

1 **Use of High-Order Sensitivity Analysis and Reduced-Form**
2 **Modeling to Quantify Uncertainty in Particulate Matter Simulations**
3 **in the Presence of Uncertain Emissions Rates: A Case Study in**
4 **Houston**

5 **Wenxian Zhang***, Marcus A. Trial, Yongtao Hu, Athanasios Nenes, Armistead G. Russell

6 *Corresponding author. Telephone: 1-404-825-0960, Email address: wzhang@trinityconsultants.com

7

8 **Abstract**

9 Regional air quality models are widely used to evaluate control strategy effectiveness. As
10 such, it is important to understand the accuracy of model simulations to establish confidence in
11 model performance and to guide further model development. Particulate matter with
12 aerodynamic diameter less than 2.5 micrometers (PM_{2.5}) is regulated as one of the criteria
13 pollutants by the National Ambient Air Quality Standards (NAAQS), and PM_{2.5} concentrations
14 have a complex dependence on the emissions of a number of precursors, including SO₂, NO_x,
15 NH₃, VOCs, and primary particulate matter (PM). This study quantifies how the emission-
16 associated uncertainties affect modeled PM_{2.5} concentrations and sensitivities using a reduced-
17 form approach. This approach is computationally efficient compared to the traditional Monte
18 Carlo simulation. The reduced-form model represents the concentration-emission response and is
19 constructed using first- and second-order sensitivities obtained from a single CMAQ/HDDM-PM
20 simulation. A case study is conducted in the Houston-Galveston-Brazoria (HGB) area. The
21 uncertainty of modeled, daily average PM_{2.5} concentrations due to uncertain emissions is
22 estimated to fall between 42% to 52% for different simulated concentration levels, and the

23 uncertainty is evenly distributed in the modeling domain. Emission-associated uncertainty can
24 account for much of the difference between simulation and ground measurements as 60% of
25 observed $PM_{2.5}$ concentrations fall within the range of one standard deviation of corresponding
26 simulated $PM_{2.5}$ concentrations. Uncertainties in meteorological fields as well as the model
27 representation of secondary organic aerosol formation are the other two key contributors to the
28 uncertainty of modeled $PM_{2.5}$. This study also investigates the uncertainties of the simulated
29 first-order sensitivities, and found that the larger the first-order sensitivity, the lower its
30 uncertainty associated with emissions. Sensitivity of $PM_{2.5}$ to primary PM has the lowest
31 uncertainty while sensitivity of $PM_{2.5}$ to VOC has the highest uncertainty associated with
32 emission inputs.

33

34 **1. Introduction**

35 Significant effort has been expended to improve air quality due to its influence on human
36 health and the environment. The United States Environmental Protection Agency (U.S. EPA)
37 sets National Ambient Air Quality Standards (NAAQS) to protect public health and the
38 environment. Particulate matter (PM), including fine particles with 2.5 micrometers in diameter
39 and smaller ($PM_{2.5}$), is currently regulated as one of the NAAQS criteria pollutants. The harmful
40 effects of PM on human health have been a focus as exposure to $PM_{2.5}$ is associated with
41 respiratory and cardiovascular disease (Zanobetti et al., 2000; Schwartz, 1994, Dockery et al.,
42 1993). A recent study (Kaiser, 2005) found that fine particles are potentially of more concern
43 than larger particles in causing respiratory disease and premature death due to their ability to
44 penetrate deep into the lung. In order to more effectively protect the public from adverse health
45 effects due to exposure to fine particles, in December 2012, U.S. EPA tightened the primary

46 NAAQS for the annual average concentration of fine particles from $15 \mu\text{g m}^{-3}$ to $12 \mu\text{g m}^{-3}$.
47 Their contribution to acidic deposition and visibility are also concerns (e.g., Galloway et al.,
48 2004; Watson 2002).

49 PM control is perhaps the most complex aspect of current air quality management. The
50 complexity comes from the many components of PM and formation routes. Regional air quality
51 models are frequently used in air quality management to evaluate the effectiveness of emissions
52 control (U.S. EPA, 2004). The accuracy of these models is limited by their representation of the
53 complex chemical and physical processes of pollutant transport and transformation, as well as
54 the lack of accuracy in inputs (e.g., emissions rates, meteorological conditions, and initial and
55 boundary conditions). Previous studies have investigated model uncertainties on ozone, focusing
56 on uncertainties due to emission estimates, initial and boundary conditions, grid size, and
57 chemical reactions (e.g., Hanna. et al., 2001, 1998; Cohan et al., 2010; Pinder et al., 2009).
58 Uncertainties in emission inventories remain a leading cause for discrepancies between models
59 and observations (Xiao et al., 2010). As such, quantification of the influence of uncertain
60 emission inventories on simulated concentrations of $\text{PM}_{2.5}$ is informative to the air quality
61 management processes as well as to guide model improvement, and the importance of that
62 information is becoming more apparent as an increased focus is placed on $\text{PM}_{2.5}$. Of further
63 importance is the uncertainty in capturing the response of air quality models to emission
64 changes, i.e., the uncertainty of the sensitivity of air quality models to emission changes.

65 Estimates of the uncertainty of air quality model prediction from uncertainties in input
66 parameters has relied heavily on Monte Carlo simulations that randomly sample model inputs
67 (according to their probability distributions) and then quantified the uncertainties of model
68 outputs (e.g. pollutant concentrations and sensitivities) by using the ensemble outputs obtained

69 from the Monte Carlo simulations (e.g., Tian et al., 2010; Pinder et al., 2009; Hanna et al., 1998,
70 2001). Initially, studies conducted the Monte Carlo simulations by running the underlying air
71 quality model multiple times (e.g., Deguillaume et al., 2008; Hanna et al., 1998 and 2001; Bergin
72 et al., 1999). However, this approach becomes computationally expensive and cumbersome for
73 three-dimensional time-dependent models applied over large domains. More recently, studies
74 have employed a reduced form model (RFM) approach, which substantially reduces the
75 computational cost (e.g. Kerl et al., 2014; Napanelok et al., 2011; Tian et al., 2010; Digar and
76 Cohan, 2010; Pinder et al., 2009). This approach constructs a reduced form model of the
77 underlying air quality model by capturing concentration-parameter responses of the original
78 model. High order direct sensitivity analysis is efficient at extracting the concentration-parameter
79 response by simultaneously providing first- and second-order sensitivity coefficients along with
80 the base concentration simulation. This advanced sensitivity technique has been implemented in
81 air quality models (e.g., the Community Multiscale Air Quality (CMAQ) model (Byun and
82 Schere, 2006) for gas and aerosol species (Zhang et al., 2012; Napanelok et al., 2011; Hakami et
83 al., 2003; Yang et al., 2007), and the Comprehensive Air Quality Model with Extensions
84 (CAMx) (ENVIRON, 2005; Cohan et al., 2010) for gas and aerosol species). It has already been
85 applied to characterize the uncertainty of modeled ozone production (Napanelok et al., 2011;
86 Tian et al., 2010; Digar and Cohan, 2010; Pinder et al., 2009) and to investigate the influence of
87 reaction constants' uncertainties on ozone sensitivities (Xiao et al., 2010; Cohan et al., 2010).

88 This paper discusses application of the RFM based on CMAQ to efficiently quantify the
89 emission-associated uncertainties of the simulated PM_{2.5} concentrations and sensitivities for an
90 air pollution episode in the Houston region. An underestimation of VOC emissions in the
91 Houston Ship Channel (HSC) has been found by a number of studies (e.g. Kim et al., 2011,

92 Cowling et al., 2007, and Ryerson et al., 2003), and Kim et al. (2011) found an overestimation of
93 NO_x emissions over the HSC in the 2005 National Emissions Inventory (NEI). The complex
94 concentration-emissions responses due to emissions from the petrochemical plants and the
95 unique geographic and meteorological conditions makes this area a good case for studying the
96 emission uncertainty. Using high-order sensitivity analysis, this paper evaluated the impact of
97 uncertain emission inventories on PM_{2.5} concentrations and sensitivities.

98 **2. Methods**

99 2.1 Modeling system

100 Air quality modeling is conducted using CMAQ (Byun and Schere, 2006) version 4.7.1
101 (CMAQ v4.7.1) with the SAPRC 99 (Carter, 2000) chemical mechanism and the AERO5 aerosol
102 module (Foley et al., 2010; Carlton et al., 2010). CMAQ v4.7.1 has been equipped with the
103 Decoupled Direct Method in Three Dimensions (DDM-3D) (Napelenok et al., 2006), which has
104 been extended to high-order DDM-3D for particulate matter (HDDM-3D/PM) by Zhang et al.
105 (2012), and it was the latest version of CMAQ that had HDDM-3D/PM implemented in when
106 this work was carried out.

107 The CMAQ model application here uses three one-way nested modeling domains (Figure
108 S1). The outer-most domain covers the entire continental United States and portions of Canada
109 and Mexico with 36- by 36-km horizontal grids; the middle domain covers eastern Texas and the
110 surrounding states of Oklahoma, Arkansas, and Louisiana with 12- by 12- km grids; the inner-
111 most domain covers southeastern Texas which contains the Houston-Galveston-Brazoria (HGB)
112 region where intense emissions from petrochemical industries occur. The three domains have 13
113 vertical layers extending approximately 16 km above ground, with seven layers below 1 km.

114 The Weather Research and Forecasting (WRF) model is widely used in atmospheric
 115 research and weather forecasting. This application used the WRF model to prepare the
 116 meteorological fields and is run with 34 layers using four-dimensional data assimilation (FDDA)
 117 techniques and the Noah land-surface model with MODIS landuse data. The Sparse Matrix
 118 Operator Kernel for Emissions (SMOKE) is used to process emissions to provide gridded,
 119 CMAQ-ready emissions. The inventory used is the U.S. National Emissions Inventory (NEI) of
 120 2005 (<ftp://ftp.epa.gov/EmisInventory/2005v4/>) (Figure 1).

