The impact of observation nudging on simulated meteorology and ozone concentrations during DISCOVER-AQ 2013 Texas campaign

Xiangshang Li [1], Yunsoo Choi [1], Beata Czader [1], Anirban Roy [1], Hyuncheol Kim [2,3], Barry Lefer [1], Shuai Pan [1]

[1]Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, 77204, USA

[2]NOAA Air Resources Laboratory, College Park, MD 20740, USA

[3]University of Maryland, Cooperative Institute for Climate and Satellite, College Park, MD, USA

Corresponding author: Xiangshang Li: xli@central.uh.edu

Abstract

Accurate meteorological fields are imperative for correct chemical transport modeling. Observation nudging, along with objective analysis, is generally considered as a low-cost and effective technique to improve meteorological simulations. However, the meteorological impact of observation nudging on chemistry has not been well characterized. This study involved two simulations to analyze the impact of observation nudging on simulated meteorology and ozone concentrations during the 2013 Deriving Information on Surface conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) Texas campaign period, using the Weather Research and Forecasting (WRF) and Community Multiscale Air Quality (CMAQ) models. The results showed improved correlations between observed and simulated parameters. For example, the index of agreement (IOA) improved by about 9% for surface temperature and 6-11% for surface zonal (U-WIND) and meridional (V-WIND) winds when observation nudging was employed. Analysis of a cold front event indicated that nudging improved the timing of wind transition during the front passage. Observation nudging also

27   reduced the model biases for the planetary boundary layer (PBL) height predictions.

28   Additionally, the IOA for CMAQ simulated surface ozone improved by 6% during the

29   simulation period. The high ozone episode on September 25th was a post-front ozone event in

30   Houston. The small-scale morning wind shifts near the Houston Ship Channel combined with

31   higher aloft ozone early morning likely caused the day's ozone exceedance. While observation

32   nudging did not recreate the wind shifts on that day and failed to reproduce the observed high

33   ozone, analyses of surface and aircraft data found that observation nudging helped model yield

34   improved ozone predictions. In a two-hour period during the event, substantially better winds in

35   the sensitivity case noticeably improved the ozone. The average IOA for ozone in the period

36   increased from just over 0.4 to near 0.7. Further work on improving the capability of nudging to

37   reproduce local meteorological events such as stagnations and wind reversals could enhance a

38   chemical transport model's skill to predict high ozone events.

39   **Keywords**: WRF, CMAQ, air quality model, DISCOVER-AQ, observation nudging

40

1. Introduction

42   Meteorological variables such as cloud fraction, winds, planetary boundary layer (PBL) heights

43   and precipitation significantly impact air quality. They influence the production, transport, and

44   deposition of various chemical species (e.g., Banta et al. 2005; Pour-Biazar et al. 2007; Cuchiara

45   et al. 2014). Hence accurate meteorological inputs are imperative for air quality modeling.

46   Common approaches of improving meteorological simulations include the selection of updated

47   and high resolution terrain data (e.g., Cheng and Byun 2008), optimization of physics and

48   dynamics options (e.g., Zhong et al. 2007) and the implementation of Four Dimensional Data

49   Assimilation (FDDA) (e.g., Stauffer and Seaman 1994).

50   Four Dimensional Data Assimilation continuously merges new observational data into model

51   simulation such that the model predictions do not drift away from observations. There are several

52   FDDA methods including nudging (e.g., Stauffer and Seaman 1994) and Variational Methods

53   (3D-VAR or 4D-VAR; e.g., Le Dimet and Talagrand 1986; Huang et al. 2009). 4D-VAR obtains

54   optimal states of the atmosphere using multi-time-level observations by globally adjusting a

55   model solution to all available observations over an interval of time. Nudging is a simple yet

56   flexible FDDA method originally developed by Stauffer and Seaman (1990, 1994), and

57   implemented in the Fifth-Generation PSU/NCAR Mesoscale Model (MM5). Not intended for

58   optimal adjustment, nudging is less computationally intensive but needs special care for the

59   nudging coefficients. Nudging involves adding an artificial tendency term to one or more model

60   prognostic equations that reflect the difference between the best estimate of the observed state

61   and the model state at a given location and time. In short, the goal is to "nudge" the model state

62   towards the observed state. There are several types of nudging such as 3D analysis nudging,

63   surface analysis nudging, and observation nudging (obs-nudging). In the case of analysis

64   nudging, the model state is nudged toward gridded analysis. The difference between 3D and

65   surface analysis nudging is that 3D analysis (at all model levels except for surface) data are used

66   in 3D analysis nudging to improve 3D fields while in surface analysis nudging only surface

67   analysis are used to improve surface fields. In observation nudging, the model predictions are

68   nudged to match better with observations at individual locations both surface and aloft. The

69   MM5 nudging codes were later improved and incorporated into the Weather Research and

70   Forecasting (WRF) model by Liu et al. (2005, 2006). The enhancements enable obs-nudging to

71   assimilate a large variety of direct or derived observations. In WRF, the inputs for obs-nudging

72   are generated by the WRF OBSGRID program. This program also performs Objective Analysis

73   (OA) to improve the quality of analysis nudging files. Objective Analysis updates first guess

74   meteorology analysis by incorporating observational data. Since obs-nudging is usually

75   performed along with OA (as in this study) to maximize the benefits of assimilating

76   observations, we also use OA to denote the combined Objective Analysis and obs-nudging

77   processes in case names.

78   The benefit of applying nudging to improve meteorological simulations has been demonstrated

79   by several studies (e.g., Deng 2009; Gilliam and Pleim 2010; Rogers et al. 2013). However, the

80   impact of the improved fields on air quality simulations has been investigated by relatively fewer

81   studies. Otte (2008) showed that the Community Multiscale Air Quality (CMAQ) model with

82   improved MM5 meteorology using analysis nudging was able to better simulate ozone chemistry

83   as reflected in model-measurement statistics. Their results indicated that better "model skill"

84   scores were achieved for daily maximum 1-hr ozone mixing ratios after analysis nudging over a

85   35-day simulation episode. Byun et al. (2008) performed over a dozen tests on obs-nudging (with

86   analysis nudging turned on) and showed obs-nudging improved both winds and temperature in

87   MM5 simulations. The study also gave an example in which improved wind fields on a given

88   day helped the CMAQ model better capture the ozone hotspot southwest of Houston. Ngan et al.

89   (2012) compared results from several CMAQ simulations coupled to the MM5 model which

90   included nudging. Their results indicated that fully nudged simulations (with both analysis

91   nudging and obs-nudging implemented) performed better than a forecast one with respect to both

92   meteorology and ozone chemistry. However, the statistics from their study cannot be used for

93   interpreting the sensitivity of obs-nudging since its base WRF case is a forecast run which used a

94   different analysis input. Previous studies by the current authors (e.g., Rappenglueck et al. 2011;

95   Czader et al. 2013) showed that obs-nudging helped correct errors in model wind fields, critical

96   to the transport of pollutants and production of secondary pollutants. To the best of the authors'

97   knowledge, there is no comprehensive study on the impact of obs-nudging on air quality

98   simulation using the WRF model.

99   This study intends to fill up the gap in the studies mentioned above by investigating the

100   sensitivity of WRF and subsequently, CMAQ simulations to observation nudging. In theory,

101   higher frequency of obs-nudging input should have a higher likelihood to capture small scale

102   events, such as local wind shifts. These events may only slightly impact local weather, yet have a

103   marked effect on chemistry. This is because local stagnation and wind convergence/reversals can

104   contribute to the pollutant build-up, as indicated by Banta et al. (1998), Cheung and Wang

105   (2001) and Tucker et al. (2010).

106   There is a significant presence of petrochemical facilities, power plants and motor vehicles in the

107   Houston-Galveston-Brazoria (HGB) region located in southeastern Texas (SETX). The major

108   pollutant in the region is ozone due to the abundant emissions of precursors like nitrogen oxides

109   (NO_x) and Volatile Organic Compounds (VOCs). Summertime ozone concentrations often rise

110   above the National Ambient Air Quality Standard (NAAQS). Consequently, HGB has been

111   designated as an ozone nonattainment region by the US Environmental Protection Agency

112   (USEPA) (http://www3.epa.gov/airquality/greenbook/hncs.html#TEXAS). The petrochemical

113   plants are largely concentrated in the Houston Ship Channel (HSC) area - just north of the

114   Galveston Bay. The VOCs emitted from the HSC area are highly reactive and have been shown

115   to contribute greatly to the high regional ozone episodes (e.g., Kleinman et al. 2002; Daum et al.

116   2004). Depending on the local meteorology, the plumes from HSC may be carried to different

117 locations in HGB and trigger high ozone events in its path. Metropolitan Houston has a high
118 level of NO_x emissions partly due to heavy urban vehicular traffic. As a result, relatively
119 frequent high ozone events occur in the area.

120 Due to the reasons listed above, the Houston-Galveston-Brazoria region has been the focus of
121 several air quality studies in the recent past (e.g., Banta et al. 2005; Parrish et al. 2009; Lefer and
122 Rappengluck 2010; Olaguer et al. 2013; Czader et al. 2013, Choi et al. 2012; Choi 2014; Choi
123 and Souri, 2015; Pan et al. 2015). It is a good place for studying ozone production and transport
124 due to the existence of a dense surface monitoring network, as well as several intensive
125 measurement field campaigns which provide ample observational data. For example in
126 September 2013, the National Aeronautics and Space Administration (NASA), joined by a
127 number of agencies and universities, conducted a field measurement campaign in SETX as part
128 of its the Deriving Information on Surface conditions from Column and Vertically Resolved
129 Observations Relevant to Air Quality (DISCOVER-AQ) program (http://www-
130 air.larc.nasa.gov/missions/discover-aq/discover-aq.html). This program has conducted several
131 air quality and meteorology measurements at different locations in the U.S. The availability of
132 dense surface observations is important for obs-nudging to correct erroneous local winds in the
133 model. The performance of obs-nudging will be handicapped without a rich set of observations.

134 This study involved performing two sets of WRF and CMAQ model simulations for the 2013
135 DISCOVER-AQ Texas time period in order to understand the impact of obs-nudging. The data
136 for nudging included multiple sets of observation data from both surface and aloft measurements.
137 We evaluated model-measurement performance statistics for both WRF and CMAQ output. The
138 paper is structured as following: Section 1 is introduction; Section 2 describes the measurement
139 data and the modeling system; Section 3 covers the evaluation protocols; Section 4 discusses the
140 general meteorological conditions during the campaign period; Section 5 presents the modeling
141 results, and Section 6 provides discussions and conclusions.

