Supplemental Information

Figure 1: (**S1**) Modeled concentration (triangles) and sensitivities (barplot) of daily maximum 8-hr average (MD8) ozone to emissions of Mobile NO_x (MNO_x), Point NO_x (PNO_x), all Other NO_x (ONO_x), Anthropogenic VOCs (AVOC), Biogenic VOCs (BVOC) and Ozone Boundary Conditions (O3 BC) at AQS site 13-089-0002 in Decatur, GA during July 1 - September 30, 2005. Negative sensitivies indicate times when an increase in emissions correspondes to a decrease in daily maximum 8-hr ozone. The difference between the modeled ozone and the total sensitivity is due to second order and cross sensitivities between these 6 parameters which are included in the analysis but not plotted here, as well as other model parameters that may influence ozone concentrations.

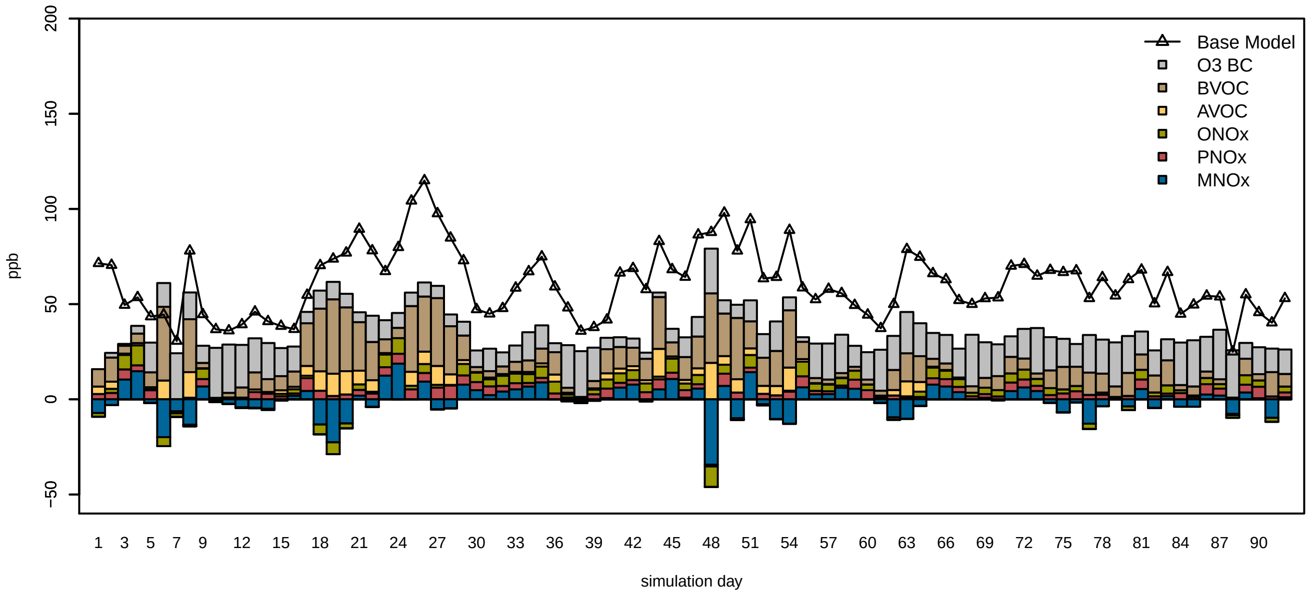
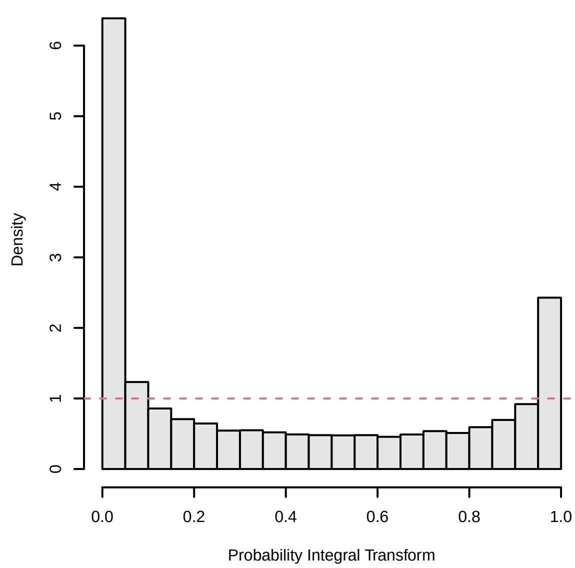


Figure 2: (**S2**) Probability Integral Transform (PIT) histogram for the predicted MD8 ozone concentrations for each site/day based on Model 3 (left) and based on Model 6 using the five-fold cross validation model fit (right). The PIT histogram is a common tool used for assessing the calibration of a probability density model prediction. A model is said to be well calibrated if it predicts a certain event has probability P and the long-term proportion of that event actually occurring is P. For example if we calculate 90% prediction intervals for ozone concentrations, we expect the observed concentrations to fall within these predicted ranges 90% of the time over many different trials (i.e. sites and/or days). For each site/day we compute the cumulative distribution function (cdf) for the observed ozone concentration based on the predicted probability density at that location and time and the histogram of these cdf values are plotted. The PIT histogram for a perfectly calibrated probabilistic model should have the form of a Uniform(0,1) distribution. The U-shaped histogram in figure (a) below indicates that the non-Bayesian RFM (Model 3) tends to be under-dispersive (i.e. the observations frequently fall outside the range of predicted values) and the model most frequently over-predicts the ozone concentrations at these sites and days. In contrast, the histogram for Model 6 shown in figure (b) is much closer to a Uniform(0,1) pattern indicating good calibration of the model for the entire cross validation dataset. These results are consistent with the cross validation statistics shown in Table 1 in the main text.