121 2.2 Reduced-form model of CMAQ

122 Uncertainty analysis performed here is based on a RFM of CMAQ. The RFM represents
 123 the relationship between pollutant concentrations and the model inputs in a straightforward way
 124 and can be used to efficiently propagate uncertainties from model inputs to outputs. Constructing
 125 the RFM involves Taylor series expansion of the pollutant concentration at a given time and
 126 location for fractional perturbations in sensitivity parameters of interest (Eq. 1) (e.g., Cohan et al.
 127 2005). The sensitivity parameters can be emissions rates, chemical reaction rates, or initial and
 128 boundary conditions. Only emissions rates are considered here for studying the emission-
 129 associated model uncertainties:

$$130 \quad C_i^* = C_{i,0} + \sum_{j=1}^J \Delta \varepsilon_j S_{i,j}^{(1)} + \frac{1}{2} \sum_{j=1}^J \Delta \varepsilon_j^2 S_{i,j,j}^{(2)} + \sum_{j \neq k} \Delta \varepsilon_j \Delta \varepsilon_k S_{i,j,k}^{(2)} + H.O.T. \quad (1)$$

131 where C_i^* and $C_{i,0}$ denote the concentration of pollutant i with and without perturbations in
 132 sensitivity parameters, respectively. $S_{i,j}^{(1)}$, $S_{i,j,j}^{(2)}$, and $S_{i,j,k}^{(2)}$ are semi-normalized sensitivity
 133 coefficients. i denotes the i^{th} species, j and k denotes the j^{th} and k^{th} emissions rates. Note that C_i^* ,
 134 $C_{i,0}$, $S_{i,j}^{(1)}$, $S_{i,j,j}^{(2)}$, and $S_{i,j,k}^{(2)}$ all vary with time and location, and the notations for time and

135 location are omitted for brevity. $\Delta\varepsilon_j$ is the relative change (e.g., 0 is no change and 1 is a 100%
136 increase) in the j^{th} emission rate. There is considerable flexibility in how the j^{th} emission rate is
137 specified: j can be the emissions of a specific species from all sources, the emissions of all
138 species from a specific source, the emissions of multiple pollutants from a specific location, or
139 various combinations. In this study, we consider the five major groups of emitted species, sulfur
140 dioxide (SO_2), nitrogen oxides (NO_x), volatile organic compounds (VOCs), ammonia (NH_3), and
141 primary PM, that impact $\text{PM}_{2.5}$ concentration. *H.O.T.* stands for higher order terms.

142 CMAQ-HDDM-3D is used to calculate local first- and second-order semi-normalized
143 sensitivity coefficients. This sensitivity technique is efficient as it simultaneously computes all
144 sensitivity coefficients along with concentrations in one model run. The controlling equations for
145 sensitivity coefficients are derived by differentiating governing equations for the concentrations
146 with respect to the sensitivity parameters (Dunker 1981). Thus, equations involving sensitivities
147 and concentrations have a similar structure and are calculated using the same numerical
148 algorithms. First- and second-order sensitivity coefficients calculated by CMAQ-HDDM-3D
149 have been evaluated by comparing them with traditional finite differences and good agreement
150 has been observed for both gas-phase species and particulate matter (Zhang et al., 2012; Hakami
151 et al., 2003). Extensive evaluation of the CMAQ-HDDM/PM sensitivities and the RFM used in
152 this study has been conducted, and demonstrated the reliability of the RFM performance in
153 replicating the results of the original CMAQ (Figures S3-S5).

154 The RFM for first order sensitivities can be constructed based on second-order
155 sensitivities (e.g., Tian et al., 2010; Cohan et al. 2005). The formulation of the RFM is similar to
156 that of the concentrations:

157
$$S_{i,j}^{(1)*} = S_{i,j,0}^{(1)} + \Delta\varepsilon_j S_{i,j,j}^{(2)} + \sum_k \Delta\varepsilon_k S_{i,j,k}^{(2)} + H.O.T., \quad (2)$$

158 where $S_{i,j}^{(1)*}$ and $S_{i,j,0}^{(1)}$ respectively denote first order sensitivity of pollutant i to emission rate j
159 with and without considering the uncertainty in the emissions rates. $S_{i,j,j}^{(2)}$ is second-order self
160 sensitivity of pollutant i to emissions rates j . $S_{i,j,k}^{(2)}$ is second-order cross sensitivity of pollutant i
161 to emissions j and k . $\Delta\epsilon_j$ is the relative emission change as described for Eq. 1. An evaluation of
162 Eq. 2 using the sensitivity output from the original CMAQ-HDDM shows an excellent
163 agreement between the two (Figure S6).

164

165 2.3 Quantification of emission-associated uncertainties

166 Monte Carlo simulations using the reduced-form CMAQ are applied to quantify the
167 emission-associated uncertainties of modeled PM_{2.5} concentrations and sensitivities. Three steps
168 are involved in the Monte Carlo simulations. The first step is to sample the emissions rates of
169 interest based on their relative uncertainties and probability distributions. The second step is to
170 propagate uncertainties through the reduced-form model and collect an ensemble of model
171 outputs. The third step is to quantify model uncertainties from the output ensemble.

172 Emissions rate uncertainties are assumed to be log-normally distributed, as is
173 approximately found for many environmental geographical variables that are constrained to be
174 positive (Hanna et al., 1998). This study focuses on the domain-wide total
175 (anthropogenic+biogenic) emissions of five major pollutants that impact atmospheric PM levels:
176 SO₂, NO_x, VOC, NH₃, and primary PM. As uncertainties explicit to NEI 2005 are unavailable,
177 the uncertainties of these emissions are obtained from previous studies, which may not be
178 exactly the same as the uncertainties in 2005 NEI but provide good references for estimating the
179 uncertainties in 2005 NEI. Hanna et al. (2001) summarized the estimates of uncertainty factors
180 for NO_x and VOC emissions from major point, mobile, biogenic, and area sources. The

181 NARSTO PM assessment (2004) provides the uncertainties in the national emission inventory
182 for SO₂, NO_x, VOC, NH₃, and primary PM from various source categories. Combining the
183 confidence levels in the NARSTO assessment with the uncertainty factors in Hanna et al. (2001),
184 the uncertainty factors of the emissions rates of the five pollutants of interest can be estimated
185 (Table 1).

186 A random number generator is applied to produce multiple sets of input emissions rates
187 given their probability distributions and uncertainties. This study selected a sampling size of
188 1000, which has been demonstrated to achieve sufficient convergence in the uncertainty analysis
189 on ozone simulation conducted by Pinder et al. (2009). Each sample includes a separate value for
190 each emitted species.

191 For every grid cell and time step, the multiple sets of sampled emissions rates were input
192 to the RFM and generate 1000 values of PM_{2.5} concentrations, from which the relative
193 uncertainty is calibrated. Among various methods to calibrate the relative uncertainty, we
194 selected the method used by Tian et al. (2010). They calculated the relative uncertainty of
195 simulated PM_{2.5} at each grid cell by dividing a half of the 68.3% confidence interval (CI)
196 (corresponding to 2 standard deviations) of the 1000 values by their median (Eq. 3) to exclude
197 potential outliers due to accumulating predictive uncertainties of the RFM at extreme emission
198 changes. The standard deviation (SD) of the simulated PM_{2.5} at each grid is equal to the product
199 of the relative uncertainty and the median (Eq. 4).

$$200 \quad \text{Relative Uncertainty} = \frac{68.3\% \text{ CI}}{2 \times \text{median}} = \frac{84.15^{\text{th}} \text{ percentile} - 15.85^{\text{th}} \text{ percentile}}{2 \times \text{median}} \quad (3)$$

$$201 \quad \text{Standard Deviation} = \frac{68.3\% \text{ CI}}{2} = \frac{84.15^{\text{th}} \text{ percentile} - 15.85^{\text{th}} \text{ percentile}}{2} \quad (4)$$

202 The method used to calculate sensitivities (sensitivities to domain wide emission changes
203 applied over the duration of the simulation) and the method used to calculate uncertainties
204 impact interpretation of the results. Calculation of uncertainties assumes that emission errors are
205 the same across space and time for each pollutant, but the errors in emissions between species is
206 independent. Thus, the actual uncertainties are likely less (much of the uncertainties in the
207 emissions are likely due to local (time and space) variations in the emissions), but can be
208 underestimated in that some of the emission uncertainties can be correlated between species (e.g.,
209 due to uncertainties in activities in a specific source that emits more than one species). We have
210 chosen to use the approach most common in prior studies.

211 **3. Results and Discussion**

212 3.1 Model evaluation

213 Surface meteorological fields simulated by WRF are evaluated by using hourly surface
214 observations in the U. S. and Canada. The bias and root mean square errors (RMSE) for the three
215 domains are well within the range considered to be acceptable for air quality modeling (Table 2)
216 (Emery et al., 2001; Hanna and Yang, 2001). Good agreement between the WRF simulation and
217 observations minimizes uncertainty due to input meteorological fields.