142 ## 2. Observational Data and Model Configurations

143 **2.1 Observational Data**

144 This study used regular measurements from the Continuous Ambient Monitoring Stations
145 (CAMS) operated by the Texas Commission on Environmental Quality (TCEQ). Additionally,

146   planetary boundary layer (PBL) and aloft ozone measurement data were obtained from the

147   DISCOVER-AQ campaign. For observation nudging, CAMS data and several data streams from

148   the Meteorological Assimilation Data Ingest System (MADIS) were used. The CAMS

149   measurement network collected real-time meteorology and pollutant data. The measured

150   parameters differ from station to station. The station density at South East Texas (SETX) is

151   relatively high. There were 63 and 52 sites with meteorological and ozone measurements

152   respectively in the 4-km domain. The network is represented in Figure 1. The sites are

153   represented by dots, with the La Porte (C556) site labeled. All CAMS observations are accessible

154   online at the TCEQ website: http://www.tceq.state.tx.us/cgi-

155   bin/compliance/monops/daily_summary.pl.
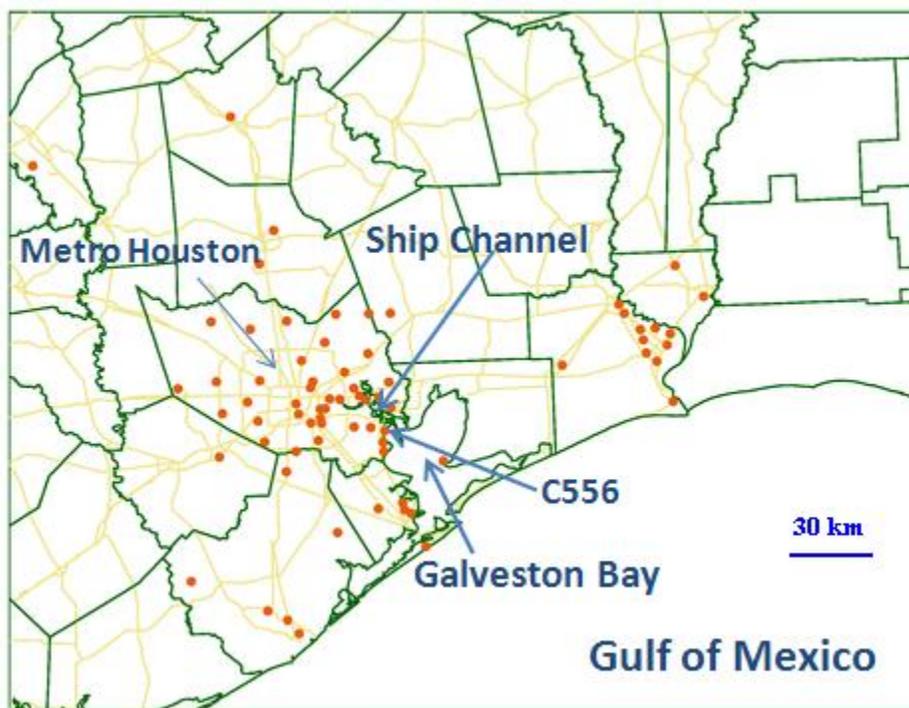
156
157
158   **Figure 1**. Locations of CAMS sites (dots) in CMAQ 4-km modeling domain during September

159   2013. Metro Houston, Houston Ship Channel, Galveston Bay and Gulf of Mexico are labeled.

160   Additionally, PBL height measurements for September were recorded at a site at the University

161   of Houston main campus. The PBL height was measured using the Light Detection and Ranging

162   (LIDAR) system. The PBL data is currently available only at this site. Primarily due to a

163   significant number of cloudy days, the PBL data coverage is incomplete (about 50%). For

164 analysis of aloft ozone, we also used measurements from aircraft P-3B, part of the rich datasets
165 collected during DISCOVER-AQ campaign. The P-3B data had over 100 parameters which are
166 accessible online.

167 **2.2 Model Configurations**

168 The modeling system consists of the WRF meteorological model (Skamarock et al., 2008), the
169 Sparse Matrix Operator Kernel Emissions (SMOKE) model for emissions modeling (Houyoux et
170 al., 2000) and the CMAQ model (Byun and Schere, 2006) for chemical transport modeling. The
171 details about model configurations are presented in the following sections. Two sets of
172 simulations were conducted, one set with obs-nudging and OA and the other without. The base
173 case, referred as "No-OA", did not employ obs-nudging or OA. The second case, "1Hr-OA",
174 performed obs-nudging and OA using hourly nudging input.

175 **2.2.1.   WRF Setup**

176 Both WRF simulations used the same nested domain and NARR (North American Regional
177 Reanalysis) as input, with grid nudging turned on. Figure 2 depicts the horizontal domain setup.
178 Two nested domains were used, with 12-km and 4-km resolution respectively. The 4-km domain
179 covered SETX and a small portion of Louisiana. The 12-km domain (red box) encompassed
180 Texas and parts of a few neighboring states. The number of grid cells for the 12-km and 4-km
181 domains were 161×145 (E-W by N-S), and 95×77 respectively. The projection type was Lambert
182 conic conformal (LCC). Three projection parameters, first latitude (33 °N), second latitude
183 (45 °N) and standard longitude (97 °W) are conformal to the projection parameters used by the
184 USEPA in developing emission inventories for air quality modeling (Mason et al. 2010). Both
185 domains had a vertical resolution of 27 eta layers based on dry hydrostatic pressures. The model
186 top is 100 hPa, corresponding to top layer pressure of the input NARR data.
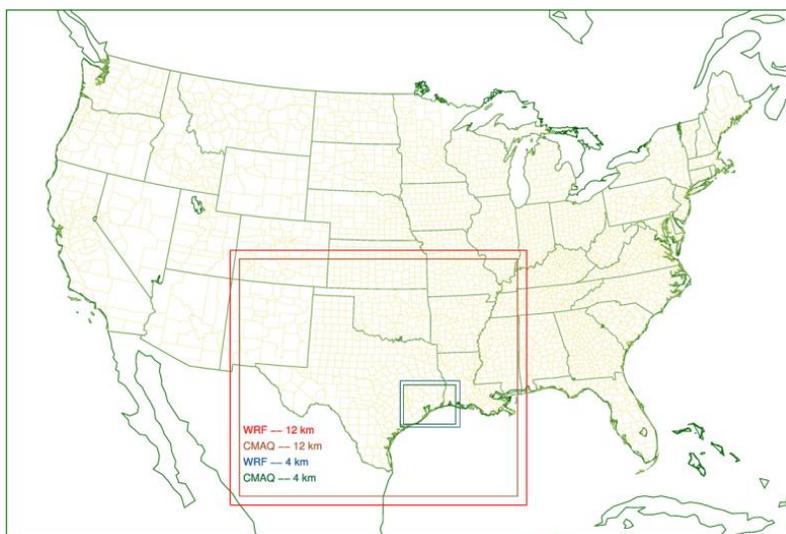
187

188   **Figure 2**. Horizontal domains of WRF and CMAQ simulation at 4 km and 12 km grid resolution

189   (the bigger domains are for 12 km WRF and CMAQ and the smaller domains for 4 km WRF and

190   CMAQ).

191   **2.2.1.1. Input Data**

192   The NARR data used for WRF simulations are downloadable from

193   http://rda.ucar.edu/datasets/ds608.0/. The data were based on an Eta 221 grid at 29 pressure

194   levels. Its horizontal resolution was 32-km and the frequency was 3 hourly. The initial and

195   boundary conditions were generated from the NARR analysis by WRF Pre-Processing System

196   (WPS). An alternative to NARR was the Eta-NAM analysis data. However, the data temporal

197   frequency was lowered from 3-hourly to 6-hourly starting in 2013. Our tests showed that it was

198   not as good as the NARR dataset - likely because of lower temporal resolution.

199   **2.2.1.2. Physics and FDDA Options**

200   Major physics options used in the model are listed in Table 1. Our past modeling experiences

201   indicated that employing the Yonsei University (YSU) for the PBL scheme and the Kain-Fritsch

202   (K-F) for the cumulous scheme gave the best results for the Houston area. The YSU scheme was

203   also one of the two PBL schemes recommended by Cuchiara et al. (2014). The K-F scheme is

204   "drier" than others and produces less number of "false" convectional thunderstorms. For grid

205   nudging options, we generally followed the recommendations in the WRF User Guide. For

206 example, the mass fields (temperature and moisture) were nudged only at layers above the PBL
207 while wind fields were adjusted at all layers including the surface layer.

208 **2.2.1.3.   Observation Nudging with MADIS and CAMS data in WRF**

209 Additional observational data are required to implement obs-nudging and OA. To generate the
210 input files for the OBSGRID program, we processed the observation data using the approach of
211 Ngan et al. (2012) and Czader et al. (2013). Observational data came from the MADIS and
212 TCEQ CAMS. MADIS is a National Oceanic and Atmospheric Administration (NOAA)
213 program which collects, integrates, quality-controls and distributes observations from NOAA
214 and other organizations. Additional information is available online, https://madis.ncep.noaa.gov/.
215 The four MADIS datasets used for obs-nudging were NOAA Profiler Network (NPN),
216 Cooperative Agency Profilers (CAP), Meteorological Terminal Aviation Routine (METAR)
217 weather report and NOAA Radiosonde (RAOB). The METAR dataset was collected by mostly
218 first-order, METAR reporting, surface monitoring stations. NPN, RAOB and CAP were the most
219 commonly used upper air datasets.

220 The processed input observation data were fed into OBSGRID to update the domain analyses
221 and generate additional surface analyses and text nudging files. Obs-nudging was performed by
222 the main WRF program after obs-nudging namelist variables are properly set. The namelist for
223 OBSGRID and relevant WRF section settings came largely from recommended values of the
224 WRF User Guide and a previous study by Ngan et al. (2012).

225 Theoretically, obs-nudging updating at a higher frequency should enhance the model's
226 performance. A typical frequency of input analysis data is 3-hourly while the frequency for
227 observational data is hourly. The 3-hourly frequency of input analyses may be the reason for the
228 default 3-hour time interval in WRF's OBSGRID settings for generating the obs-nudging files.
229 Since there were few existing obs-nudging studies related to air quality and we were not aware of
230 any reference to the adoption of a 1-hour input frequency, we assumed that all previous studies
231 used the default 3-hour interval. As the WRF model allows the interval to be set to 1-hour or
232 smaller when corresponding observational data were available, we tested both 1-hour and 3-hour
233 scenarios. The results indicated that 1-hour obs-nudging had slightly better performance than the
234 3-hour one. As a result, this study adopted 1-hour temporal frequency for observation nudging.

235   The variables that were nudged were temperature, moisture, and the two wind components

236   (zonal U-WIND and meridional V-WIND). Obs-nudging for moisture was not performed in this

237   study. This was based on our past experiences since performing moisture nudging sometimes

238   triggers excessive artificial thunderstorms which disrupted the model flow fields.

