Atmospheric
Chemistry
and Physics

Supplement of

Technical note: Effects of uncertainties and number of data points on line fitting – a case study on new particle formation

Santtu Mikkonen et al.

Correspondence to: Santtu Mikkonen (santtu.mikkonen@uef.fi)

Supplement S1 – Alternative simulating methods for test data

To investigate the effect of alternative data distributions on the linear regression, the identical data simulation and analysis was repeated for a noise-free H_2SO_4 equally distributed in log-space. The range for the noise-free $\log10(H_2SO_4)$ was chosen 6.5, 8.5 to generate simulated observations in a more narrow shape of scatter compared to the lognormally generated dataset.

The resulting figures, created in similar manner than figs 1 and 2 in the main text showing the overall fits and effect of increasing uncertainty indicate that when the range of data is wide enough, the OLS method works equally well with the EIV methods. However, Fig S1 shows that when the H_2SO_4 concentrations are in level observed in atmosphere, the high uncertainties cause OLS to underestimate the slope again. In addition, Fig S2 shows that when the level of uncertainty is increased, OLS is underestimating the slope more than the EIV methods.
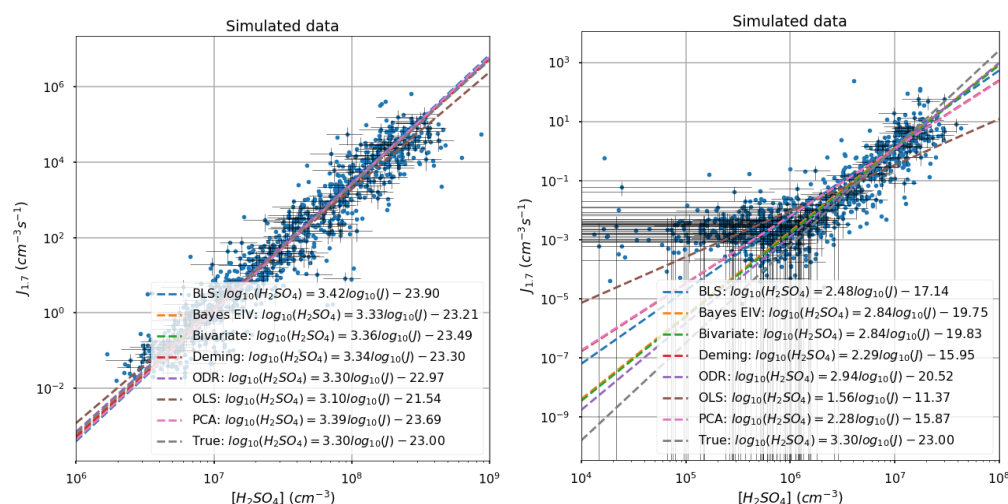

Figure S1. Regression lines fitted to the simulated data with all methods in comparison. In left panel the $\log10(H_2SO_4)$ varies in range 6.5-8.5 whereas in right panel the concentration is in level typically observed in atmosphere. Whiskers in data points refer to the measurement error used for simulation.
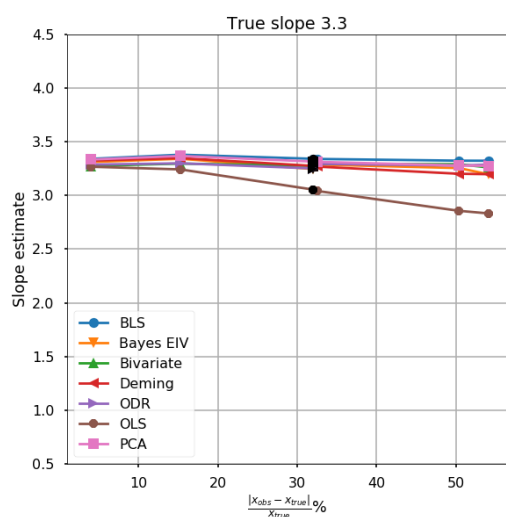

Figure S2. Sensitivity test for increasing uncertainty in simulated data when $\log10(H_2SO_4)$ varies in range 6.5-8.5. Black markers show the initial data set shown in Fig. S1.