Supplemental Material

1. Why DISP estimation requires that $dQ^{\max} \geq 4$ (One-dimensional case)

The Least Squares (LS) estimate of the expectation value E(x) of a scalar random variable x is simply the average of its n observations, avg(x_i). It is assumed that x has a finite variance; standard deviation of x is denoted by σ.

The LS object function Q, as a function of a test value t, is defined by

$$Q(t) = \sum_{i=1}^{n} \left(t - x_i\right)^2 / \sigma^2 \tag{1}$$

The LS estimate $\hat{x}$ of E(x) is that value of t that minimizes expression (1). It is seen that $\hat{x} = \overline{x} = avg\left(x_i\right)$. The following identity for Q is well known (“Steiner’s rule”):

$$\begin{aligned} Q(t) &= \sum_{i=1}^{n} \left(t - x_i\right)^2 / \sigma^2 \\ &= \sum_{i=1}^{n} \left(\overline{x} - x_i\right)^2 / \sigma^2 + n(t - \overline{x})^2 / \sigma^2 \\ &= Q^{\min} + dQ \end{aligned} \tag{2}$$

It follows that

$$\begin{aligned} Q\left(\mathrm{E}(x)\right) &= Q^{\min} + dQ \\ &= Q^{\min} + n\left(\mathrm{E}(x) - \overline{x}\right)^2 / \sigma^2 \end{aligned} \tag{3}$$

Average of n observations has standard deviation = $\sigma / \sqrt{n}$. With sufficiently large n, the average is also approximately normally distributed. Thus

$$\mathrm{prob}\left(\left|\mathrm{E}(x) - \overline{x}\right| > 2\sigma / \sqrt{n}\right) \approx 0.05 \tag{4}$$

$$\mathrm{prob}\left(\left(\mathrm{E}(x) - \overline{x}\right)^2 > 4\sigma^2 / n\right) \approx 0.05 \tag{5}$$

$$\mathrm{prob}\left(dQ > 4\right) \approx 0.05 \tag{6}$$

It is seen that for the one-dimensional case, constraining dQ to be less than 4 defines a 95% confidence interval for the true value.

2. Why DISP estimation requires that $dQ^{\max} \geq 4$ (The multivariate linear model)

The observations are now fitted by the equation

$$\mathbf{Xc} = \mathbf{b} = \mathbf{b}^0 + \mathbf{e} \tag{7}$$

where column vector **c** is to be estimated. Column vector **b** contains measured values. Vector $\mathbf{b}^0$ represents the unknown true data values; **e** contains measurement errors. It is assumed that cov(**e**)=**1**.

The multivariate case is reduced to one-dimensional cases by using the singular value decomposition SVD(**X**). Singular components are statistically independent of each other; thus they may be estimated separately as a number of one-dimensional cases. In the direction of each singular component of **c**, the 95% confidence interval is obtained from the constraint $dQ<4$. Note that these are *not* joint confidence intervals.

This is not a complete analysis of the multivariate case. This analysis suffices to demonstrate that *dQ* limits smaller than 4 will not provide satisfactory confidence limits because the probability of the obtained interval not containing the true value will be too high. However, this analysis does *not* guarantee that the limit of 4 is sufficiently large for *dQ* in practical work.

3. Factor profiles for simulated datasets

Table S-1 provides the simulated data factor profiles.

Table S-1. Profiles for Factors Used in Simulation, µg/m^3

Species	Coal Combustion	Aged Sea Salt	Copper	Soil
Ca	0.0120	0.0025	0.0027	0.1175
Cl	0.0000	0.0149	0.0007	0.0000
Cu	0.0001	0.0000	0.0025	0.0007
EC	0.0421	0.0065	0.0000	0.0034
Fe	0.0093	0.0015	0.0044	0.0975
K	0.0000	0.0000	0.0056	0.0270
Mn	0.0004	0.0004	0.0000	0.0022
Ni	0.0002	0.0000	0.0000	0.0000
OC	0.6543	0.0197	0.1132	0.0000
Pb	0.0000	0.0001	0.0030	0.0006
$PM_{2.5}$	2.2371	0.0974	0.4186	1.8205
S	0.3134	0.0078	0.0462	0.0508
Se	0.0001	0.0000	0.0002	0.0000
Si	0.0411	0.0060	0.0051	0.3093
Ti	0.0011	0.0001	0.0000	0.0081
Zn	0.0005	0.0001	0.0017	0.0011