239   **2.2.2. Emissions Processing**

240   For anthropogenic sources we utilized the USEPA's National Emissions Inventory of 2008

241   (NEI2008). Motor vehicle emissions for this inventory were processed using the USEPA's

242   Motor Vehicle Emissions Simulator (MOVES) model (USEPA, 2015). The inventory was

243   processed using the Sparse Matrix Operator Kernel Emissions (SMOKE) model v3.1 (Houyoux

244   et al., 2000) to obtain gridded emission rates and speciated for the Carbon Bond 05 (CB05)

245   chemical mechanism for use in the CMAQ model. The biogenic emissions were modeled using

246   the Biogenic Emissions Inventory System (BEIS) v3.14. Although NEI2008 might have

247   overestimated NO_x emissions in Houston (e.g., Choi 2012; Czader et al. 2015), we used base

248   NEI2008 without adjustment because the adjustment of the NO_x emission also has large

249   uncertainty (Czader et al. 2015). Pan et al. (2015) showed that the CMAQ ozone performance

250   using NEI2008 appears reasonable.

251   **2.2.3. CMAQ Configurations**

252   The USEPA's CMAQ (Byun and Schere 2006) version 5.0.1 was adopted for this study. Several

253   air quality studies focusing on the Houston area have used this model (e.g., Foley et al. 2010;

254   Czader et al. 2013, 2015; Choi 2014; Pan et al. 2015; Diao et al. 2016; Souri et al. 2016). The

255   model's horizontal domains were slightly smaller than its WRF counterpart in order to avoid the

256   discontinuity near the domain boundary. The domains are depicted in Figure 2 as green and

257   brown boxes, respectively. The chemical boundary conditions for all the species in the 4-km

258   domain were derived from 12-km domain air quality forecasting results, available online at

259   http://spock.geosc.uh.edu. The model used the same vertical structure as WRF. Additional model

260   configurations are listed in Table 2. Chemical processes were simulated with the CB05 chemical

261   mechanism with cloud/aqueous chemistry, active chlorine chemistry and an updated toluene

262   mechanism. For aerosol modeling, the fifth-generation CMAQ aerosol mechanism (AERO5)

263 which includes sea salt modeling was selected. The total number of included species is 132, with

264 70 reactive gas-phase, 49 aerosol and 13 non-reactive species.

3. Evaluation Metrics

266 To assess model performance against observations, we computed a set of five statistics including

267 Pearson correlation, index of agreement (IOA, Willmott 1981), mean bias (MB), root mean

268 square error (RMSE), and Mean Absolute Error (MAE). This list is similar to one used by Li et

269 al (2008) for model performance evaluation. The goal is to have a comprehensive comparison

270 between model and observation time series. The set of five statistics was divided into three

271 groups:

272 1) Evaluation of the magnitude of model results vis-a-vis in-situ data, in measurement units

273 • Mean Bias (MB)

274 • Mean Absolute Error (MAE)

275 • Root Mean Square Error (RMSE)

276 2) Measuring how close the model values follow changes in the observations, unitless

277 • Correlation

278 3) A composite performance index, Index of Agreement (IOA or d), unitless

279 IOA is considered a better performance index than correlation as it takes into account the

280 difference in the means and standard deviation. For example, when correlations are similar,

281 lower model biases would yield higher IOA values. Additionally for this study, the mean and the

282 standard deviation of model values and observations were included as a reference.

4. General Meteorological and Ozone Conditions in September 2013

284 The weather during September 2013 was relatively dry with mostly southerly, easterly or

285 southeasterly winds. For this study, date format is mm/dd unless stated otherwise. From 09/05 to

286 09/19, there was a lack of influence of strong synoptic weather systems. Shifting wind patterns

287 were observed during the period: light northeasterly in the early morning gradually turned

288 clockwise to southeasterly in the afternoon and evening hours. In this period, winds shifted from

289 southeast to near east and there were more clouds after 09/10. The only cold front arrived on the

290 early morning of 09/21. Figure 3 shows the regional average temperatures for the period and it

291 can be seen that 09/21 had the lowest daily maximum temperature. The influences of the cold air

292 intrusion lasted till early 09/21. Winds turned into southerly in the afternoon of 09/25 and

293 warming continued in the next few days until 09/28.
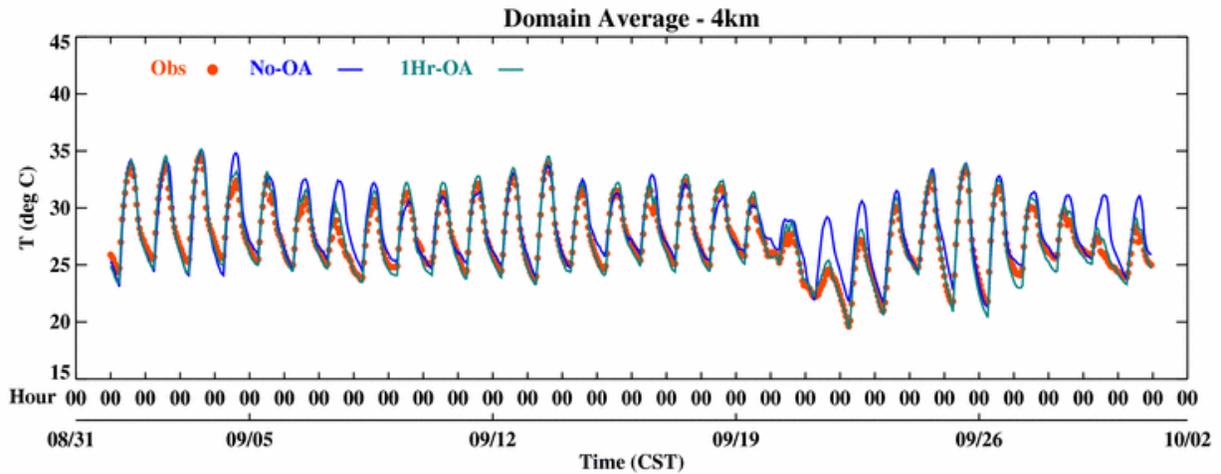
294

295 **Figure 3**. Regional hourly temperature averaged over all available hourly CAMS observations,

296 two model cases also included for September of 2013.

297 Light rain events occurred on 09/02, 09/10, 09/16, 09/19 to 09/21 and 09/28 to 09/30. 09/20 and

298 09/21 events consisted of widespread light to medium showers. Besides the above-mentioned

299 dates, there were a few other days with sporadic drizzles. A majority of the days between 09/01

300 and 09/20 were either sunny or cloudy. The periods from 09/08 to 09/10 and 09/18 to 09/20 had

301 more clouds than other days. The period from 09/21 to 09/30 was influenced by a cold front

302 passage. The days between 09/22 and 09/24 were sunny and cold. The surface wind reversed

303 direction during midday of 09/25 and brought clouds back from 09/26 to 09/30. High ozone

304 events in SETX during fall are typically associated with the passage of a cold front (e.g.,

305 Rappenglueck et al. 2008). The only ozone event with hourly surface ozone exceeding 120 ppb

306 (parts per billion) in September which occurred on 09/25 fell in this category.

307 Figure 4 plots the hourly regional averaged ozone. On most days, the in-situ averaged ozone

308 concentrations were below 70 ppb. Since the winds after dawn consistently pushed the

309 precursors from the industrial area to the southwest of the city, the wind pattern did not favor the

310 local ozone production. The daytime winds also contained a persistent easterly component which

311 moved the pollutants away from the Houston metropolitan area. In the first 10-day period, low

312 background ozone originating from the Gulf of Mexico contributed to the low-ozone days. With

313 overcast skies on 09/19 and 09/20, hourly high ozone values dipped below 30 ppb. The two

314 highest ozone days characterized by post-frontal ozone events were the 25th and the 26th of

315 September respectively.
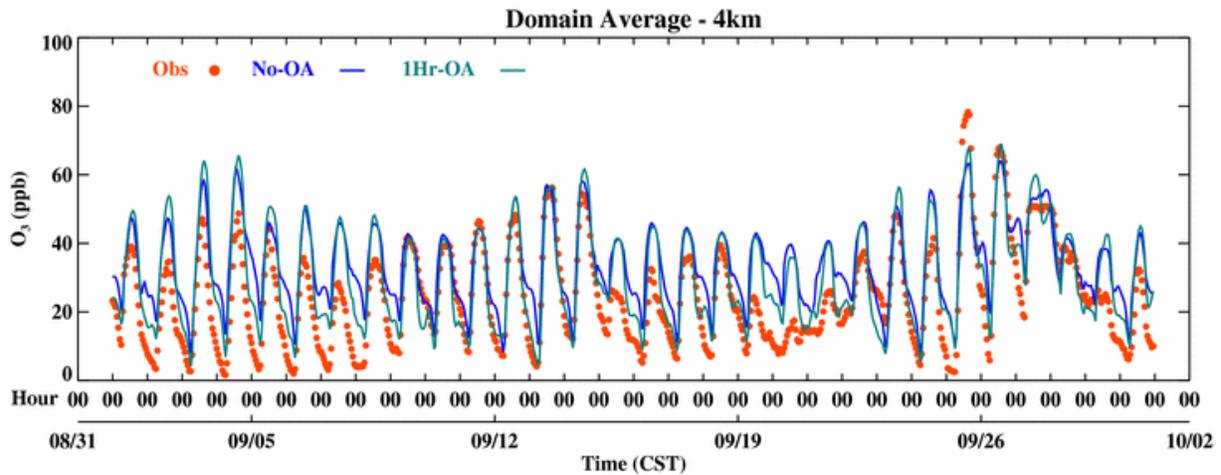
316

317 **Figure 4**. The hourly regional averaged ozone for the two cases (No-OA and 1hr-OA) at the

318 stations which include observation surface O_3 over the 4km domain for September of 2013.

5. Results

320 To evaluate the WRF simulation, we calculated statistics for surface temperature and winds in

321 the 4-km domain. In addition, we plotted out the PBL height time-series for the one site we had

322 observations. For CMAQ evaluation, we calculated the surface ozone statistics for the whole

323 month. Also, we plotted vertical ozone profile and calculated biases for ozone aloft on 09/25.

5.1. Meteorology

5.1.1. Temperature

326 The comparison of regional average hourly temperature for the simulation period is shown in

327 Figure 3. The regional observed averaged surface temperature was calculated by averaging the

328 hourly temperature from ~60 CAMS sites in the 4-km model domain. The base case temperature

329   was too high compared to the in-situ measurements. For example, the "No-OA" base case

330   maximum temperature for the 21st was 30°C compared to 25°C for the in-situ data. The high

331   biases in the base case are sharply reduced in the "1Hr-OA" sensitivity case and temperature

332   matched better with the observations for several time periods, especially for 09/20 to 09/23. The

333   statistics of hourly surface temperature are listed in Table 3. With higher IOA and lower mean

334   biases (MB), the "1Hr-OA" case was clearly better than the base case "No-OA". The IOA of

335   "1Hr-OA" was about 9% higher than the base case.

