +/- 0.10	6.41	-0.01
Livonia	4	1.48 +/- 0.12	1.04 +/- 0.04	0.82 +/- 0.04	8.19	-0.69
Lostwood	4	1.25 +/- 0.08	0.96 +/- 0.05	0.99 +/- 0.04	9.37	-1.06
M.K. Goddard	4	1.34 +/- 0.07	1.00 +/- 0.03	0.84 +/- 0.03	6.18	-0.16
Mammoth Cave NP	4	1.69 +/- 0.10	0.95 +/- 0.04	0.45 +/- 0.04	7.97	-0.34
Meadview	4	2.14 +/- 0.17	0.91 +/- 0.07	0.17 +/- 0.11	8.99	0.14
Medicine Lake	4	1.55 +/- 0.11	1.05 +/- 0.05	0.72 +/- 0.05	11.09	0.03
Monture	4	1.59 +/- 0.04	1.19 +/- 0.08	0.11 +/- 0.22	9.89	-0.92
Moosehorn NWR	4	1.57 +/- 0.08	0.95 +/- 0.03	0.52 +/- 0.12	7.84	0.01
Mount Baldy	4	1.39 +/- 0.04	1.05 +/- 0.03	0.44 +/- 0.19	8.88	-0.29
Mount Hood	4	1.66 +/- 0.04	1.20 +/- 0.06	0.54 +/- 0.11	11.27	0.23
Mount Rainier NP	4	1.56 +/- 0.04	1.38 +/- 0.10	0.04 +/- 0.28	10.28	0.75
Mount Zirkel Wilderness	4	2.32 +/- 0.15	0.72 +/- 0.07	0.09 +/- 0.18	14.94	-2.20
Nebraska NF	4	1.51 +/- 0.11	1.12 +/- 0.07	0.72 +/- 0.04	10.13	-0.15
North Absaroka	4	1.83 +/- 0.12	0.99 +/- 0.09	0.41 +/- 0.16	14.77	-0.55
North Cascades	4	1.75 +/- 0.05	1.07 +/- 0.07	0.83 +/- 0.29	10.76	0.23
Northern Cheyenne	4	1.71 +/- 0.07	1.16 +/- 0.06	0.16 +/- 0.11	11.39	0.31
Okefenokee NWR	4	1.60 +/- 0.07	1.04 +/- 0.03	-0.23 +/- 0.19	7.07	-0.02
Olympic	4	1.60 +/- 0.06	1.14 +/- 0.09	0.40 +/- 0.12	9.79	-0.28
Omaha	4	1.47 +/- 0.11	1.02 +/- 0.05	1.06 +/- 0.03	7.15	0.07
Organ Pipe	4	1.46 +/- 0.12	1.15 +/- 0.04	0.29 +/- 0.08	6.92	0.22
Pasayten	4	1.70 +/- 0.05	1.23 +/- 0.06	0.11 +/- 0.12	11.26	0.60
Petrified Forest NP	4	1.57 +/- 0.08	1.12 +/- 0.05	0.37 +/- 0.14	8.12	-0.09
Phoenix	4	1.24 +/- 0.04	0.97 +/- 0.07	0.69 +/- 0.05	5.41	0.17
Pinnacles NM	4	1.55 +/- 0.06	1.02 +/- 0.07	0.60 +/- 0.04	9.07	-0.53
Point Reyes National Seashore	4	1.63 +/- 0.11	1.03 +/- 0.07	0.65 +/- 0.04	8.85	-0.70