218 CMAQ performance is evaluated using data retrieved from the Air Quality System (AQS;
219 available at <http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqdata.htm>). Simulated 8-hr
220 averaged ozone and daily averaged PM_{2.5} concentrations are compared with monitoring data
221 inside the 4 km domain. The mean normalized bias (MNB) and mean normalized error (MNE)
222 for 8-hr averaged ozone are -7 ppb and 19%, respectively, well within suggested performance
223 levels (EPA 2007) (Table S1). The definition of MNB and MNE is shown by Eqs. 5 and 6,
224 where C_m and C_o are the modeled and observed PM_{2.5} concentrations, respectively. PM_{2.5}

225 simulations are commonly evaluated by using mean fractional bias (MFB) and mean fractional
 226 error (MFE) (see Eqs 7, 8, and Fig. 2). For this simulation, MFB and MFE are -30% and 54%,
 227 respectively, which are within the acceptable range according to the guidance of EPA (2007).
 228 The MFBs of daily averages of sulfate, nitrate, ammonium, organic carbon (OC), and EC are -
 229 65%, -122%, -59%, -25%, and 47%, respectively. Graphically, these MFBs and MFEs are shown
 230 and compared against suggested criteria using bugle plots (Boylan and Russell, 2006) (Figure 2).
 231 CMAQ performance of PM_{2.5} is further detailed in Table S1. Nitrate aerosols have higher MFB
 232 and MFE than other PM_{2.5} species. This is largely due to the limitation of the model in
 233 representing the complex processes involved in nitrate aerosol formation (e.g., Kwok et al. 2013,
 234 Civerolo et al., 2010, Zhang et al., 2009, Yu et al., 2005).

$$235 \quad MNB = \frac{1}{N} \sum_{i=1}^N \frac{C_m - C_o}{C_o} \quad (5)$$

$$236 \quad MNE = \frac{1}{N} \sum_{i=1}^N \left| \frac{C_m - C_o}{C_o} \right| \quad (6)$$

$$237 \quad MFB = \frac{2}{N} \sum_{i=1}^N \frac{C_m - C_o}{C_m + C_o} \quad (7)$$

$$238 \quad MFE = \frac{2}{N} \sum_{i=1}^N \left| \frac{C_m - C_o}{C_m + C_o} \right| \quad (8)$$

239

240 3.2 Uncertainty of modeled concentrations

241 Using the sampling results to drive the reduced-form CMAQ gives an ensemble of
 242 pollutant concentrations for each grid cell at every hour. In this study, daily averages of PM_{2.5}
 243 concentrations are studied because that is one metric used in the NAAQS, and it is commonly
 244 used for model performance evaluation. Since the response of PM_{2.5} concentration to precursor
 245 emissions varies spatially and temporally, the median and uncertainty are calculated for 171,864
 246 ensembles across the entire modeling domain over 30 days in the episode. We group the results

247 based on the base PM_{2.5} levels to assess if uncertainties vary with simulated level. Emission-
248 related relative uncertainties for different levels of PM_{2.5} fall into a range from 42% to 52%
249 (Figure 3). This is consistent with the spatial distribution of the relative errors of modeled PM_{2.5}
250 (Figure 4), which shows that most areas have a relative uncertainty in model simulated daily
251 average PM_{2.5} around 40%. Hot spots in the PM_{2.5} concentration field do not have higher relative
252 uncertainties. Instead, the relative uncertainty distributes more evenly over the entire domain,
253 indicating that CMAQ has similar relative uncertainty over a wide range of PM_{2.5} concentrations.

254 While the daily relative uncertainties calculated here may be high, the uncertainty in the
255 mean PM is likely much lower, though quantifying such would require understanding the
256 temporal and spatial correlations in the emissions uncertainties.

257

258 3.3 Comparison with observations

259 The emission associated uncertainty of simulated PM_{2.5} can be used to investigate how
260 much the difference between the model simulated PM_{2.5} and ground measurements of PM_{2.5} can
261 be explained by emissions uncertainties (Figure 3). The red bars indicate the normalized mean
262 error (NME) between the simulations and available observations of daily PM_{2.5} concentrations
263 fall into various PM_{2.5} levels. The errors between simulated and observed PM_{2.5} are slightly
264 higher than the medians of emissions-associated uncertainties for observed PM_{2.5} concentrations
265 that are between 5 and 15 $\mu\text{g m}^{-3}$. For observed PM_{2.5} levels that are between 15 and 30 $\mu\text{g m}^{-3}$,
266 the differences between simulated and observed PM_{2.5} are below the medians of emissions-
267 associated uncertainties. This suggests that the emission-associated uncertainty can explain a
268 large fraction of the simulation errors. Five observation sites in the Houston Ship Channel region
269 were selected for more detailed study (Figure S2). The five sites have available continuous PM_{2.5}

270 monitoring data and a diversity of land use types, which include urban, suburban residential,
271 agricultural, and industrial. They are located to the south, north, and west of the Houston Ship
272 Channel, so they are able to represent various impacts from the areas with intense industrial
273 emissions and as those emissions evolve and are impacted by emissions from other sources,
274 including biogenic.

275 Modeled and observed daily averages of $PM_{2.5}$ concentrations at the five sites are
276 compared for the entire episode (Figure 5). The error bars represent the emission-associated
277 uncertainties, expressed as the SD (Eq. 3). The dashed lines correspond to the 95% confidence
278 interval (CI), which are obtained by calculating the 2.5th and 97.5th percentiles of the ensemble
279 results. Across the five sites, the percentage of observations that fall in the range of the SD is
280 below 60%, and the percentage of observations that fall in the 95% CI is 85% (Table 3), both
281 slightly less than expected if all of the error were due to uncertainties in the emissions, alone.
282 Both Houston East (AQS#482011034) and Channelview (AQS#482010026) have over 90
283 percent of the observed concentrations falling in the 95% CI, implying that uncertainties in the
284 emission rates can explain much, but not all, of the difference between the simulated and
285 observed $PM_{2.5}$ concentrations at the two sites (Figures 5a and 5b). The two sites which are south
286 of the Houston Ship Channel (Figures 5c and 5d) have about 70% of the observations in the 95%
287 CI. Modeled $PM_{2.5}$ concentrations are consistently biased high during the last ten days of the
288 episode. These two sites are close to Galveston Bay so the sea breeze can have a large impact on
289 the air pollutant concentrations. The Kingwood site (Figure 5e) exhibits a low bias for $PM_{2.5}$ in
290 the simulation. The reason may be due to its location, which, unlike the other four sites near the
291 Houston Ship Channel, is about 25 miles northeast of Houston's downtown and is located in an
292 area with substantial biogenic VOC emissions. Studies have shown that the current air quality

293 model tends to underestimate secondary organic aerosols at monitoring sites. Here, the
294 comparison between model simulated and observed concentrations also indicates a low bias. 19
295 out of 31 monitors had observed daily average PM_{2.5} higher than the upper bound of the standard
296 deviation of the model simulation. This is due in part to the model representation of the chemical
297 reactions, thermodynamic, and the formation of secondary organic aerosols.

298 During the same period, simulated meteorological fields had higher errors and bias than on
299 the other days: Temperature is biased high in the last ten days of the episode, and wind direction
300 shows larger deviation from observations compared to the first half of the episode. The
301 correlations between the root mean square error of the meteorological fields and that of PM_{2.5}
302 concentrations indicate that the error in PM_{2.5} simulation is related more to errors in wind
303 direction ($R^2 = 0.2$) than to errors in temperature, wind speed, and relative humidity. In addition,
304 the low bias in the PM_{2.5} simulation for August 29-31 at all the five sites is due, in part, to biases
305 associated with meteorological fields. The system simulated a precipitation event occurred from
306 August 29 to 31, which came from north and swept the Houston Ship Channel and the western
307 shore of the Galveston Bay and the heavy rainfall decreased the simulated PM_{2.5} concentration.
308 However, the same reduction is not found in the observed PM_{2.5} concentrations at the five
309 monitoring sites, so the biased-low PM_{2.5} is, in part, attributable to the error in the simulated
310 precipitation and estimated scavenging. The daily precipitation observational data from CPC
311 Unified Gauge-Based Analysis of Daily Precipitation ([http://www.esrl.noaa.gov/psd/cgi-](http://www.esrl.noaa.gov/psd/cgi-bin/db_search/SearchMenus.pl)
312 [bin/db_search/SearchMenus.pl](http://www.esrl.noaa.gov/psd/cgi-bin/db_search/SearchMenus.pl)) shows that the observed precipitation is located right above the
313 Galveston Bay and moved to the East. The monitors around the Houston Ship Channel were not
314 highly affected by the rain. In contrast, the simulated precipitation is centered at the HGB area,
315 so washout of PM_{2.5} led to low simulated concentrations of PM_{2.5}.

316 To further investigate the contribution of uncertain emissions to simulated PM_{2.5}, we
317 plotted the histogram of the difference between simulated and observed PM_{2.5} concentrations
318 (i.e., simulated minus observed daily average PM_{2.5} concentrations) at any observation site with
319 continuous PM_{2.5} monitoring data (Figure 6). The probability distribution is close to a log-normal
320 distribution. The mean of the histogram is -1.5 $\mu\text{g m}^{-3}$, which means the simulated PM_{2.5}
321 concentration is generally biased low and is explained by the secondary organic aerosol bias. The
322 standard deviation of the histogram is 8.6 $\mu\text{g m}^{-3}$, while the average standard deviation of PM_{2.5}
323 concentrations due to emissions uncertainties is 7.3 $\mu\text{g m}^{-3}$. The comparable standard deviations
324 suggest that emission-associated uncertainty can explain a large portion, but not all, of the
325 uncertainty in PM_{2.5} simulations in a regional air quality model.

326

327 3.4 Uncertainties in PM sensitivities

328 Accurately calculating pollutant concentration sensitivities is important because they are
329 directly linked to estimating control strategy effectiveness. The uncertainty of PM_{2.5} sensitivities
330 can be estimated using the same approach of quantifying the emission-associated uncertainty of
331 modeled PM_{2.5}. The sampled emissions rates are propagated through the RFM described by Eq. 2
332 for quantification of the emission-associated uncertainty of first order sensitivity of PM_{2.5}. The
333 method described by Eq. 3 is applied to calculate the relative uncertainty of the modeled first
334 order sensitivities of PM_{2.5} to the emissions rates of the five major pollutants (i.e., NO_x, primary
335 PM, SO₂, NH₃, and VOC). For each of these sensitivities, their uncertainties are grouped based
336 on the magnitudes of the associated sensitivities, and each group is represented by a bin in the
337 box plot (Figure 7).