5.1.2. Winds

337   In ozone chemistry, winds affect the accumulation of precursors and hence the resulting ozone

338   production (e.g., Banta et al. 2005, 2011; Darby 2005). They are also responsible for dispersing

339   high ozone and bringing in background ozone. Prevailing summer time southerly to

340   southeasterly winds in the HGB region significantly lower the ozone concentrations in the

341   metropolitan area. Therefore, high ozone events usually occur when such wind patterns change.

342   Cold front intrusions coming as early as late August blow pollutants to the south. As a result, an

343   area of high ozone develops in the Gulf. Following cold fronts weakening and the weather

344   warming up, reversing winds can bring high ozone back to land. High ozone may also occur

345   during intra-day recirculation events when pollutants previously blown away from industrial

346   zone are brought back by reversing winds. Correctly simulating these recirculation events is

347   particularly important in predicting the high ozone event caused by post-front conditions. The

348   ozone event in the HSC area on 09/25 was likely due to a combination of local recirculation

349   caused by onset of the bay breeze and increased background ozone brought in by transport.

350   Due to the land-water thermal contrast and the different size of the Galveston Bay and the Gulf

351   of Mexico, the western shore of the Galveston Bay often experiences a successive onset of bay

352   breeze and sea breeze in the summer. The bay breeze is typically a weaker easterly while the sea

353   breeze is a stronger southeasterly. The sea breeze usually arrives one to a few hours later after

354   the bay breeze. The bay breeze and the subsequent sea breeze phenomena in Houston were

355   described by Banta et al. (2005).

356   The statistics of zonal (U-WIND) and meridional (V-WIND) wind components are listed in

357   Table 3. The purpose of choosing U and V over wind speed and direction is to avoid the

358   anomalies in the wind direction statistics. For example, although wind direction of 5 and 355
359   degrees are close, the statistics suggest that they are distinctively different.

360   For both U and V components of wind, "1Hr-OA" had higher correlation and IOA than "No-
361   OA". The model performance on U and V are similar, with the correlation in a range of 0.76 to
362   0.81 for all the cases. For comparison, the performance of the OA case ("M1") in Ngan et al.
363   (2012) is very close to that in this study, with a correlation of 0.75 for U and 0.82 for V. In terms
364   of IOA, the OA case is higher by 5-6% in U and 10-11% in V over the base case. This can be
365   explained by the much reduced wind biases in the OA case.

366   The base case had consistently stronger winds, especially the southerly component, than the
367   observations. This was reflected in the mean bias "MB", as well as the model mean "M_M".
368   Winds were reduced significantly after OA was performed. In fact, the high southerly bias in
369   "No-OA" turned slightly negative after OA. Winds originating from the Gulf were also stronger
370   in the base case, which played a role in raising the ozone level comparing to the "1hr-OA"
371   sensitivity case. Figure 5 illustrated the slowing down of southerly winds after observation
372   nudging. As a result of nudging, the wind vectors matched the observations better.
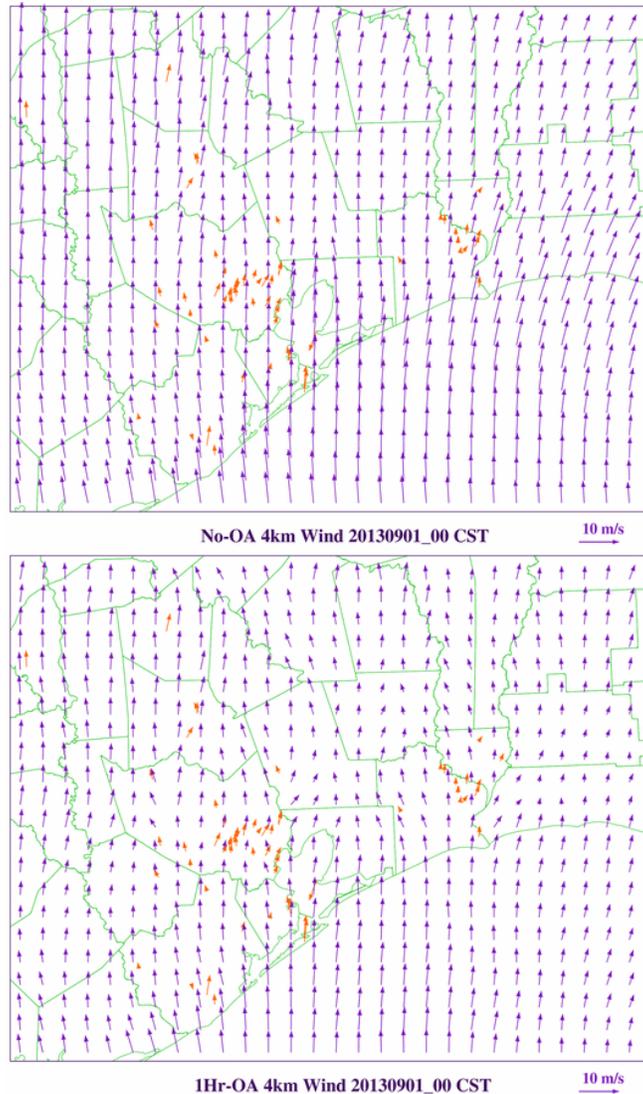
Figure 5. Model and observed winds at 09/01_00 CST: No-OA (top) and 1Hr-OA (bottom). Model winds are blue arrows and the observations are orange arrows. Stronger southerly winds, especially along coastal region, were reduced in the OA case.

5.1.3. PBL height

Atmospheric pollutants are largely confined in the PBL as most of the emissions sources are close to the ground level. Hence the PBL height plays a critical role in mixing and spreading the pollutants. Haman et al. (2014) studied the relationship between ozone level and PBL height at a Houston CAMS site and found that nighttime and early morning PBL heights were consistently lower on high ozone days than on low ozone days. Czader et al. (2013) pointed out that the

model underprediction of PBL during nighttime may have caused the CO overprediction at the same site. CO is a good proxy for understanding model transport since it has low reactivity and a relatively long lifetime in the troposphere (typically 2 months). Cuchiara el al. (2014) conducted four WRF/Chem sensitivity tests using different PBL schemes over southeast Texas. While no preferred PBL scheme was identified for WRF simulations, the YSU scheme performed better than others in terms of ozone prediction.

Haman et al. (2012) showed that the daily maximum PBL height at the University of Houston site mentioned in section 2.1 reached its highest values of slightly over 2000 m in August. In September, typical daily maximum PBL height was 1500 m at 15 CST while daily minimum was just below 200 m between 00 CST and 06 CST. The comparison of observed and model PBL height is shown at Figure 6. Our results indicated that the model tended to overpredict the daily maximum PBL height; obs-nudging helped to reduce the overprediction. For the daily minimum PBL height, "No-OA" case had slightly high biases while the OA case matched quite well with in-situ height data. The observed minimum PBL height was lower than that reported by Haman et al. (2012), likely due to the cloudy conditions prevailing in September 2013. There was no apparent explanation on the reduced daytime PBL biases in the OA case than the base case, but it is likely the results of improved winds and temperatures in PBL.
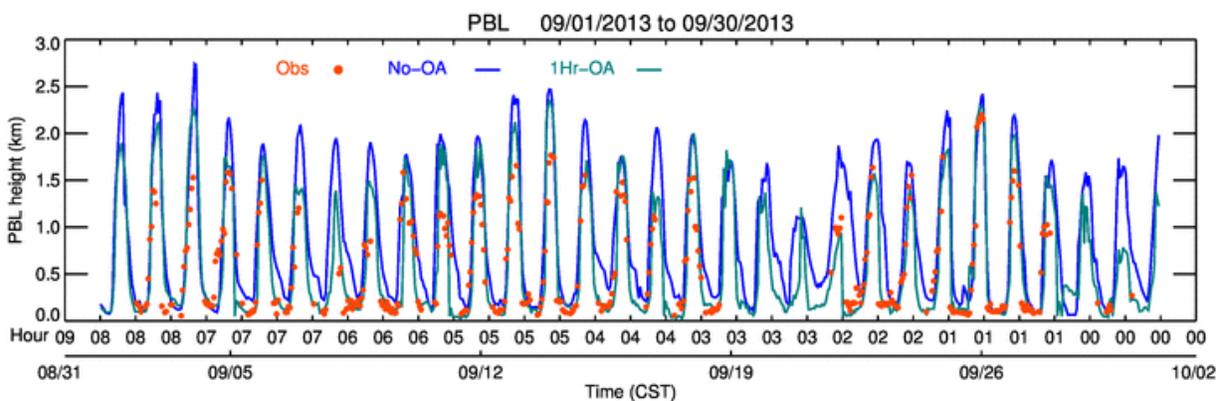

Figure 6. Planetary Boundary Layer (PBL) height time series at CAMS C695 for September 2013.

5.1.4. Cold Front Passage

404   The surface winds on 09/20 were overwhelmingly southerly in the region and reversed on 09/21

405   due to the arrival of a cold front. The hour-by-hour wind shifts for 11 sites in HGB on 21

406   September are plotted in Figure 7. The sites are sorted by latitude with the southernmost site,

407   Galveston C1034, located at the bottom row. There was only one site, Deer Park C35, showing

408   weak southerly at 00 CST while all the others had mostly weak northerly. Starting from 01 CST,

409   winds in the entire HGB area turned northerly to northeasterly and continued gaining strength in

410   the next few hours, indicating cold air had taken over the region.

411   Both cases performed reasonably well on 09/21 and the timing of wind shift was captured

412   accurately; although "No-OA" lagged behind by ~ 1 hr. The winds turned weak northerly at 00

413   CST for most sites but the "No-OA" case still showed the wind direction to be all southerly.

414   Besides the timing, the northeasterly winds in the "No-OA" case sometimes were too strong;

415   obs-nudging helped moderate the winds. The reduced V-wind bias in "1Hr-OA" was also evident

416   in the wind model-measurement statistics on 09/21.