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Presque Isle	4	1.71 +/- 0.06	0.89 +/- 0.03	0.31 +/- 0.10	6.85	-0.06
Proctor Maple R. F.	4	1.65 +/- 0.09	0.99 +/- 0.03	0.60 +/- 0.06	8.08	-0.37
Quabbin Summit	4	1.65 +/- 0.12	0.96 +/- 0.04	0.38 +/- 0.09	8.01	-0.10
Quaker City	4	1.61 +/- 0.09	0.98 +/- 0.03	0.67 +/- 0.04	6.48	-0.20
Queen Valley	4	1.34 +/- 0.09	1.07 +/- 0.04	0.69 +/- 0.03	8.04	0.18
Redwood NP	4	1.70 +/- 0.05	1.09 +/- 0.10	0.45 +/- 0.24	7.57	-1.08
Rocky Mountain NP	4	1.44 +/- 0.11	0.96 +/- 0.07	0.64 +/- 0.07	14.04	-2.04
Sac and Fox	4	1.38 +/- 0.10	0.94 +/- 0.05	0.95 +/- 0.03	9.43	-0.47
Saguaro NM	4	1.42 +/- 0.09	0.91 +/- 0.04	0.37 +/- 0.04	7.24	0.13
Saguaro West	4	1.27 +/- 0.09	0.94 +/- 0.06	0.41 +/- 0.04	6.06	-0.21
San Gabriel	4	1.64 +/- 0.10	0.83 +/- 0.08	0.53 +/- 0.04	11.87	-0.40
San Gorgonio Wilderness	4	1.20 +/- 0.11	0.87 +/- 0.08	0.87 +/- 0.03	11.27	-1.21
San Pedro Parks	4	1.38 +/- 0.09	1.13 +/- 0.06	0.20 +/- 0.17	12.38	0.76
San Rafael	4	1.50 +/- 0.10	0.94 +/- 0.07	0.60 +/- 0.04	10.81	-1.19
Seney	4	1.31 +/- 0.11	0.97 +/- 0.04	0.73 +/- 0.03	8.60	0.61
Shamrock Mine	4	1.72 +/- 0.08	1.11 +/- 0.05	0.23 +/- 0.12	8.15	-0.74
Shenandoah NP	4	1.66 +/- 0.10	0.97 +/- 0.03	0.60 +/- 0.05	7.42	0.03
Shining Rock Wilderness	4	1.59 +/- 0.14	0.98 +/- 0.04	0.35 +/- 0.17	9.40	-0.92
Sierra Ancha	4	1.36 +/- 0.06	1.00 +/- 0.05	0.25 +/- 0.09	8.78	-0.43
Sikes	4	1.66 +/- 0.06	1.09 +/- 0.03	0.03 +/- 0.12	6.96	-0.35
Sipsy Wilderness	4	1.80 +/- 0.06	0.94 +/- 0.03	0.36 +/- 0.04	5.81	-0.15
Snoqualmie Pass	4	1.71 +/- 0.05	1.05 +/- 0.07	0.28 +/- 0.07	10.01	-0.72
St. Marks	4	1.69 +/- 0.08	1.05 +/- 0.03	-0.23 +/- 0.23	7.05	-0.12
Starkey	4	1.50 +/- 0.03	1.09 +/- 0.07	0.76 +/- 0.04	8.75	-0.05
Sula Peak	4	1.62 +/- 0.05	1.08 +/- 0.07	0.22 +/- 0.14	11.05	-1.08
Sycamore Canyon	4	1.26 +/- 0.05	1.12 +/- 0.06	0.31 +/- 0.12	7.35	-0.30
Tallgrass	4	1.47 +/- 0.09	0.98 +/- 0.04	0.81 +/- 0.03	8.73	-0.70
Theodore Roosevelt	4	1.54 +/- 0.08	0.93 +/- 0.05	0.99 +/- 0.04	8.36	0.15