338 In general, the relative uncertainty of each first-order sensitivity decreases when the
339 magnitude of the first-order sensitivity becomes higher. The averages of the median uncertainties
340 for sensitivities of PM_{2.5} to primary PM, NO_x, SO₂, NH₃, and VOC are 16%, 32%, 55%, 75%,
341 and 128%, respectively. These uncertainties are on a relative basis. Taking $\frac{\partial PM_{2.5}}{\partial NO_x}$ for example, its
342 relative uncertainty is 32%, and its simulated value is 1.3 $\mu\text{g m}^{-3}$, so its standard deviation can be
343 estimated by $32\% \times 1.3 \mu\text{g m}^{-3} \approx 0.4 \mu\text{g m}^{-3}$, which means its uncertainty on an absolute basis
344 can be expressed as $1.3 \pm 0.4 \mu\text{g m}^{-3}$. Comparing the five first-order sensitivities, sensitivity of
345 PM_{2.5} to primary PM is the largest one. The spatial average of first-order sensitivity of PM_{2.5} to
346 primary PM on the day of September 8 in the HGB area is 5.2 $\mu\text{g m}^{-3}$, followed by SO₂ at 1.8 μg
347 m^{-3} , NO_x at 1.3 $\mu\text{g m}^{-3}$, NH₃ at 0.9 $\mu\text{g m}^{-3}$, and VOC at 0.8 $\mu\text{g m}^{-3}$. The standard deviation in the
348 sensitivity caused by uncertainty in emissions rates is comparable (around 1.1 $\mu\text{g m}^{-3}$) for all
349 species. The same set of perturbations in emissions is applied to each first-order sensitivity, so
350 the magnitude of second-order determines the difference in first-order sensitivities and thus the
351 relative uncertainty. Second-order sensitivities associated with VOC are larger than the others
352 (Table 4). Combined with first-order sensitivity of PM_{2.5} to VOC being the smallest, its relative
353 uncertainty becomes the largest. Second-order sensitivities associated with NO_x are relatively
354 smaller than the others, so the uncertainty of its first order sensitivity is relatively smaller.

355 Comparison of the uncertainties of the five first-order sensitivities indicates that a first-
356 order sensitivity to a given emission source has less uncertainty associated with emissions if the
357 magnitude of first-order sensitivity is large, i.e., this source has a relatively larger contribution to
358 PM_{2.5} concentration than the other sources, and if first-order sensitivity is much larger (over 10
359 times in this study) than the magnitudes of its associated second-order sensitivities. This sheds
360 light on whether or not to include certain second-order terms in the RFM: if a second-order

361 sensitivity has comparable magnitude as its associated first-order sensitivities, it needs to be
362 included in the RFM since it has similar contribution to error propagation as the first-order ones.

363

364

365 **4. Conclusions**

366 A reduced form model of CMAQ is applied to propagate the emission uncertainties in
367 Monte Carlo simulations to estimate the related uncertainties in simulated $PM_{2.5}$ concentrations
368 and $PM_{2.5}$ sensitivities. The reduced form model is constructed based on first- and second-order
369 sensitivities obtained from high-order DDM sensitivity analysis in the CMAQ model. The
370 application of the reduced form model saves a substantial amount of computational time (2 - 3
371 orders of magnitude) in this uncertainty analysis. 1000 possible combinations of emissions rates
372 of five major pollutants are sampled based on a log-normal distribution function with uncertainty
373 factors estimated from a literature search. The ensemble output from the reduced form model is
374 used to quantify the model uncertainty associated with emission rates. The relative uncertainty of
375 modeled 24-hour average of $PM_{2.5}$ concentration is estimated to fall between 42% and 52% for
376 different simulated $PM_{2.5}$ levels. The spatial distribution of the relative uncertainty is fairly
377 uniform over the entire modeling domain.

378 Comparison of the normalized mean error between simulated and observed $PM_{2.5}$ for
379 different concentration levels suggests that the emission-associated uncertainties can account for
380 a majority of the model error, though the persistent low bias in the summer is attributed to a bias
381 in SOA formation. The time series of simulated and observed daily $PM_{2.5}$ concentrations found
382 that the observations are well captured by model simulation when the emission uncertainties are

383 included. In total, 85% of the measured PM_{2.5} concentrations fall into the 95% confidence
384 interval due to the uncertainty in emission rates and 60 % of the measured PM_{2.5} concentrations
385 fall into the standard deviations. The temporal and spatial trends are well captured in the base
386 simulation. This suggests that much, but not all, of the difference between the observed and
387 simulated concentrations can be attributed to the uncertainty in emission rates. The histogram of
388 errors between simulated and observed PM_{2.5} concentrations has a shape close to a log-normal
389 distribution, with estimated the mean and standard deviation to be $-1.5 \mu\text{g m}^{-3}$ and $8.6 \mu\text{g m}^{-3}$,
390 respectively, and the standard deviation due to uncertain emissions is $7.3 \mu\text{g m}^{-3}$, which is
391 comparable to the standard deviation of the model error. While the emissions uncertainties,
392 alone, can explain most of the errors in the model results, other factors, including errors in the
393 meteorological inputs and model parameters, will contribute. We have attempted to minimize
394 the impacts of errors in the meteorological inputs by using detailed meteorological model results
395 that have been thoroughly evaluated well (Angevine et al., 2009, Byun et al., 2011).

396 This paper also demonstrated how to use the RFM approach for estimating the emission-
397 associated uncertainty of sensitivities. Averaged over the HGB area, the emission-associated
398 relative uncertainty of first order sensitivities of PM_{2.5} to primary PM, NO_x, SO₂, NH₃, and VOC
399 emissions are 16%, 32%, 55%, 75%, and 128%, respectively. The first-order sensitivity of PM_{2.5}
400 to primary PM emissions is much higher than the other first-order sensitivities, so it has the
401 lowest uncertainty. Uncertainty of a first-order sensitivity to a precursor depends on its
402 associated second-order sensitivities, which determine the magnitude of standard deviation of
403 first-order sensitivity. For example, the first-order sensitivity of PM_{2.5} to VOCs has the largest
404 uncertainty associated with emissions because the VOCs-associated second-order sensitivities
405 have the highest magnitude compared to the others. This indicates that first-order sensitivity has

406 less uncertainty associated with emissions if it is much larger in magnitude (over 10 times in this
407 study) than its associated second-order sensitivities, and that it would be safe to only include this
408 first-order sensitivity in building a RFM which is a function of emission changes.

409 Although the emission-associated uncertainties can explain much of the errors in $PM_{2.5}$
410 simulations, meteorological conditions and model representation of chemical reactions also play
411 an important role. Part of the uncertainty in emissions rates is due to the uncertainty in
412 meteorological fields. In this study, model errors at certain times and locations are associated
413 with relatively high bias in simulated temperature and wind directions. A correlation analysis
414 indicates that the domain-wide model errors in $PM_{2.5}$ simulation is more related to errors in wind
415 direction than to errors in temperature, wind speed, and relative humidity. Besides
416 meteorological conditions, model representation of chemical reactions also contributes to model
417 bias in the $PM_{2.5}$ concentration. The biased-low $PM_{2.5}$ simulation north of Houston suggests that
418 there is a bias in the model representation of the formation of secondary organic aerosol and/or a
419 bias in biogenic VOC emissions in the inventory.

420

421

Tables

422 Table 1. Uncertainty factors and associated σ (standard deviations of log-transformed data) of
423 emission rates of five major pollutants.

Emissions	Uncertainty Factor	σ
SO ₂	1.62	0.243
NO _x	1.67	0.258
VOC	2.11	0.373
NH ₃	2.74	0.505
Primary PM	2.71	0.500

424

425

426 Table 2. Evaluation of WRF-generated meteorological fields from Aug 10 to Sep 14, 2006 with
 427 the Techniques Development Laboratory (TDL) surface observations.
 428

Model Domain Resolution	Surface Wind Speed		Surface Wind Direction		Surface Air Temperature		Surface Humidity	
	Bias (m s⁻¹)	RMSE* (m s⁻¹)	Bias (deg.)	Gross Error (deg.)	Bias (K)	RMSE (K)	Bias (g kg⁻¹)	Gross Error (g kg⁻¹)
36km	0.23	1.76	2.49	32.78	-0.43	1.96	0.44	0.92
12km	0.53	2.27	14.90	52.08	-0.92	3.34	-0.13	1.18
4km	0.47	1.91	10.50	53.17	0.63	2.29	0.29	1.46

429 *RMSE: root mean square error

430

431 Table 3. Summary of the comparison between simulated and observed daily average PM_{2.5}
432 concentrations.
433

Site Name	Site Number	Fraction of Days within One Standard Deviation	Fraction of days within 95% CI
Houston East	482011034	83%	93%
Channel View	482010026	69%	97%
Deer Park	482011039	59%	84%
Park Place	482010416	50%	75%
Kingwood	482011042	38%	75%
All Sites		60%	85%