20130921: Houston Wind - No-OA
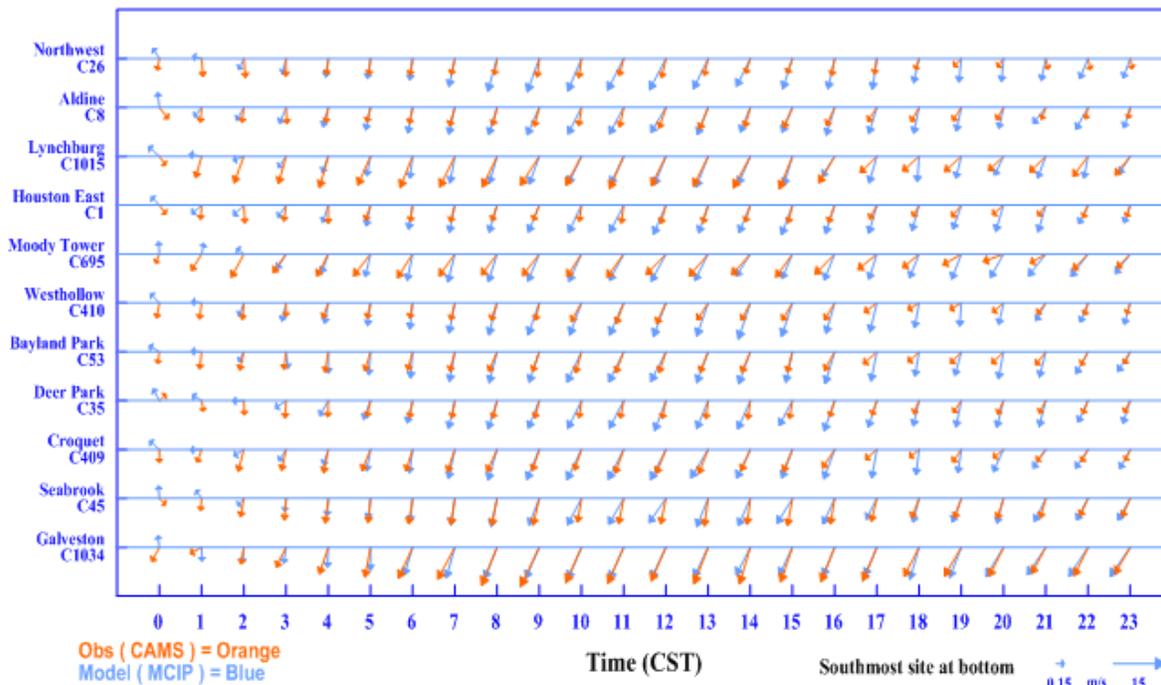
20130921: Houston Wind - 1Hr-OA
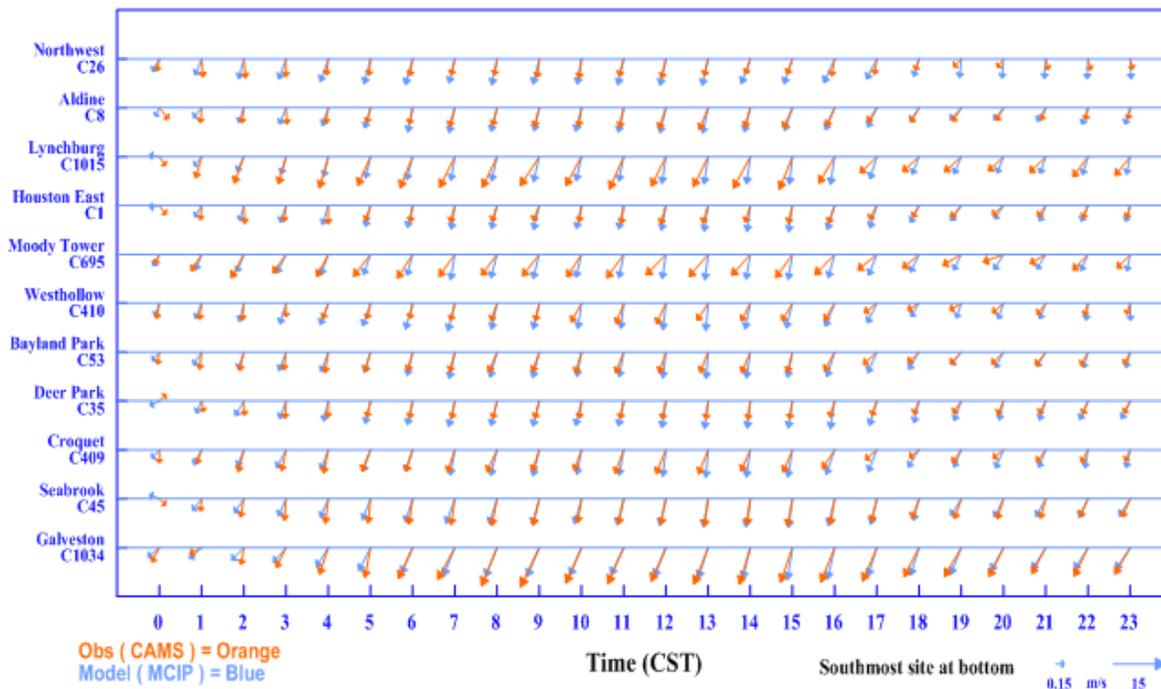
417

418 **Figure 7**. Hourly model (blue) and CAMS (orange) winds at 11 sites on 09/21: No-OA (top) and

419 1hr-OA (bottom). The 1hr-OA case is better at 00 CST to 02 CST and 17 CST to 20 CST.

420 **5.2.** **Ozone**

5.2.1. Regional Average Hourly Ozone

Figure 4 plots the regional average hourly ozone, which was defined similarly to the average temperature. Overall, observed ozone concentrations were low and the model did a reasonably good job on capturing the timing of intra-day variations. However, both cases tended to overpredict the daily highs and daily lows, especially in the first 8 days and between 09/15 and 09/21. An obvious departure is the 25th – both cases underpredicted the daily high. During the model high bias period, the OA case usually did better in reaching the daily low although it overpredicted the high a bit more than the base case. The night time biases were reduced likely because the lower southerly winds in the OA case transported less ozone from the Gulf to the land.

Our results suggested that the modeled ozone concentrations were likely higher in the Gulf than those in the real world. During the periods 09/02 – 09/04 and 09/07 – 09/08, the incoming ozone from the Gulf was markedly lower than "normal" values. Since the model ozone had fixed boundary values, the model was unable to capture these daily ozone variations at the boundary. The model showed the highest biases during the period of 09/19 – 09/20, which is likely due to overcast skies and uncertainties in models cloud fields and high background ozone values. Despite the overprediction, the biases in OA case are notably lower than the base "No-OA" case during the nights of 19th and 20th. A future study to upgrade the accuracy of cloud fraction using remote sensing data (e.g., Moderate Resolution Imaging Spectroradiometer, or MODIS) should be helpful in explaining the biases.

There were a few days with elevated ozone due to post-front meteorology conditions. The only exceedance happened on 09/25, which was likely caused by meteorological events in Houston and the Galveston Bay. The overall ozone on 09/26 was higher than normal after southerly winds transported the ozone back from the Gulf, raising ozone concentrations over the entire region. A more detailed analysis of model predictions on 09/25 will be presented in subsection of 5.2.3.

5.2.2. Performance Statistics

The ozone statistics are listed in Table 4. Both cases had very similar correlation of 0.72 and 0.73. However, the mean biases in the OA case were lower than the base case by 3.2 ppb, which helped raise the IOA from 0.78 to 0.83. The model standard deviation increased in the OA case

450   and matched better with that of the in-situ data. The IOA improvement of the "1Hr-OA" case

451   over the base case for ozone was slightly less as compared to that for temperature and winds.

5.2.3.   High ozone episode after the passage of a front

453   In SETX, high ozone events during the fall season usually occurred after the passage of a cold

454   front (e.g., Rappengl ück et al. 2008; Ngan and Byun 2011; Ngan et al. 2012; Haman et al. 2014).

455   Two factors may have contributed to the post-front ozone events: 1) Following a cold spell, light

456   winds and sunny skies create an ideal condition for ozone production and accumulation. 2) Wind

457   reversal may transport back the pollutants that were previously blown into the Gulf.

458   During the DISCOVER-AQ period, the two days with highest ozone concentrations were the

459   25th and 26th of September as indicated in Figure 4. The 1-hour maximum ozone on the 25th was

460   localized and higher by about 40 ppb than the 26th. In addition to heightened background ozone

461   on the 25th, the major contributor was the production resulting from favorable weather

462   conditions: sunny, overall light winds and shifting winds over the industrial area. The light

463   morning land breeze carried pollutants from the ship channel area to the Galveston Bay. As the

464   day warmed up, the bay breeze started to develop and carry pollutants back to the land. This

465   localized circulation was described by Banta et al. (2005). Ngan et al. (2012) reported the same

466   phenomenon in their Texas Air Quality Study-II 2006 study.

467   Figure 8 shows the ozone time series for the La Porte (C556 in Figure 1) site located in the HSC

468   area. The highest hourly ozone for September, 151 ppb occurred here at 13 CST on the 25th.

469   Ozone rose from 10 ppb to 150 ppb between 09-12 CST. Such a sharp increase in ozone was

470   likely the result of increased photochemical activity under favorable meteorological conditions in

471   an area with accumulated precursors. Figures 9 and 10 depict the wind and ozone concentrations

472   at 08 CST and 13 CST.

473   The wind plots of Figure 9 indicate that the observed winds in the HGB region at 08 CST were

474   light northerly for sites located on the north side, while they were westerly in the middle and

475   south. The base case winds were all northerly while the OA case had northwest winds for the

476   north side and west winds for the middle and south. Hence, the model winds in OA case are

477   more realistic than the winds in base case. The 09 CST winds were similar to those of 08 CST.

478   As a result, the ozone statistics in Table 5 showed that the OA case had much better correlation

479 and IOA than the base case during 08-09 CST. This example demonstrated the ability of obs-
480 nudging to correct erroneous winds. However, later events showed it may not always be able to
481 perform consistently.