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)
Three Sisters Wilderness	4	1.64 +/- 0.04	1.24 +/- 0.08	0.38 +/- 0.15	9.69	-0.99
Thunder Basin	4	1.78 +/- 0.08	0.92 +/- 0.06	0.61 +/- 0.07	7.24	-0.27
Tonto NM	4	1.62 +/- 0.07	1.06 +/- 0.04	0.23 +/- 0.05	6.71	-0.32
Trinity	4	1.56 +/- 0.04	1.19 +/- 0.10	0.37 +/- 0.08	9.89	-0.92
UL Bend	4	1.85 +/- 0.07	0.89 +/- 0.05	0.91 +/- 0.05	9.37	-0.25
Upper Buffalo Wilderness	4	1.64 +/- 0.08	0.99 +/- 0.04	0.65 +/- 0.04	7.83	-0.49
Viking Lake	4	1.44 +/- 0.09	1.03 +/- 0.04	1.03 +/- 0.02	6.44	-0.11
Voyageurs NP #2	4	1.50 +/- 0.08	0.91 +/- 0.04	1.00 +/- 0.03	8.54	-0.51
Weminuche Wilderness	4	1.35 +/- 0.12	1.05 +/- 0.07	0.69 +/- 0.27	13.73	-0.55
Wheeler Peak	4	1.81 +/- 0.08	0.98 +/- 0.05	0.14 +/- 0.16	11.14	-1.58
White Mountain	4	1.40 +/- 0.10	1.12 +/- 0.05	0.64 +/- 0.07	9.57	-0.01
White Pass	4	1.90 +/- 0.08	1.20 +/- 0.07	-0.19 +/- 0.21	14.93	-1.02
White River NF	4	1.62 +/- 0.17	1.16 +/- 0.08	0.20 +/- 0.33	16.34	-0.39
Wichita Mountains	4	1.62 +/- 0.09	1.04 +/- 0.04	0.70 +/- 0.03	8.41	-1.04
Wind Cave	4	1.32 +/- 0.08	1.08 +/- 0.07	0.37 +/- 0.08	11.75	0.23
Yellowstone NP 2	4	1.51 +/- 0.08	1.11 +/- 0.08	0.52 +/- 0.07	12.26	-1.93
Zion Canyon	4	1.78 +/- 0.07	1.09 +/- 0.05	0.29 +/- 0.07	8.62	-0.58

Table S6. Quarter-specific regression results flagged for single outlier year (n = 28) or temporal trend (n = 7) in residual errors. Values in parentheses represent regression results when the outlier year was removed from the dataset. Regressions for which all coefficients changed by less than 0.1 when the outlier year was removed are highlighted in gray. These 10 cases are regarded as high confidence results so they also appear in Table S5.

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)	outlier year
Agua Tibia	1	1.38 +/- 0.12 (1.30 +/- 0.16)	1.08 +/- 0.08 (1.15 +/- 0.09)	0.56 +/- 0.04 (0.52 +/- 0.05)	9.52	0.07	2002
Salt Creek	1	0.74 +/- 0.29 (0.71 +/- 0.17)	1.22 +/- 0.12 (1.03 +/- 0.07)	0.98 +/- 0.07 (1.08 +/- 0.05)	10.70	-0.26	2004
Shamrock Mine	1	1.52 +/- 0.16 (1.64 +/- 0.05)	1.06 +/- 0.07 (0.94 +/- 0.03)	0.43 +/- 0.11 (0.22 +/- 0.14)	10.59	0.08	2002
St. Marks	1	1.63 +/- 0.06	1.02 +/- 0.03	0.06 +/- 0.19	7.53	-0.47	N/A
Big Bend NP	2	2.16 +/- 0.12 (2.18 +/- 0.13)	0.96 +/- 0.03 (0.93 +/- 0.03)	0.24 +/- 0.28 (0.34 +/- 0.26)	6.72	0.19	2002
Blue Mounds	2	1.67 +/- 0.09 (1.67 +/- 0.10)	0.95 +/- 0.04 (1.00 +/- 0.04)	0.74 +/- 0.04 (0.63 +/- 0.05)	9.62	-0.23	2003
Boundary Waters Canoe Area	2	1.69 +/- 0.06 (1.98 +/- 0.06)	1.06 +/- 0.03 (0.94 +/- 0.03)	0.20 +/- 0.10 (0.28 +/- 0.09)	8.25	-0.01	2003
Brigantine NWR	2	2.07 +/- 0.11 (2.17 +/- 0.13)	0.99 +/- 0.03 (0.97 +/- 0.03)	0.35 +/- 0.13 (0.31 +/- 0.13)	6.98	-0.26	2008
Bryce Canyon NP	2	1.62 +/- 0.06 (1.60 +/- 0.06)	0.97 +/- 0.05 (1.08 +/- 0.05)	0.27 +/- 0.15 (0.19 +/- 0.16)	7.56	-0.21	2005
Caney Creek	2	1.87 +/- 0.10 (1.87 +/- 0.09)	1.03 +/- 0.03 (1.00 +/- 0.03)	0.01 +/- 0.16 (-0.08 +/- 0.14)	6.51	-0.15	2002
Connecticut Hill	2	1.28 +/- 0.14	1.15 +/- 0.03	0.40 +/- 0.10	7.92	-0.64	N/A
El Dorado Springs	2	1.78 +/- 0.07 (1.78 +/- 0.06)	0.98 +/- 0.03 (0.96 +/- 0.02)	0.27 +/- 0.06 (0.29 +/- 0.05)	6.24	-0.24	2002
Fort Peck	2	1.95 +/- 0.08 (1.91 +/- 0.08)	0.90 +/- 0.04 (0.92 +/- 0.04)	0.28 +/- 0.06 (0.28 +/- 0.06)	7.60	0.41	2002
Great Gulf	2	1.70 +/- 0.07 (1.63 +/- 0.07)	1.06 +/- 0.03 (1.06 +/- 0.03)	0.04 +/- 0.33 (0.10 +/- 0.34)	7.36	-0.15	2007