434

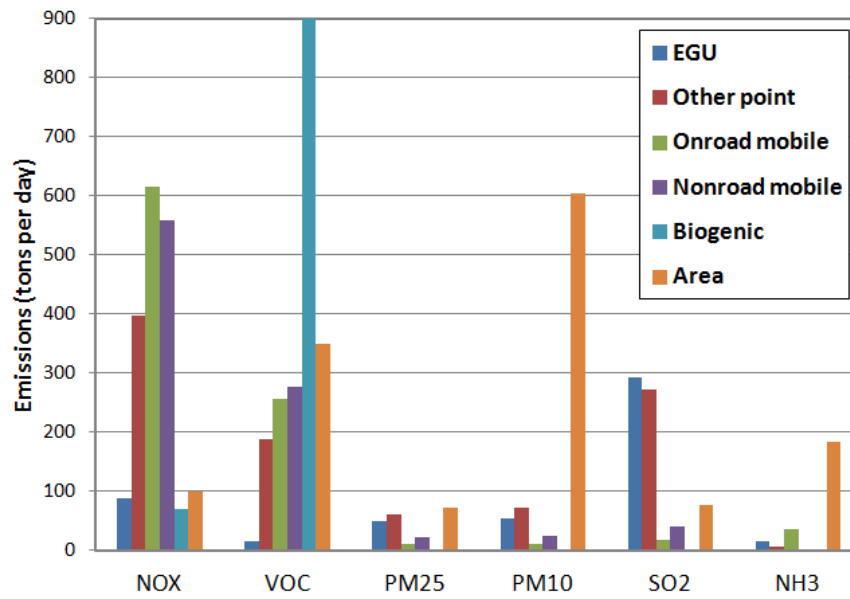
435

436 Table 4. First- and second-order sensitivities of PM_{2.5} to emissions. First order sensitivities are
 437 calculated as $S_i^{(1)} = \frac{\partial C_{PM_{2.5}}}{\partial E_i}$, second-order self-sensitivities are calculated as $S_{i,i}^{(2)} = \frac{\partial^2 C_{PM_{2.5}}}{\partial E_i^2}$, and
 438 second-order cross-sensitivities are calculated as $S_{i,j}^{(2)} = \frac{\partial^2 C_{PM_{2.5}}}{\partial E_i \partial E_j}$. E_i and E_j are the i^{th} and j^{th}
 439 emissions rates, respectively. The values are daily averages over the HGB area on September 8th,
 440 2006. The unit is $\mu\text{g m}^{-3}$.
 441

E_i	First Order	Second Order (E_j)				
		NOx	PM	SO ₂	NH ₃	VOC
NOx	1.3	0.48	0.003	-0.06	-0.002	0.07
PM	5.2		-0.76	-0.37	-0.4	-0.47
SO ₂	1.8			-0.76	-0.4	-0.59
NH ₃	0.9				-0.88	-0.52
VOC	0.8					-1.16

442

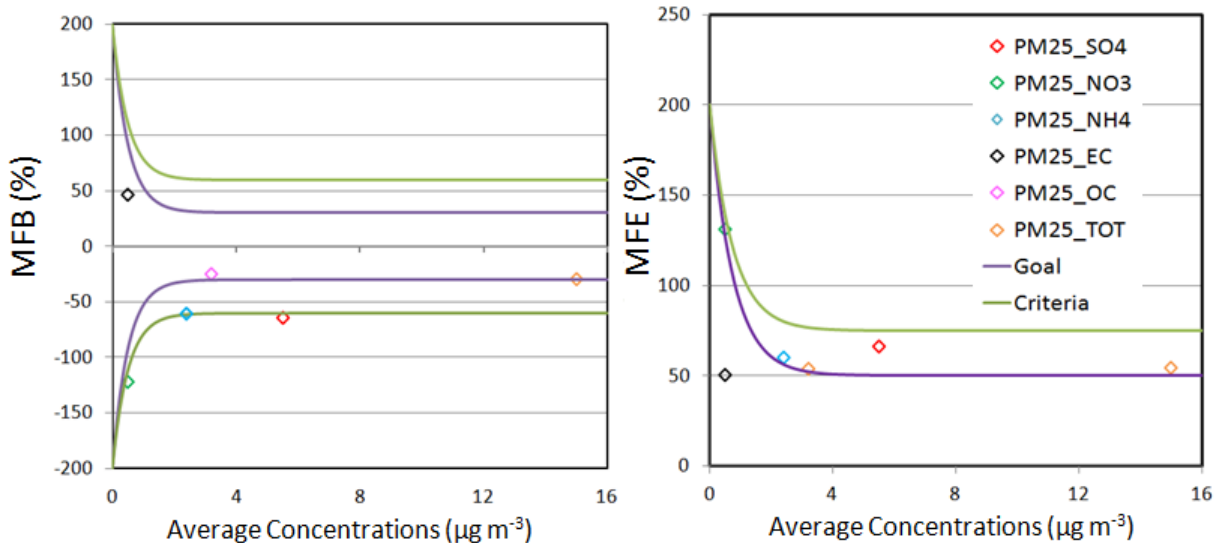
Figures



444

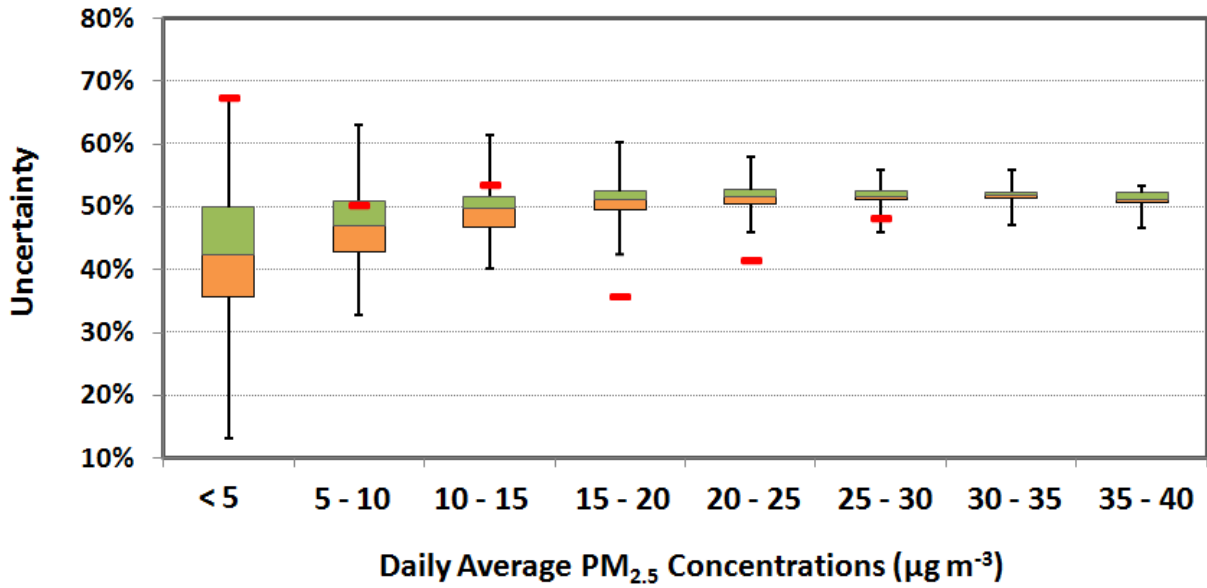
445 Figure 1. Emission rates of six major pollutants categorized by emission sources in 2005 NEI.
 446 The emission rates are the daily averages of the domain-wide emissions.

447



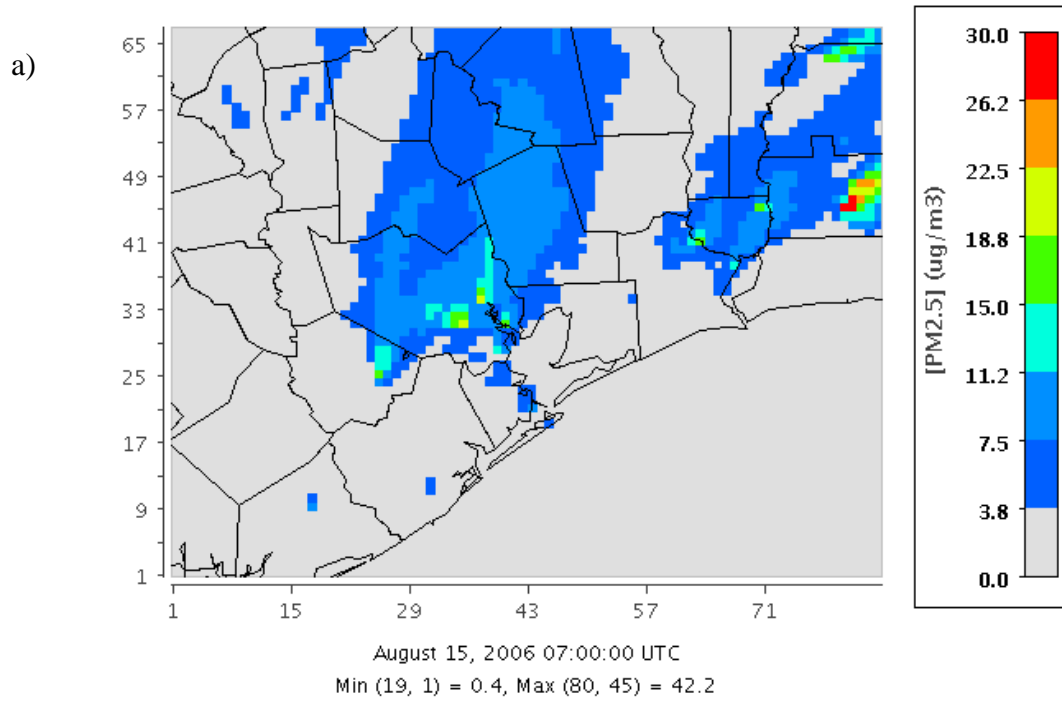
448
 449 Figure 2. Bugle plots of CMAQ performance of 24-hour average total and speciation
 450 concentrations of PM_{2.5}. MFB stands for mean fractional bias and MFE stands for mean
 451 fractional error. Goal (purple line) and Criteria (green line) are obtained from Boylan and
 452 Russell, 2006. MFB is calculated using $\frac{2}{N} \sum_{i=1}^N \frac{C_m - C_o}{C_o + C_m}$, and MFE is calculated using
 453 $\frac{2}{N} \sum_{i=1}^N \frac{|C_m - C_o|}{C_o + C_m}$, where C_m and C_o are the modeled and observed PM_{2.5} concentrations,
 454 respectively. Different markers represent PM_{2.5} and its species, e.g., PM25_TOT means total
 455 PM_{2.5}, and PM25_SO4 means sulfate aerosol.