482 The bay breeze started to develop at 10 CST near the C556 site. The early onset was likely to be
483 related to warming up on the previous afternoon on 09/24 as indicated in Figure 3. At 10 CST
484 most other sites to the west of HSC experienced light northwest winds while those at HSC
485 originated from the northeast. Combined with the easterly bay breeze, a convergence zone was
486 formed just below C556, where emissions from the HSC area stalled and accumulated. At 13
487 CST, the whole region had light winds and the bay breeze was well developed. The highest
488 ozone indeed appeared in C556 and its vicinity. The rapid increase of ozone concentration for
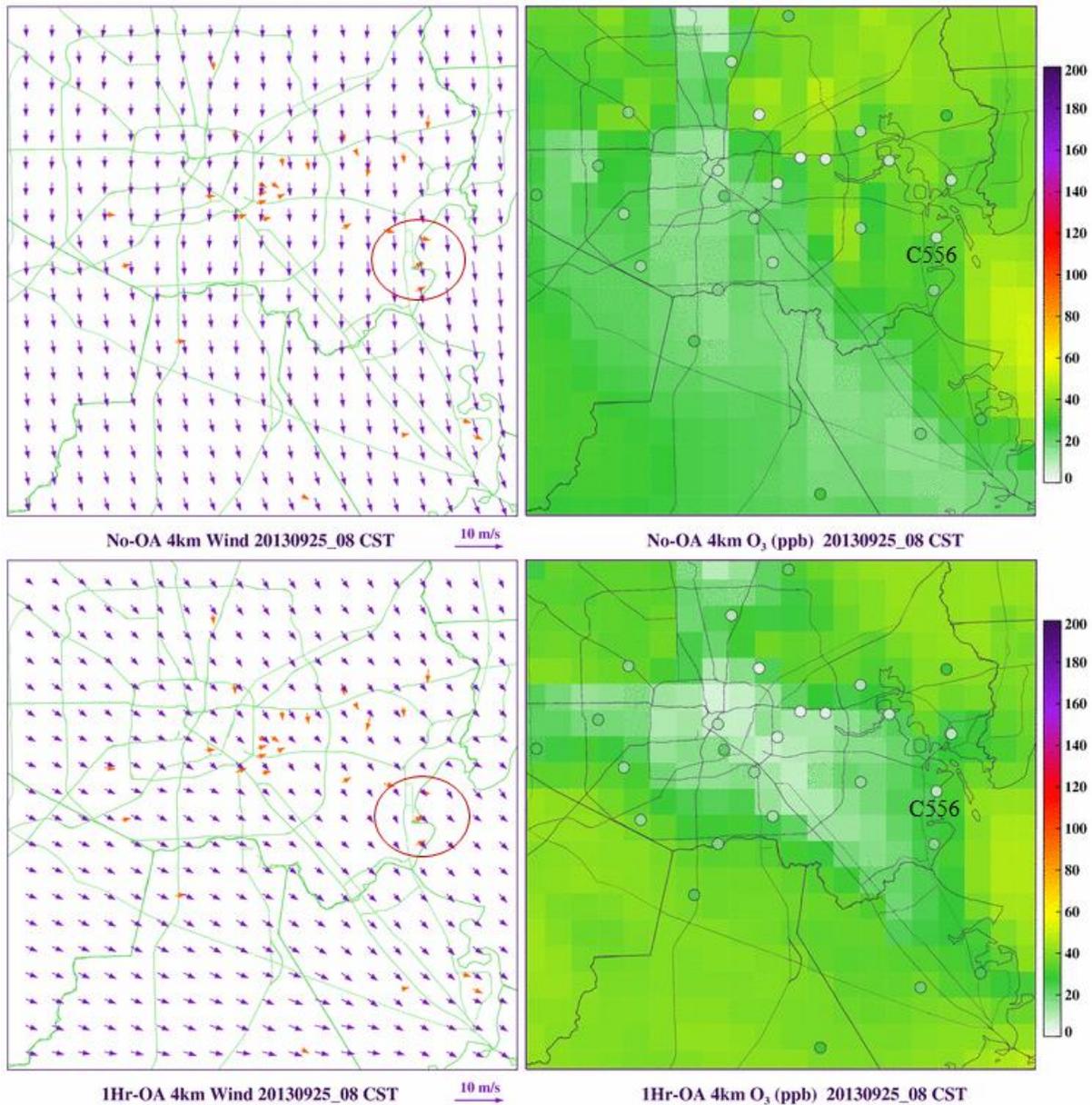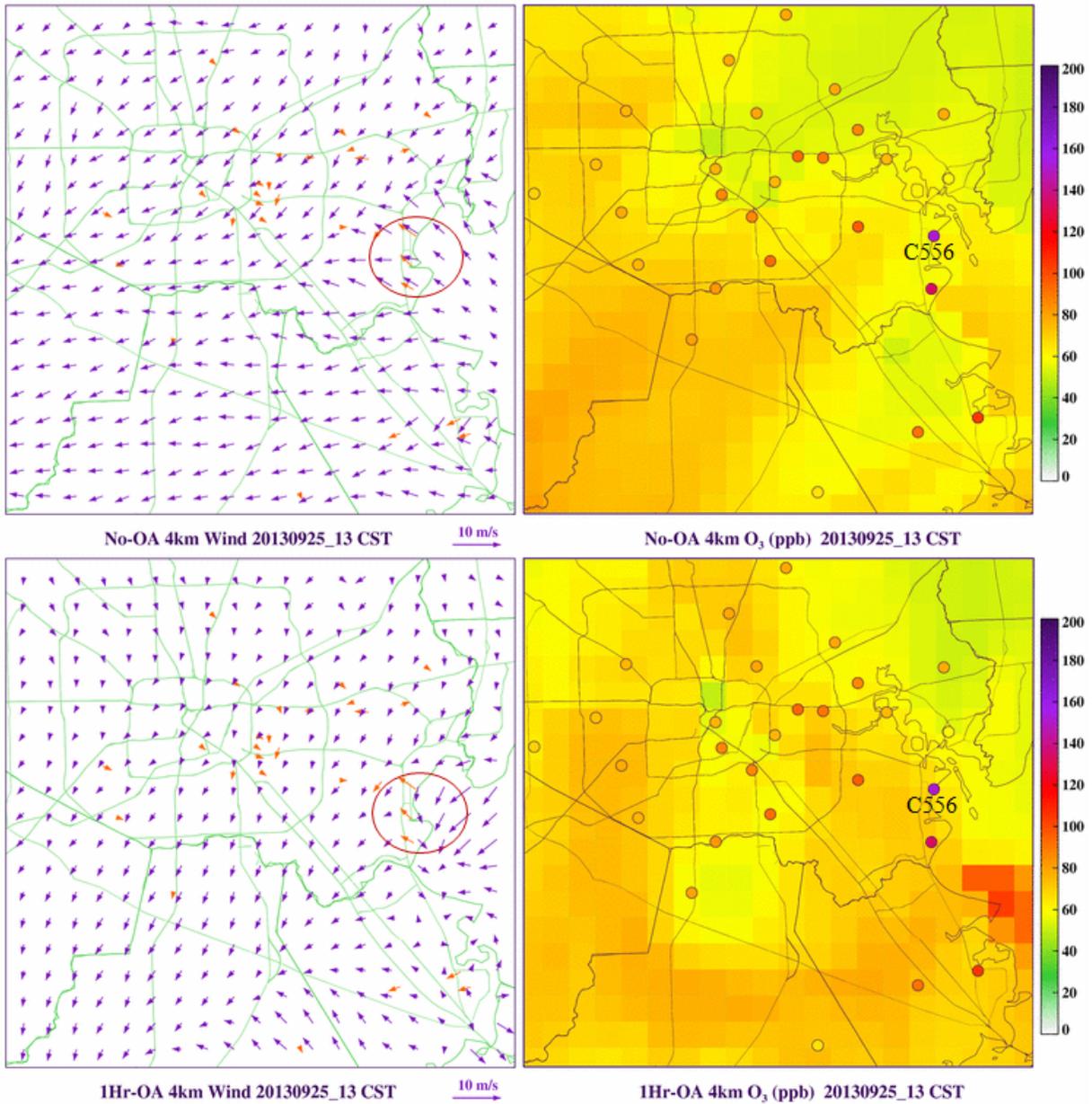489 C556 between 09-13 CST is shown in Figure 8.
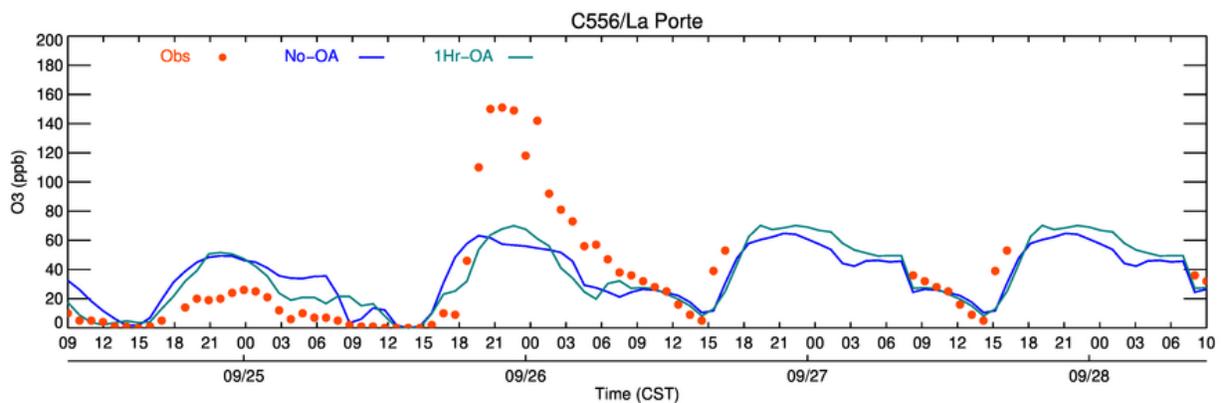
490

491 **Figure 8**. Ozone time series of La Porte (C556) between 09/24_00 to 09/28_00 CST of 2013.

492 It is important to note that both modeled cases missed the wind shifts in the HSC area, and the
493 resulting convergence zone near C556. This could explain the model's inability to recreate the
494 sharp ozone increase at C556. Figure 9 shows that the ozone concentrations around HSC area are
495 quite low (~10 ppb) at 08 CST. A further examination showed that while both model cases
496 missed the wind shift and convergence, the patterns were different. The base case winds differed
497 considerably from the observed ones for most of the morning: instead of a weak westerly, it had
498 stronger northwesterly to northerly. By 08 CST, winds were almost uniformly northerly in the
499 base case while they were weak west-northwesterly in the OA case (Figure 9). The oval in
500 Figure 9's top-left panel shows the mismatch of winds around C556 in the base case. As a result,

501   the NO$_x$ produced in the city was carried further to the southeast in the model in the base case.

502   Until 13 CST, base case winds did not shift directions by much. The OA case got the early hour

503   weak northwesterly right, but missed the bay breeze onset between 10 and 13 CST (oval in

504   Figure 10). The OA case could not reproduce the small-scale wind reversal near C556,

505   suggesting there is a limitation in the current WRF OA's capability in rectifying the winds when

506   poorly simulated. On the other hand, the OA case did improve the spatial ozone pattern, as the

507   high ozone area was closer to HSC after OA (Figure 10).

No-OA 4km Wind 20130925_08 CST   10 m/s   No-OA 4km O$_3$ (ppb) 20130925_08 CST

1Hr-OA 4km Wind 20130925_08 CST   10 m/s   1Hr-OA 4km O$_3$ (ppb) 20130925_08 CST

508

509   **Figure 9**. Zoom-in ozone concentrations (right) and wind plots (left) at 13 CST 25 September for

510   "No-OA" (top) and "1Hr-OA" (bottom). Ozone observation is in the small circle; wind

511   observation is indicated by an orange arrow. La Porte site C556 is labeled. The numerical range

512   of the right-side color scale is 0 to 200 ppb. Higher values than 200 ppb have the same color as

513   200 ppb.