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)	outlier year
Wilderness							
Monture	2	1.52 +/- 0.05 (1.65 +/- 0.06)	0.90 +/- 0.08 (0.96 +/- 0.08)	2.23 +/- 0.79 (1.80 +/- 0.72)	10.09	0.43	2004
New York City	2	1.75 +/- 0.14	0.99 +/- 0.05	0.58 +/- 0.08	6.91	-0.62	N/A
Northern Cheyenne	2	2.06 +/- 0.07 (2.07 +/- 0.07)	0.96 +/- 0.04 (0.93 +/- 0.03)	0.13 +/- 0.09 (0.14 +/- 0.07)	7.45	-0.09	2002
Okefenokee NWR	2	1.61 +/- 0.07 (1.65 +/- 0.06)	1.00 +/- 0.03 (1.02 +/- 0.03)	0.13 +/- 0.22 (0.12 +/- 0.19)	6.98	0.01	2005
Swanquarter	2	2.03 +/- 0.08	0.93 +/- 0.02	0.28 +/- 0.14	6.02	-0.61	N/A
Boundary Waters Canoe Area	3	1.81 +/- 0.04 (1.89 +/- 0.05)	0.99 +/- 0.04 (0.96 +/- 0.03)	1.08 +/- 0.61 (1.03 +/- 0.53)	7.84	-0.64	2003
Bridgton	3	1.84 +/- 0.06 (1.89 +/- 0.07)	1.14 +/- 0.02 (1.11 +/- 0.03)	-0.48 +/- 0.45 (-0.73 +/- 0.45)	6.90	-0.22	2002
Cadiz	3	1.91 +/- 0.17 (1.83 +/- 0.16)	1.07 +/- 0.03 (1.07 +/- 0.03)	-0.05 +/- 0.55 (0.02 +/- 0.59)	7.81	-0.50	2002
Caney Creek	3	1.81 +/- 0.08 (1.82 +/- 0.07)	0.97 +/- 0.03 (0.96 +/- 0.02)	0.45 +/- 0.28 (0.31 +/- 0.24)	6.07	-0.18	2002
Chassahowitzka NWR	3	2.11 +/- 0.17	0.93 +/- 0.04	-0.56 +/- 0.35	6.98	-0.21	N/A
Guadalupe Mountains NP	3	1.76 +/- 0.13 (1.93 +/- 0.13)	0.98 +/- 0.04 (0.97 +/- 0.04)	0.57 +/- 0.29 (0.67 +/- 0.31)	7.13	-0.19	2005
Ikes Backbone	3	1.33 +/- 0.10 (1.61 +/- 0.12)	1.01 +/- 0.08 (1.03 +/- 0.08)	0.56 +/- 0.34 (0.28 +/- 0.33)	10.26	-0.24	2005
Mesa Verde NP	3	1.90 +/- 0.09 (1.97 +/- 0.08)	1.11 +/- 0.08 (1.07 +/- 0.08)	-0.29 +/- 0.52 (0.05 +/- 0.64)	9.03	0.14	2002
Phoenix	3	1.19 +/- 0.07 (1.45 +/- 0.07)	1.08 +/- 0.06 (0.99 +/- 0.05)	0.29 +/- 0.18 (0.22 +/- 0.15)	6.20	0.03	2004
Thunder Basin	3	1.88 +/- 0.04	1.14 +/- 0.06	-0.74 +/- 0.50	5.87	-0.78	N/A
Badlands NP	4	1.22 +/- 0.08 (1.35 +/- 0.08)	1.02 +/- 0.06 (0.93 +/- 0.05)	0.49 +/- 0.08 (0.60 +/- 0.07)	11.60	-0.34	2007
Ikes Backbone	4	1.28 +/- 0.06 (1.62 +/- 0.08)	1.06 +/- 0.05 (1.04 +/- 0.05)	0.44 +/- 0.05 (0.32 +/- 0.05)	10.74	0.90	2005
Mesa Verde NP	4	1.79 +/- 0.13 (1.65 +/- 0.11)	1.14 +/- 0.08 (1.21 +/- 0.06)	0.32 +/- 0.12 (0.23 +/- 0.13)	11.63	0.32	2002