(a) PIT Histogram for Model 3



(b) PIT Histogram for Model 6

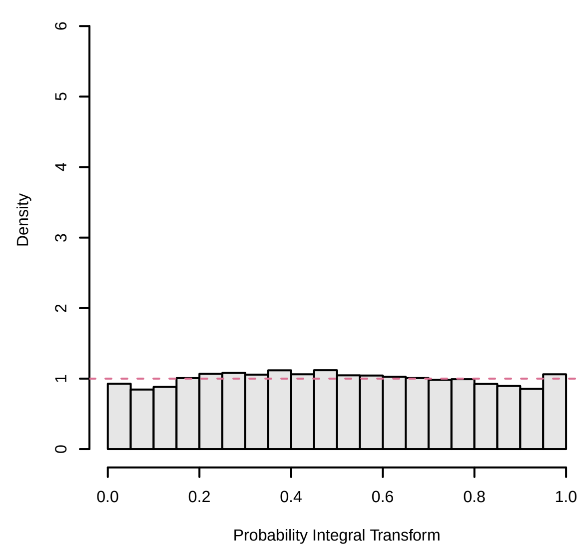


Figure 3: (**S3**) 5th (top row), 50th (middle row), and 95th (bottom row) percentiles of the daily first order sensitivity coefficients across the domain (ppb). The daily sensitivity coefficients are created by averaging over the same eight-hour time period used to calculate the maximum eight-hour average ozone concentrations from the base model simulation. Negative sensitivities indicate days when an increase in emissions corresponds to a decrease in daily maximum 8-hr ozone.

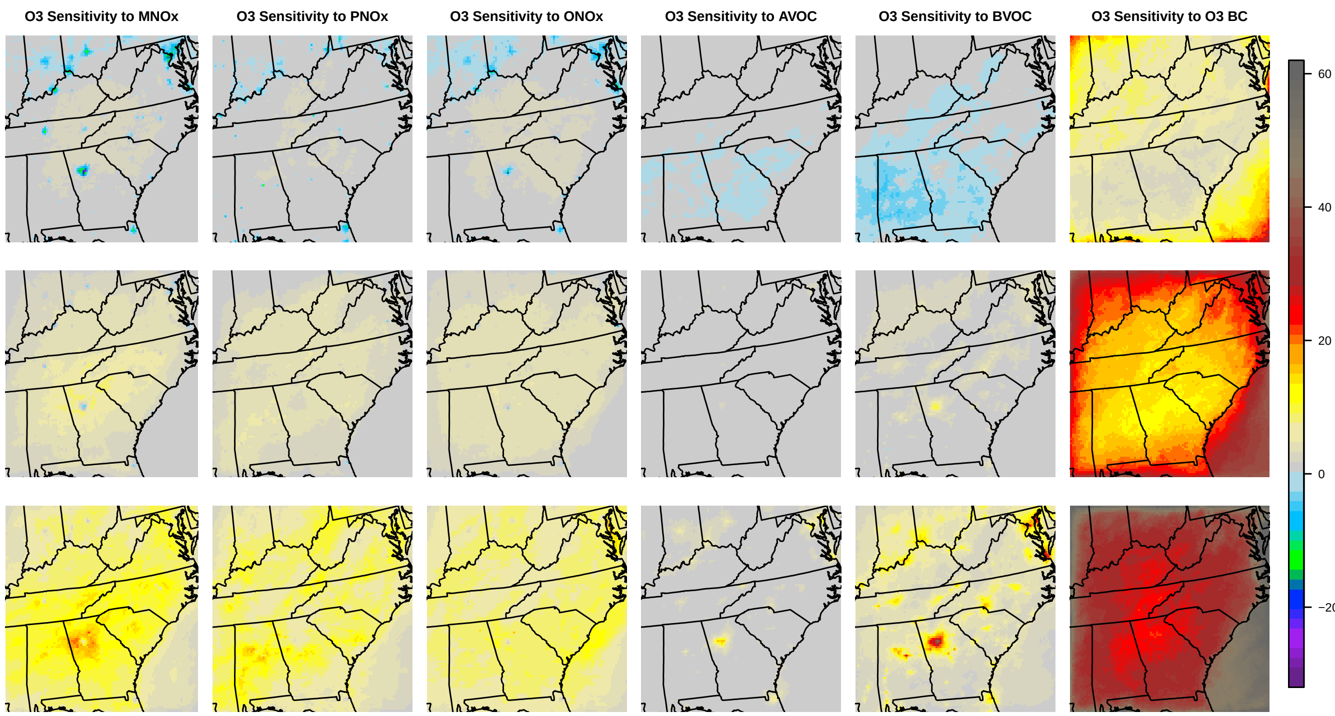


Figure 4: (**S4**) Impact of 4 emission reduction scenarios on fourth highest summer MD8 ozone concentration. Top row shows the change in the fourth highest MD8 using the base simulation and the RFM output. Middle and bottom rows show the posterior mean and standard deviation, respectively, of the change in fourth highest MD8 based on the Bayesian RFM. All units are ppb.