No-OA 4km Wind 20130925_13 CST   10 m/s   No-OA 4km O₃ (ppb) 20130925_13 CST

1Hr-OA 4km Wind 20130925_13 CST   10 m/s   1Hr-OA 4km O₃ (ppb) 20130925_13 CST

514
515

516   **Figure 10**. Zoom-in ozone concentrations (right) and wind plots (left) at 13 CST 25 September

517   for "No-OA" (top) and "1Hr-OA" (bottom). Ozone observation is in the small circle; wind

518   observation is indicated by an orange arrow. Bay breeze is shown in the orange oval. The

519   numerical range of the right-side color scale is 0 to 200 ppb. Higher values than 200 ppb have

520   the same color as 200 ppb.

521   The ozone measurements from aircraft P3-B provided a more complete picture for the ozone

522   evolution on 09/25. During the day, the aircraft flew around the industrial area, Galveston Bay

523   and Galveston Island for about 9 hours. Figures 11 and 12 plot the ozone concentrations along

524   aircraft tracks at 08 and 13 CST. Surface layer ozone from the "No-OA" case is provided as

525   background for reference. At 08 CST, an ozone level of 60-80 ppb aloft was already observed at

526   three locations (three loops in Fig.11): Galveston Island, Smith Point and the inner city. Another

527   high of ~90 ppb could be seen above the HSC area. Ozonesonde observations over HGB showed

528   the aloft ozone concentrations were typically ~40-60 ppb (e. g., Li and Rappengluck 2014) at the

529   height level of a few hundred meters to 4 km. The higher-than normal ozone aloft suggested a

530   post-front ozone recirculation condition. Such high ozone aloft might raise surface ozone as a

531   growing PBL downwardly mixed the air aloft with near surface air. At 13 CST, high ozone over

532   100 ppb was observed at multiple locations. The highest aloft ozone of ~ 160 ppb occurred

533   southwest of Smith Point in the Galveston Bay. Such high increase in ozone concentrations was

534   likely the result of active photochemistry in the industrial zone and around Galveston Bay;

535   indicating a high level of precursor accumulation in the area.

2013-09-25_08 CST   P3B Ozone with No-OA Surface

536
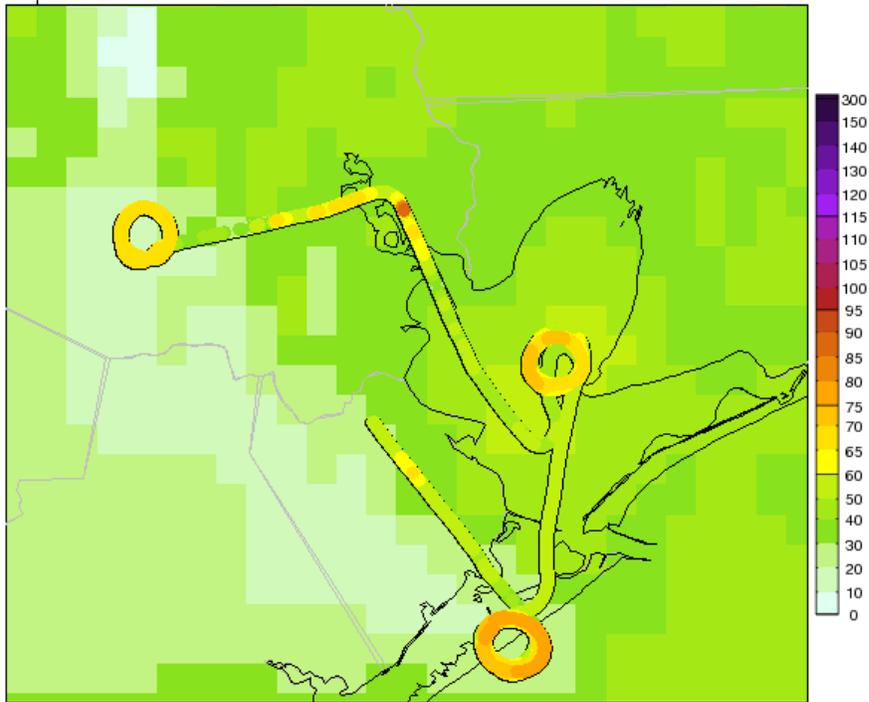
Figure 11. Ozone along aircraft tracks at 08 CST of September 25th, overlaid upon model No-OA surface ozone. The numerical range of the right-side color scale is 0 to 300 ppb. Higher values than 300 ppb have the same color as 300 ppb.

2013-09-25_13 CST   P3B Ozone with No-OA Surface

540
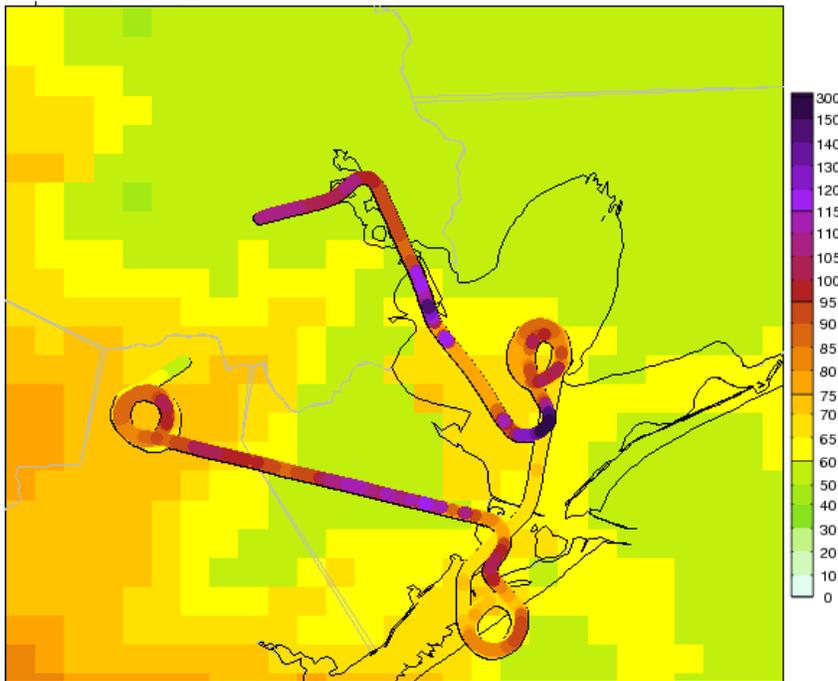
Figure 12. Ozone along aircraft tracks at 09/25_13 CST of September 25th, overlaid upon model
"No-OA" surface ozone. Plumes can be seen as dark purple circles in Galveston Bay. The
numerical range of the right-side color scale is 0 to 300 ppb. Higher values than 300 ppb have
the same color as 300 ppb.

Figure 13 shows hourly ozone vertical profiles from 08 CST to 16 CST on September 25th, with
ozone being displayed on the x-axis and height on the y-axis. The observed ozone was averaged
over multiple measurements in the same model cell, so that they could be compared to model
output. Next, both model and observed ozone values were averaged over all the grid cells in the
same model layer, such that one dot represents the average ozone of all the cells in the same
layer. The 08 and 09 CST profiles showed there was a high ozone layer with average ozone of
~65 ppb stretching from 450 m to 1200 m in height. In comparison, all model runs had lower
ozone in this layer. The model biases as shown in Figure 14 were about -10 ppb at 08 CST and
grew to -20 ppb at 09 CST. The large discrepancy between low surface ozone and ozone aloft
was unusual and may be explained by the arrival of a high ozone air mass from aloft. The
observed ozone rose continuously in the following hours; however simulated ozone stagnated
around 60 ppb from the surface up to 2000 m until 15 CST. At 16 CST, the ozone of the OA case

557   in the lowermost (0-1 km) layer rose 20 ppb over the previous hours yet the base case ozone

558   increased only a few ppb. Although different in magnitude, the aloft ozone had a few similar

559   features to the surface ozone. Firstly, the model missed the observed high ozone in the afternoon

560   by a large margin. For example, the base case underpredicted the 0-1 km level ozone by up to 50

561   ppb. The primary cause for the lower ozone production was likely the models wind fields as both

562   model and observations had a clear sky in the industrial area and the Galveston Bay. Secondly,

563   nudging clearly helped reducing the ozone biases aloft. In most plots of Figure 14, the OA case

564   had lower biases than the base case. The largest difference was at 16 CST when nudging reduced

565   biases from ~45 ppb to ~30 ppb in the 300 – 1000 m layer.

566   While it is easy to understand the improvements in temperature and winds after obs-nudging was

567   applied, it is more difficult to explain how other variables such as precipitation and clouds

568   reacted to obs-nudging. The indirect impact of these meteorological variables on ozone was

569   harder to assess. In our study, we did not evaluate clouds quantitatively as there were no

570   digitized cloud fraction data available for our modeling domains. A preliminary analysis on

571   convection showed that there were occasions in which the model missed the convection or

572   precipitation and there were other occasions in which the model created artificial convection.

573   The convection cells were usually visible as a "star-burst" from surface wind vector plots –

574   arrows going out to different directions from a center. However, the mismatch in convection

575   appeared to be not a serious issue since only a few occurrences were observed in the month of