Site	quarter	βoc	βsulf	βnit	NME (%)	NMB (%)	outlier year
Salt Creek	4	1.17 +/- 0.20 (1.26 +/- 0.15)	1.16 +/- 0.09 (0.99 +/- 0.07)	0.93 +/- 0.08 (1.03 +/- 0.06)	10.45	-0.17	2003
Swanquarter	4	1.67 +/- 0.12 (1.62 +/- 0.12)	0.98 +/- 0.03 (0.99 +/- 0.03)	0.02 +/- 0.15 (-0.08 +/- 0.15)	6.28	-0.13	2008
Yosemite NP	4	1.56 +/- 0.03	1.05 +/- 0.06	0.59 +/- 0.05	8.93	-1.00	N/A

Table S7. Quarter-specific regression results flagged because of physically unrealistic coefficients. Four cases with a * are already flagged as low confidence due to an influential outlier year or temporal trend in ε_i (see Table S6). New York City and Washington D.C. regressions from all quarters are included here because their multiyear β_{soil} values are physically unrealistic (See Table S2).

Site	quarter	β_{oc}	β_{sulf}	β_{nit}	NME (%)	NMB (%)
Badlands NP	1	**0.54 +/- 0.07**	1.03 +/- 0.04	0.72 +/- 0.08	12.00	1.42
Cloud Peak	1	**0.93 +/- 0.30**	0.98 +/- 0.09	0.83 +/- 0.13	16.48	0.56
Fort Peck	1	**0.99 +/- 0.17**	0.87 +/- 0.05	0.90 +/- 0.04	11.01	-0.93
Hance Camp at Grand Canyon NP	1	**0.45 +/- 0.08**	1.51 +/- 0.05	0.59 +/- 0.06	11.21	1.56
Ikes Backbone	1	**0.76 +/- 0.08**	1.34 +/- 0.07	0.34 +/- 0.05	11.66	1.11
Kalmiopsis	1	1.37 +/- 0.03	0.75 +/- 0.08	**1.66 +/- 0.30**	9.25	-0.75
Meadview	1	**0.87 +/- 0.13**	1.33 +/- 0.07	0.39 +/- 0.07	10.61	1.72
New York City	1	1.59 +/- 0.14	0.76 +/- 0.05	1.34 +/- 0.06	4.82	-0.02
North Absaroka	1	**0.90 +/- 0.19**	1.17 +/- 0.07	0.61 +/- 0.09	14.93	-0.26
Petrified Forest NP	1	**0.96 +/- 0.18**	0.90 +/- 0.12	1.05 +/- 0.20	14.88	2.40
Phoenix	1	1.31 +/- 0.04	**0.57 +/- 0.10**	0.67 +/- 0.04	6.01	0.20
Rocky Mountain NP	1	**0.64 +/- 0.17**	1.00 +/- 0.09	0.89 +/- 0.06	13.72	-0.24
Saguaro NM	1	**0.93 +/- 0.13**	1.11 +/- 0.08	0.50 +/- 0.07	8.59	1.52
Salt Creek*	1	**0.74 +/- 0.29**	1.22 +/- 0.12	0.98 +/- 0.07	10.70	-0.26
San Gorgonio Wilderness	1	**0.81 +/- 0.15**	0.84 +/- 0.09	0.93 +/- 0.03	9.02	0.40