456

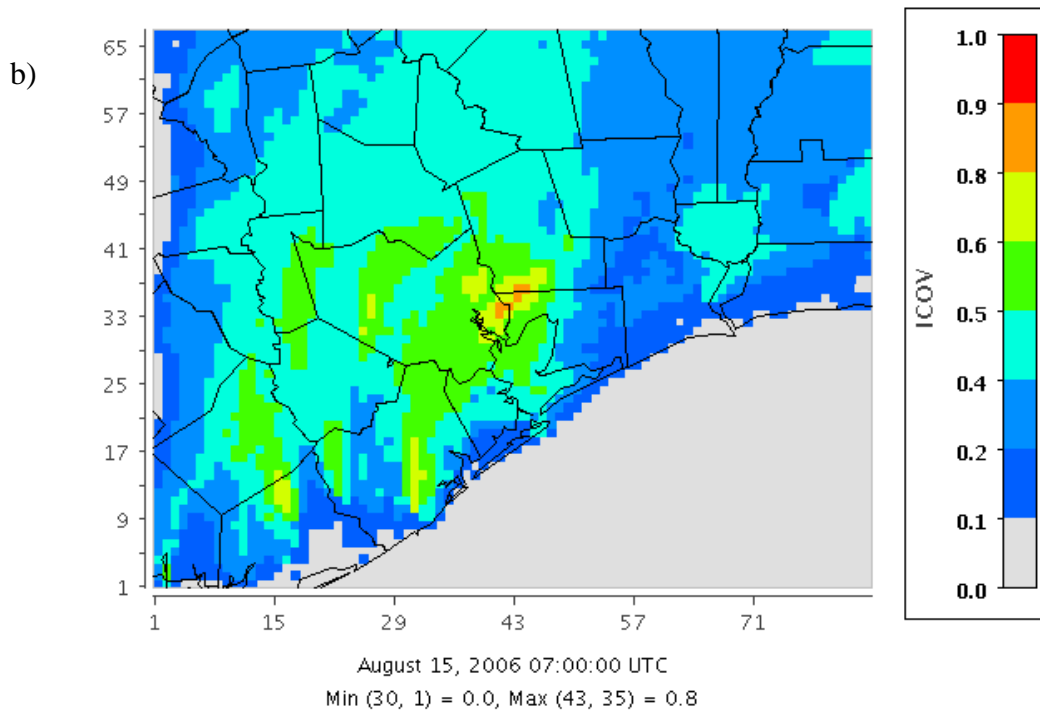


457

458 Figure 3. Relative uncertainty in simulated PM_{2.5} concentrations over the modeling domain. The
 459 box shows median, 25th and 75th percentiles. The line between the green and orange boxes
 460 represents the median. The whiskers indicate a 95% confidence interval. The red bars indicate
 461 the normalized mean errors between observed and simulated PM_{2.5} concentrations at all the
 462 observation sites in the modeling domain.
 463



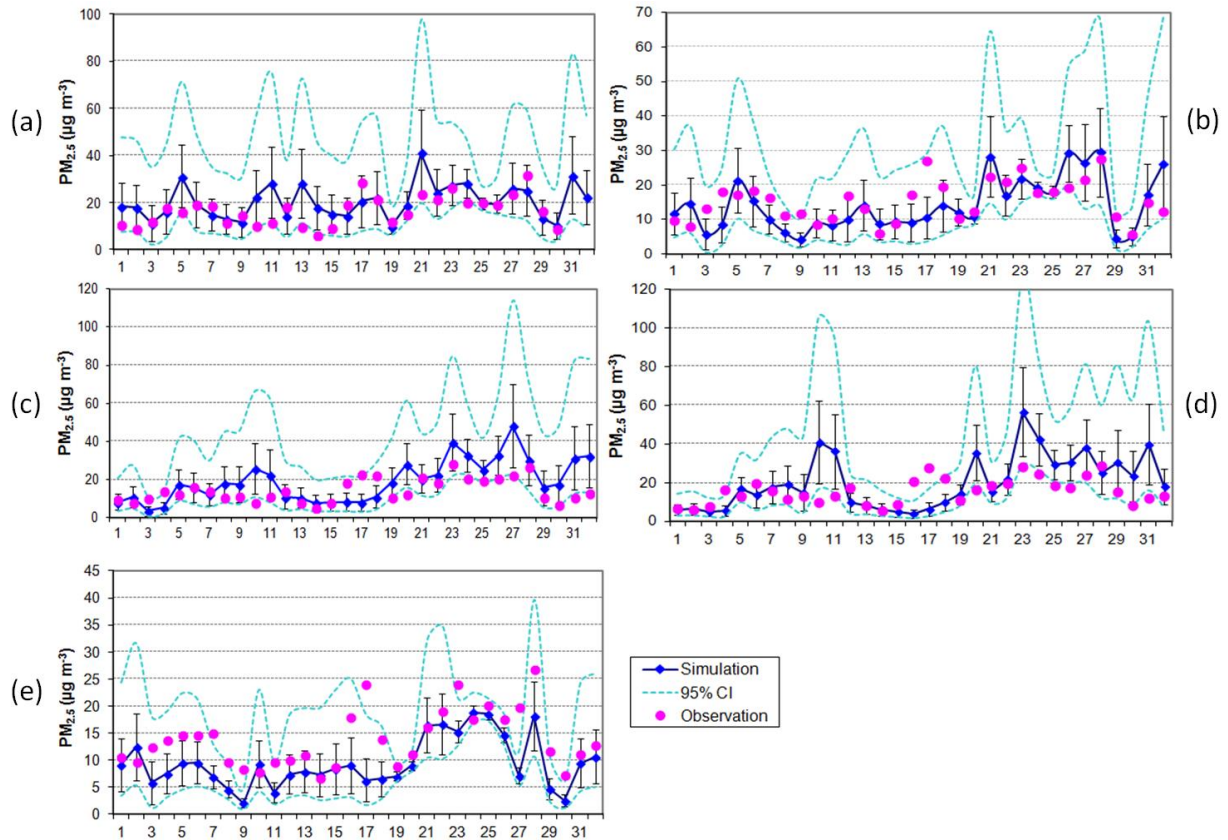
464



465

466 Figure 4. Spatial distribution of a) daily average PM_{2.5} concentrations and b) uncertainty on
467 August 15, 2006.