Site	quarter	β_{oc}	β_{sulf}	β_{nit}	NME (%)	NMB (%)
San Rafael	1	**0.92 +/- 0.08**	0.94 +/- 0.07	0.77 +/- 0.06	11.48	-0.40
Sequoia NP	1	**0.72 +/- 0.14**	0.84 +/- 0.13	1.10 +/- 0.03	9.20	-0.22
Snoqualmie Pass	1	**0.23 +/- 0.06**	**1.55 +/- 0.09**	1.02 +/- 0.07	19.60	0.86
Washington D.C.	1	1.55 +/- 0.06	0.88 +/- 0.03	1.08 +/- 0.04	5.59	-0.23
Wheeler Peak	1	**0.58 +/- 0.26**	1.27 +/- 0.17	0.54 +/- 0.23	15.48	0.02
Yellowstone NP 2	1	**0.92 +/- 0.11**	1.00 +/- 0.06	0.74 +/- 0.05	12.55	-1.07
Bridgton	2	1.85 +/- 0.08	1.01 +/- 0.03	**-0.43 +/- 0.23**	7.50	-0.23
Chiricahua NM	2	**1.00 +/- 0.10**	1.15 +/- 0.06	0.88 +/- 0.26	7.17	0.01
Death Valley NP	2	1.82 +/- 0.16	0.71 +/- 0.08	**1.41 +/- 0.28**	6.26	-0.26
Douglas	2	1.04 +/- 0.17	1.10 +/- 0.14	**1.49 +/- 0.56**	5.27	0.36
Gila Wilderness	2	1.64 +/- 0.06	**0.52 +/- 0.07**	**2.43 +/- 0.29**	7.25	-0.37
Hells Canyon	2	1.56 +/- 0.07	0.81 +/- 0.11	**1.87 +/- 0.65**	9.47	-0.96
Mohawk Mt.	2	**0.76 +/- 0.12**	1.18 +/- 0.04	0.72 +/- 0.30	11.47	0.95
Monture*	2	1.52 +/- 0.05	0.90 +/- 0.08	**2.23 +/- 0.79**	10.09	0.43
New York City*	2	1.75 +/- 0.14	0.99 +/- 0.05	0.58 +/- 0.08	6.91	-0.62
Sawtooth NF	2	1.65 +/- 0.07	1.03 +/- 0.13	**-2.00 +/- 1.03**	9.74	-1.19
Shining Rock Wilderness	2	1.88 +/- 0.17	1.05 +/- 0.04	**-0.64 +/- 0.33**	7.89	-0.68
Sycamore Canyon	2	1.59 +/- 0.10	0.73 +/- 0.09	**1.52 +/- 0.22**	6.20	-0.15
Washington D.C.	2	2.05 +/- 0.10	1.00 +/- 0.03	0.52 +/- 0.07	6.91	-0.52
Bandelier NM	3	1.83 +/- 0.08	1.04 +/- 0.05	**-0.96 +/- 0.56**	7.76	-0.58
Capitol Reef NP	3	2.04 +/- 0.07	1.04 +/- 0.05	**-0.81 +/- 0.46**	7.22	-0.86
Casco Bay	3	1.69 +/- 0.05	1.29 +/- 0.03	**-0.76 +/- 0.25**	7.21	-0.65
Chassahowitzka NWR*	3	2.11 +/- 0.17	0.93 +/- 0.04	**-0.56 +/- 0.35**	6.98	-0.21
Cloud Peak	3	1.96 +/- 0.04	1.14 +/- 0.07	**-0.89 +/- 0.30**	6.36	-0.78
Cohutta	3	1.30 +/- 0.17	1.04 +/- 0.04	**2.92 +/- 0.74**	6.43	-0.10
Death Valley NP	3	1.75 +/- 0.06	0.69 +/- 0.05	**1.88 +/- 0.30**	7.24	-0.02
Flathead	3	1.82 +/- 0.03	1.09 +/- 0.10	**-1.26 +/- 0.57**	7.27	-1.13
Glacier NP	3	1.90 +/- 0.06	1.48 +/- 0.19	**-4.94 +/- 1.74**	8.00	-0.14