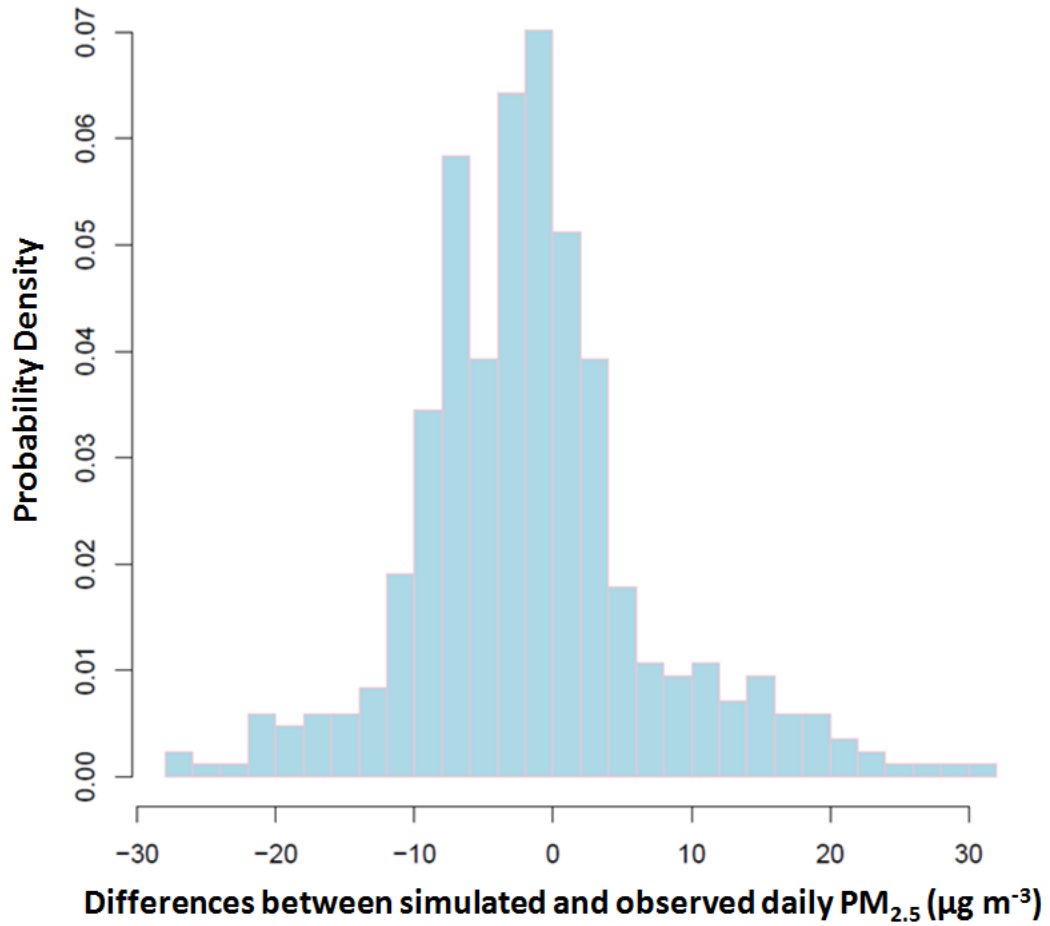
468



469
470

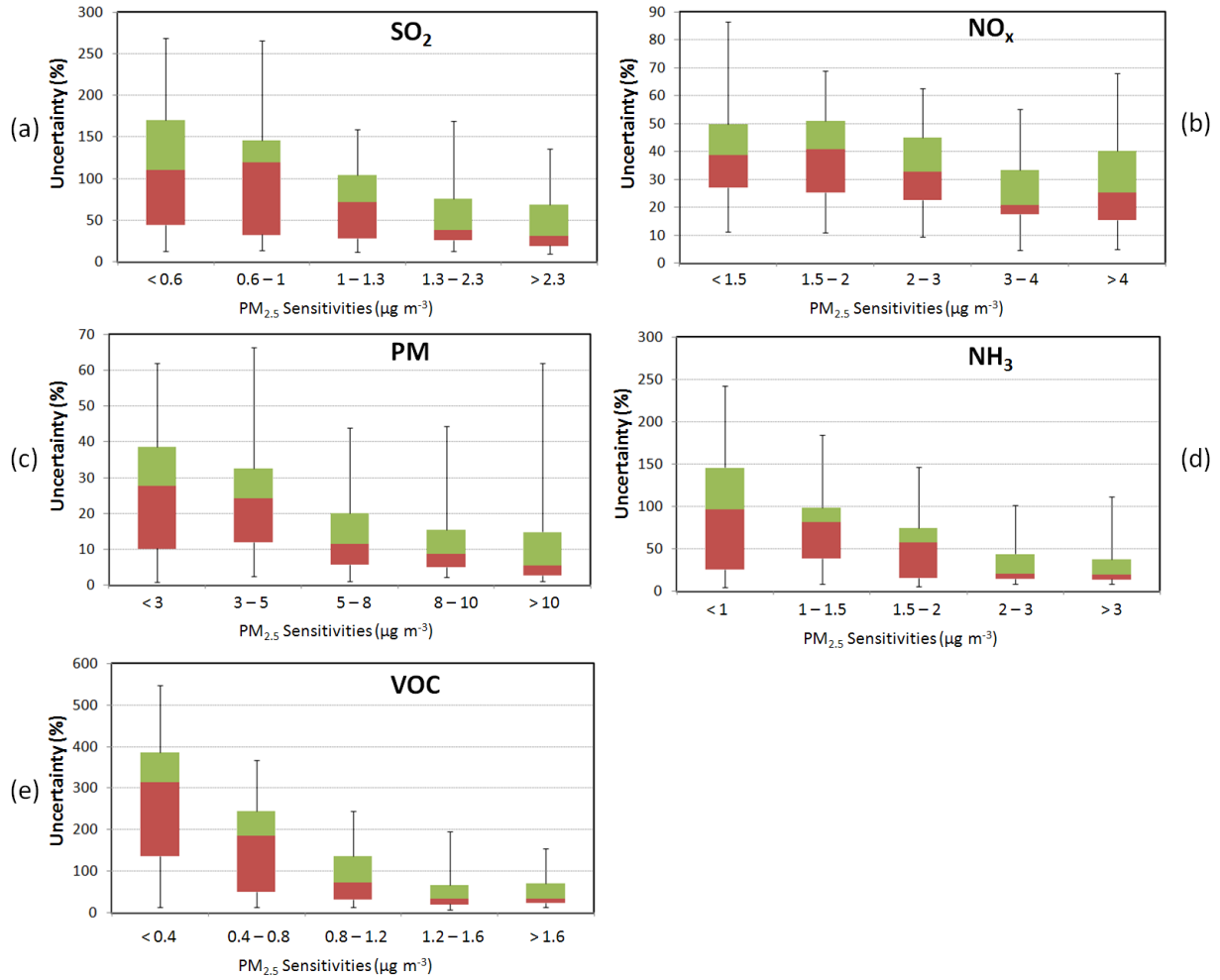
471 Figure 5. Time series of daily average $PM_{2.5}$ concentrations for the five AQS sites: a) Houston
 472 East (AQS#: 482011034), b) Channelview (AQS#: 482010026, c) Deer Park (AQS#:
 473 482011039), d) Park Place (AQS#: 482010416), e) Kingwood (AQS#: 482011042). Blue line
 474 with dots stand for simulated $PM_{2.5}$ concentrations; magenta dots stand for observed $PM_{2.5}$
 475 concentrations; light blue dashed line stands for the 95% CI of simulated $PM_{2.5}$ concentrations;
 476 error bars correspond to 68.3% CI of $PM_{2.5}$ concentrations, which is equivalent to one standard
 477 deviation range.

478
479



480
481
482
483
484
485
486
487
488
489
490

Figure 6. Histogram of differences between CMAQ simulated and observed PM_{2.5} concentrations at all the monitoring sites in the 4km modeling domain.



491
492

493 Figure 7. Uncertainty in simulated PM_{2.5} sensitivities due to uncertainties in domain-wide
 494 emissions rates of (a) SO₂, (b) NO_x, (c) PM, (d) NH₃, and (e) VOC of the HGB area on
 495 September 8, 2007. The box shows median, 25th and 75th percentiles. The line between the green
 496 and orange boxes represents the median. The whiskers indicate a 95% confidence interval.

497 **Acknowledgements:**

498

499 This research was made possible by funding from Phillips 66 Inc., Health Effects Institute, and
500 US EPA STAR grant R834799. While this work was supported, in part, by grant from the US
501 EPA, its contents are solely the responsibility of the grantee and do not necessarily represent the
502 official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial
503 products or services mentioned in the publication. We also acknowledge the contribution of
504 Sergey L. Napelenok of US EPA and Di Tian of Georgia EPD for helpful discussions.

505 **Reference**

- 506 Angevine, W.M., Zagar, M., Brioude, J., 2009. How well can we model the Houston area for air
507 pollution studies? *Proceedings, Eighth Symposium on the Urban Environment*, January
508 2009, Phoenix, AZ. http://ams.confex.com/ams/89annual/techprogram/paper_143826.htm.
- 509 Banta, R. M., Senff, C. J., Alvarez, R. J., Langford, A. O., Parrish, D. D., Trainer, M. K., White,
510 A. B. (2011). Dependence of daily peak O₃ concentrations near Houston, Texas on
511 environmental factors: Wind speed, temperature, and boundary-layer depth. *Atmospheric*
512 *Environment*, 45(1), 162-173. doi: 10.1016/j.atmosenv.2010.09.030.
- 513 Bergin, M. S., Noblet, G. S., Petrini, K., Dhieux, J. R., Milford, J. B., & Harley, R. A. (1999).
514 Formal uncertainty analysis of a Lagrangian photochemical air pollution model.
515 *Environmental Science & Technology*, 33(7), 1116-1126. doi: 10.1021/es980749y.
- 516 Bergin, M. S., Russell, A. G., Odman, M. T., Cohan, D. S., & Chameldes, W. L. (2008). Single-
517 Source Impact Analysis Using Three-Dimensional Air Quality Models. *Journal of the Air &*
518 *Waste Management Association*, 58(10), 1351-1359. doi: 10.3155/1047-3289.58.10.1351.
- 519 Boylan, J. W., & Russell, A. G. (2006). PM and light extinction model performance metrics,
520 goals, and criteria for three-dimensional air quality models. *Atmospheric Environment*,
521 40(26), 4946-4959. doi: 10.1016/j.atmosenv.2005.09.087.
- 522 Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and
523 other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling
524 system. *Applied Mechanics Reviews*, 59(1-6), 51-77. doi: 10.1115/1.2128636.
- 525 Byun, D., Kim, H., Ngan, F. (2011) Improvement of Meteorological Modeling by Accurate
526 Prediction of Soil Moisture in the Weather Research and Forecasting (WRF) Model.
527 Technical Report submitted to TCEQ, available at
528 [https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/5828](https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/582886246FY1009-NOAA_WRF_Soil_Moisture_20110331.pdf)
529 [86246FY1009-NOAA_WRF_Soil_Moisture_20110331.pdf](https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/582886246FY1009-NOAA_WRF_Soil_Moisture_20110331.pdf).
- 530 Carlton, A. G., Bhave, P. V., Napelenok, S. L., Edney, E. D., Sarwar, G., Pinder, R. W.,
531 Houyoux, M. (2010). Model Representation of Secondary Organic Aerosol in
532 CMAQv4.7. *Environmental Science & Technology*, 44(22), 8553-8560. doi:
533 10.1021/es100636q
- 534 Carter, W.P.L. (2000). Documentation of the SAPRC99 chemical mechanism for VOC and
535 reactivity assessment, Air Pollution Research Center and College of Engineering, Center for
536 Environmental Research and Technology, University of California, Riverside, CA.

- 537 Civerolo, K., Hogrefe, C., Zalewsky, E., Hao, W., Sistla, G., Lynn, B., Rosenzweig, C., Kinney,
538 P.L. (2010). Evaluation of an 18-year CMAQ simulation: Seasonal variations and long-term
539 temporal changes in sulfate and nitrate. *Atmospheric Environment*, 44, 6745 – 3752.
- 540 Cohan, D. S., Koo, B., & Yarwood, G. (2010). Influence of uncertain reaction rates on ozone
541 sensitivity to emissions. *Atmospheric Environment*, 44(26), 3101-3109. doi:
542 10.1016/j.atmosenv.2010.05.034.
- 543 Cowling, E. B., Furiness, C., Dimitriadis, B., Parrish, D., and Estes, M. (2007). Final Rapid
544 Science Synthesis Report: Findings from the Second Texas Air Quality Study, Southern
545 Oxidants Study Office of the Director at North Carolina State University Location: Raleigh,
546 NC, USA
- 547 Deguillaume, L., Beekmann, M., and Derognat, C. (2008). Uncertainty evaluation of ozone
548 production and its sensitivity to emission changes over the Ile-de-France region during
549 summer periods. *Journal of Geophysical Research-Atmospheres*, 113(D2). doi:
550 10.1029/2007jd009081
- 551 Digar, A., & Cohan, D. S. (2010). Efficient Characterization of Pollutant-Emission Response
552 under Parametric Uncertainty. *Environmental Science & Technology*, 44(17), 6724-6730.
553 doi: 10.1021/es903743t
- 554 Dockery, D. W., Pope, C. A., Xu, X. P., Spengler, J. D., Ware, J. H., Fay, M. E., Speizer, F. E.
555 (1993). An Association Between Air-Pollution and Mortality in Six United-States Cities.
556 *New England Journal of Medicine*, 329(24), 1753-1759. doi:
557 10.1056/nejm199312093292401.
- 558 Emery, et al. (2001) Enhanced meteorological modeling and performance evaluation for two
559 Texas ozone episodes, ENVIRON, Novato, CA
560
561
- 562 ENVIRON. (2005) Users' Guide, Comprehensive Air Quality Model with Extensions (CAMx),
563 Version 4.20; ENVIRON International Corporation.
564
565
- 566 Foley, K. M., Roselle, S. J., Appel, K. W., Bhave, P. V., Pleim, J. E., Otte, T. L., Bash, J. O.
567 (2010). Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling
568 system version 4.7. *Geoscientific Model Development*, 3(1), 205-226.
569

- 570 Galloway, T. S., Brown, R. J., Browne, M. A., Dissanayake, A., Lowe, D., Jones, M. B., and
571 Depledge, M. H. (2004). A multibiomarker approach to environmental assessment.
572 *Environmental Science & Technology*, 38(6), 1723-1731. doi: 10.1021/es030570.
- 573 Hakami, A., Odman, M. T., & Russell, A. G. (2003). High-order, direct sensitivity analysis of
574 multidimensional air quality models. *Environmental Science & Technology*, 37(11), 2442-
575 2452. doi: 10.1021/es020677h.
- 576 Hanna, S. R., Chang, J. C., & Fernau, M. E. (1998). Monte Carlo estimates of uncertainties in
577 predictions by a photochemical grid model (UAM-IV) due to uncertainties in input
578 variables. *Atmospheric Environment*, 32(21), 3619-3628. doi: 10.1016/s1352-
579 2310(97)00419-6
- 580 Hanna, S. R., Lu, Z. G., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Hansen, D. A.
581 (2001). Uncertainties in predicted ozone concentrations due to input uncertainties for the
582 UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric*
583 *Environment*, 35(5), 891-903. doi: 10.1016/s1352-2310(00)00367-8
- 584 Hanna, S. R., & Yang, R. X. (2001). Evaluations of mesoscale models' simulations of near-
585 surface winds, temperature gradients, and mixing depths. *Journal of Applied Meteorology*,
586 40(6), 1095-1104. doi: 10.1175/1520-0450(2001)040<1095:eommso>2.0.co;2
- 587 Kaiser, J. (2005). Mounting evidence indicts fine-particle pollution. *Science*, 307(5717), 1858-
588 1861.
- 589 Kim, S.-W., McKeen, S. A., Frost, G. J., Lee, S.-H., Trainer, M., Richter, A., Angevine, W. M.,
590 Atlas, E., Bianco, L., Boersma, K. F., Brioude, J., Burrows, J. P., de Gouw, J., Fried, A.,
591 Gleason, J., Hilboll, A., Mellqvist, J., Peischl, J., Richter, D., Rivera, C., Ryerson, T.,
592 te Lintel Hekkert, S., Walega, J., Warneke, C., Weibring, P., and Williams, E.: Evaluations
593 of NO_x and highly reactive VOC emission inventories in Texas and their implications for
594 ozone plume simulations during the Texas Air Quality Study 2006, *Atmos. Chem. Phys.*, 11,
595 11361-11386, doi:10.5194/acp-11-11361-2011, 2011.
- 596 Kwok, R.H.F., Napelenok, S.L., Baker, K.R. (2013). Implementation and evaluation of PM_{2.5}
597 source contribution analysis in a photochemical model. *Atmospheric Environment*, 80, 398
598 – 407.
- 599 Latha, K. M., & Badarinath, K. V. S. (2005). Shortwave radiative forcing efficiency of urban
600 aerosols - a case study using ground based measurements. *Chemosphere*, 58(2), 217-220.
601 doi: 10.1016/j.chemosphere.2004.09.013.

- 602 Lefer, B., Rappengluck, B., Flynn, J., & Haman, C. (2010). Photochemical and meteorological
603 relationships during the Texas-II Radical and Aerosol Measurement Project (TRAMP).
604 *Atmospheric Environment*, 44(33), 4005-4013. doi: 10.1016/j.atmosenv.2010.03.011.
- 605 Napelenok, S. L., Cohan, D. S., Hu, Y. T., & Russell, A. G. (2006). Decoupled direct 3D
606 sensitivity analysis for particulate matter (DDM-3D/PM). *Atmospheric Environment*, 40(32),
607 6112-6121. doi: 10.1016/j.atmosenv.2006.05.039.
- 608 Napelenok, S. L., Foley, K. M., Kang, D. W., Mathur, R., Pierce, T., and Rao, S. T. (2011).
609 Dynamic evaluation of regional air quality model's response to emission reductions in the
610 presence of uncertain emission inventories. *Atmospheric Environment*, 45(24), 4091-4098.
611 doi: 10.1016/j.atmosenv.2011.03.030
- 612 NARSTO (2004) Particulate Matter Science for Policy Makers: A NARSTO Assessment. P.
613 McMurry, M. Shepherd, and J. Vickery, eds. Cambridge University Press, Cambridge,
614 England. ISBN 0 52 184287 5.
- 615 Pinder, R. W., Gilliam, R. C., Appel, K. W., Napelenok, S. L., Foley, K. M., and Gilliland, A. B.
616 (2009). Efficient Probabilistic Estimates of Surface Ozone Concentration Using an
617 Ensemble of Model Configurations and Direct Sensitivity Calculations. *Environmental*
618 *Science & Technology*, 43(7), 2388-2393. doi: 10.1021/es8025402.
- 619 Ryerson, T. B., Trainer, M., Angevine, W. M., Brock, C. A., Dissly, R. W., Fehsenfeld, F. C.,
620 Sueper, D. T. (2003). Effect of petrochemical industrial emissions of reactive alkenes and
621 NO_x on tropospheric ozone formation in Houston, Texas. *Journal of Geophysical Research-*
622 *Atmospheres*, 108(D8). doi: 10.1029/2002jd003070
- 623 Schwartz, J. (1994). PM₁₀, Ozone, and Hospital Admissions for the Elderly in Minneapolis-Paul,
624 Minnesota. *Archives of Environmental Health*, 49(5), 366-374.
- 625 Tian, D., Cohan, D. S., Napelenok, S., Bergin, M., Hu, Y. T., Chang, M., & Russell, A. G.
626 (2010). Uncertainty Analysis of Ozone Formation and Response to Emission Controls
627 Using Higher-Order Sensitivities. *Journal of the Air & Waste Management Association*,
628 60(7), 797-804. doi: 10.3155/1047-3289.60.7.797.
- 629 U.S. EPA (2004) Air quality criteria for particulate matter. EPA/600/P-99/002aF.
- 630 U.S. EPA. (2007). Guidance on the Use of Models and Other Analyses for Demonstrating
631 Attainment of Air Quality Goals for Ozone, PM_{2.5}, and Regional Haze. EPA -454/B-07-002.

- 632 Xiao, X., Cohan, D. S., Byun, D. W., & Ngan, F. (2010). Highly nonlinear ozone formation in
633 the Houston region and implications for emission controls. *Journal of Geophysical*
634 *Research-Atmospheres*, 115. doi: 10.1029/2010jd014435.
- 635 Watson, J. G. (2002). Visibility: Science and regulation. *Journal of the Air & Waste*
636 *Management Association*, 52(6), 628-713.
- 637 Wert, B. P., Trainer, M., Fried, A., Ryerson, T. B., Henry, B., Potter, W., Wisthaler, A. (2003).
638 Signatures of terminal alkene oxidation in airborne formaldehyde measurements during
639 TexAQS 2000. *Journal of Geophysical Research-Atmospheres*, 108(D3). doi:
640 10.1029/2002jd002502.
- 641 Yang, Y. J., J. G. Wilkinson, and A. G. Russell (1997) Fast, direct sensitivity analysis of
642 multidimensional photochemical models. *Environmental Science & Technology*, 31, 2859-
643 2868.
- 644 Yu, S, Dennis, R, Roselle, S., Nenes, A., Walker, J., Eder, B., Schere, K., Swall, J. and Robarge,
645 W. (2005). An assessment of the ability of three-dimensional air quality mode with current
646 thermodynamic equilibrium models to predict aerosol NO₃. *Journal of Geophysical*
647 *Research*, 110, D07S13, doi:10.1029/2004JD004718.
- 648 Zanobetti, A., Schwartz, J., and Dockery, D. W. (2000). Airborne particles are a risk factor for
649 hospital admissions for heart and lung disease. *Environmental Health Perspectives*, 108(11),
650 1071-1077. doi: 10.1289/ehp.001081071.
- 651 Zhang, W., Capps, S. L., Hu, Y., Nenes, A., Napelenok, S. L., and Russell, A. G. (2012).
652 Development of the high-order decoupled direct method in three dimensions for particulate
653 matter: enabling advanced sensitivity analysis in air quality models. *Geoscientific Model*
654 *Development*, 5(2), 355-368. doi: 10.5194/gmd-5-355-2012
- 655
656 Zhang, Y., Vijayaraghavan, K., Wen, X., Snell, H.E., and Jacobson, M.Z. (2009). Probing into
657 regional ozone and particulate matter pollution in the United States: 1. A 1 year CMAQ
658 simulation and evaluation using surface and satellite data. *Journal of Geophysical Research*,
659 114, D22304, doi:10.1029/2009JD011898.