Site	quarter	β_{oc}	β_{sulf}	β_{nit}	NME (%)	NMB (%)
James River Face Wilderness	3	2.09 +/- 0.11	1.04 +/- 0.02	**-0.41 +/- 0.25**	6.22	-0.24
M.K. Goddard	3	1.77 +/- 0.08	1.13 +/- 0.02	**-0.46 +/- 0.29**	6.84	-0.26
Mohawk Mt.	3	1.75 +/- 0.09	1.03 +/- 0.03	**-0.55 +/- 0.21**	8.31	0.14
New York City	3	1.75 +/- 0.14	1.13 +/- 0.04	0.38 +/- 0.17	7.61	-0.36
Point Reyes National Seashore	3	1.51 +/- 0.11	0.86 +/- 0.05	**1.49 +/- 0.16**	9.38	-0.44
Presque Isle	3	1.85 +/- 0.04	1.05 +/- 0.02	**-1.28 +/- 0.52**	6.12	-0.39
Proctor Maple R. F.	3	2.03 +/- 0.06	1.05 +/- 0.02	**-1.05 +/- 0.52**	6.09	0.24
Quaker City	3	1.90 +/- 0.13	1.09 +/- 0.02	**-0.70 +/- 0.40**	6.59	0.01
Saguaro West	3	1.68 +/- 0.35	0.70 +/- 0.14	**2.11 +/- 0.63**	10.17	-1.59
Theodore Roosevelt	3	1.95 +/- 0.05	1.07 +/- 0.05	**-0.46 +/- 0.18**	6.58	-0.90
Trinity	3	1.66 +/- 0.04	0.83 +/- 0.08	**1.75 +/- 0.36**	9.51	1.35
Voyageurs NP #2	3	1.54 +/- 0.06	1.08 +/- 0.06	**1.90 +/- 1.20**	10.12	-0.18
Washington D.C.	3	2.05 +/- 0.14	1.08 +/- 0.03	-0.09 +/- 0.27	7.36	-0.35
Addison Pinnacle	4	**0.26 +/- 0.07**	1.19 +/- 0.03	1.11 +/- 0.06	9.05	0.14
Cape Romain NWR	4	1.81 +/- 0.09	1.03 +/- 0.04	**-0.63 +/- 0.19**	7.98	-0.56
New York City	4	1.42 +/- 0.11	0.87 +/- 0.05	1.17 +/- 0.06	6.02	0.07
Sequoia NP	4	1.43 +/- 0.07	**0.45 +/- 0.13**	1.04 +/- 0.02	8.85	-0.19
Washington D.C.	4	1.48 +/- 0.07	0.95 +/- 0.04	1.06 +/- 0.06	6.95	-0.42

S6. References

Colberg, C. A., Luo, B. P., Wernli, H., Koop, T., and Peter, T.: A novel model to predict the physical state of atmospheric $H_2SO_4/NH_3/H_2O$ aerosol particles, Atmospheric Chemistry and Physics, 3, 909-924, 2003.

Fuller, W. A.: Measurement error models, John Wiley & Sons, New York, 1987.

Hand, J. L., and Malm, W. C.: Review of the IMPROVE equation for estimating ambient light extinction coefficients - final report, Colorado State University, CIRA, 146, 2006.

Khlystov, A., Stanier, C. O., Takahama, S., and Pandis, S. N.: Water content of ambient aerosol during the Pittsburgh air quality study, Journal of Geophysical Research-Atmospheres, 110, 10.1029/2004jd004651, 2005.

McDade, C. E.: IMPROVE standard operating procedure, Crocker Nuclear Laboratory, University of California, Davis, CASOP 351-2, 258, 2008.

Pinder, R W., Dennis, R. L., and Bhave, P. V.: Observable indicators of the sensitivity of $PM_{2.5}$ nitrate to emission reductions—Part I: Derivation of the adjusted gas ratio and applicability at regulatory-relevant time scales, Atmospheric Environment, 42, 1275-1286, doi:10.1016/j.atmosenv.2007.10.039, 2008.

Wexler, A. S., and Clegg, S. L.: Atmospheric aerosol models for systems including the ions H^+, NH_4^+, Na^+, SO_4^{2-}, NO_3^-,Cl-, Br^-, and H_2O, Journal of Geophysical Research-Atmospheres, 107, 14, 4207
10.1029/2001jd000451, 2002.

White, W. H.: IMPROVE data advisory: Shift in EC/OC split with 1 January 2005 TOR hardware upgrade, 2007.