576   September.

577
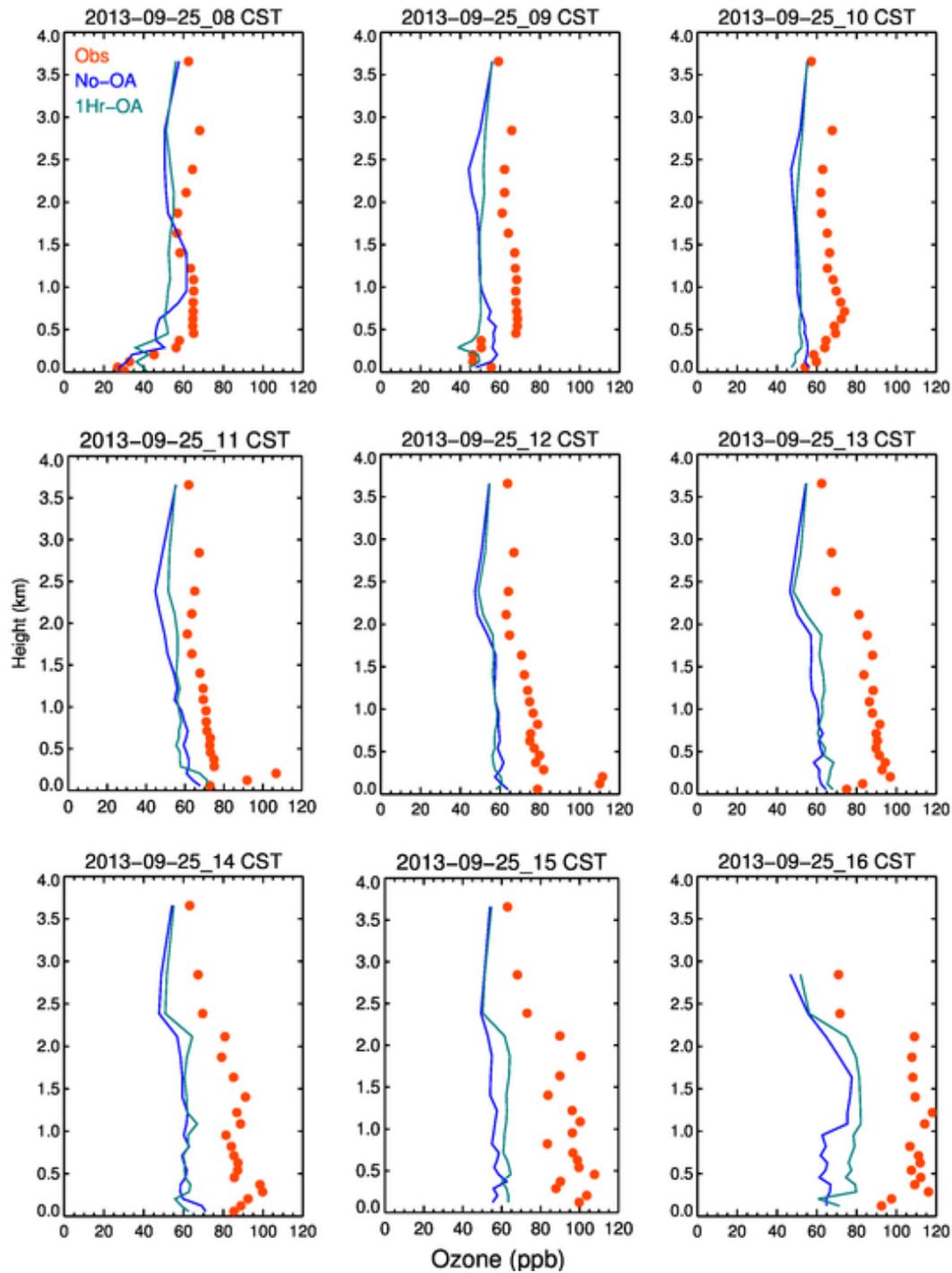
Figure 13. Vertical ozone profiles from 09/25_08 CST to 09/25_16 CST of 2013 for two cases of No-OA and 1Hr-OA compared with corresponding observations. There are 9 plots, each for one hour, with date and hour as caption.

581
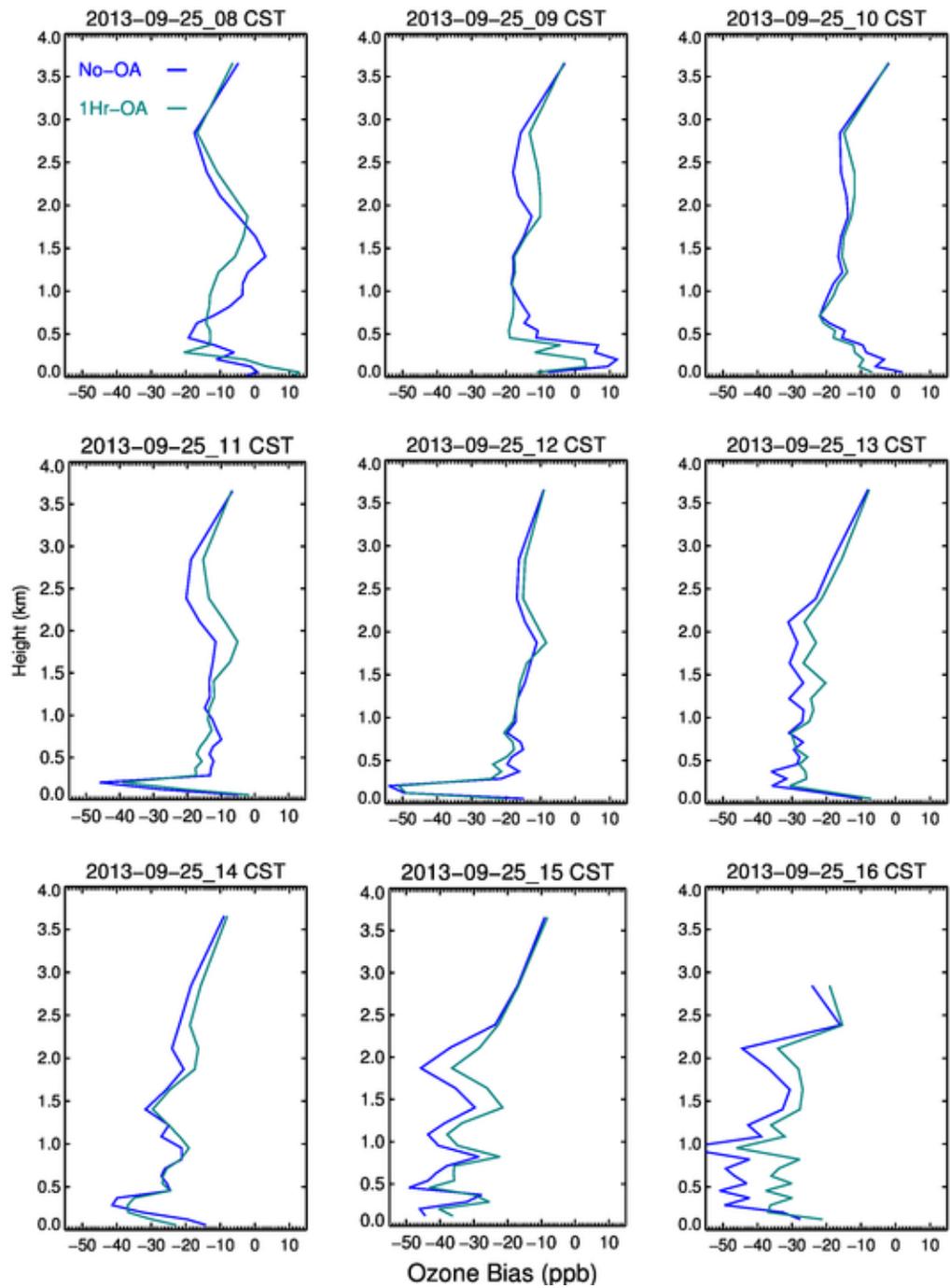
Figure 14. Model vertical ozone biases from 09/25_08 CST to 09/25_16 CST of 2013 for two

cases of No-OA and 1Hr-OA. There are 9 plots, each for one hour, with date and hour as caption.

6.   Conclusions and Discussions

585 In this study, we performed two Weather Research and Forecasting (WRF) and Community
586 Multiscale Air Quality (CMAQ) model simulations to explore model sensitivity to observation
587 nudging. In evaluating meteorological and ozone conditions, we found that obs-nudging
588 improved the meteorology and ozone performance as shown in the index of agreement (IOA) of
589 temperature, winds, and ozone. While the base case winds were overall well simulated, obs-
590 nudging significantly reduced the high wind biases (especially the meridional wind) shown in the
591 base case. For planetary boundary layer height, obs-nudging reduced high biases in both daily
592 maximum and daily minimum values. In the end, the combined changes in meteorology lowered
593 the ozone biases by about 3 ppb, a 35% reduction. There were short time periods (such as
594 between 07 and 09 CST on 09/25/2013) when the simulated base case model winds differed
595 significantly from observational data and obs-nudging significantly corrected the meteorological
596 simulation problems, leading to much better ozone simulation. However, model ozone biases are
597 also impacted by emissions and lateral boundary conditions.

598 The only high ozone episode in the simulation period was related to the cold front passage. The
599 small-scale winds and high aloft ozone concentrations on 09/25 likely contributed to the ozone
600 exceedance in the area. It is also possible that an unreported emission event played a role. Since
601 the maximum surface ozone at La Porte was much higher than the morning-time ozone aloft, the
602 active local ozone production was likely the dominant factor. Analyses of ozone aloft on 09/25
603 showed while there was high ozone aloft and large negative model biases, the OA case tended to
604 have smaller biases, especially in late hours of the day.

605 Small-scale meteorological events such as wind transition and recirculation have been cited (e.g.,
606 Banta et al. 2005; Darby 2005) for their contributions to high ozone events. The model capability
607 in reproducing these events is critical in simulating such high ozone episodes. The base case did
608 not recreate the 25 September small-scale events likely due to the complex winds and a lack of
609 local information which can be used to steer model state closer to reality. On the other hand, the
610 inability of the sensitivity case to replicate the local winds is likely a result of the shortcomings
611 of the nudging process pending further investigation. An ongoing study by the current authors
612 suggests that errors in the meteorological fields from the default grid nudging files are important
613 sources. Methods are being tested to improve the quality of grid nudging files. Early results
614 showed that the bay breeze which caused the wind reversal around La Porte was well captured

615   through improved grid nudging files. In addition, more observational data (e.g., more sites and

616   higher data frequency) and more testing on the combination of nudging settings should help

617   improve the obs-nudging performance. Also, the impact of obs-nudging on precipitation and

618   clouds should be further investigated to understand their effect on chemistry.

624 **References**

625 Banta, R. M., Senff, C. J., White, A. B., Trainer, M., McNider, R. T., Valente, R. J., Mayor, S.

626 D., Alvarez, R. J., Hardesty, R. M., and Parrish, D.: Daytime buildup and nighttime transport of

627 urban ozone in the boundary layer during a stagnation episode, J. Geophys. Res.-Atmos., 103,

628 22519-22544, 1998.

629 Banta, R. M., Senff, C. J., Nielsen-Gammon, J., Darby, L. S., Ryerson, T. B., Alvarez, R. J.,

630 Sandberg, S. R., Williams, E. J., and Trainer, M.: A bad air day in Houston, B. Am. Meteorol.

631 Soc., 86, 657-669, doi:10.1175/bams-86-5-657, 2005.

632 Banta, R. M., Senff, C. J., Alvarez, R. J., Langford, A. O., Parrish, D. D., Trainer, M. K., Darby,

633 L. S., Hardesty, R. M., Lambeth, B., Neuman, J. A., Angevine, W. M., Nielsen-Gammon, J.,

634 Sandberg, S. P., and White, A. B.: Dependence of daily peak O_3 concentrations near Houston,

635 Texas on environmental factors: Wind speed, temperature, and boundary-layer depth, Atmos.

636 Environ., 45, 162-173, doi:10.1016/j.atmosenv.2010.09.030, 2011.

637 Byun, D., and Schere, K. L.: Review of the governing equations, computational algorithms, and

638 other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling

639 system, Appl. Mech. Rev., 59, 51-77, doi:10.1115/1.2128636, 2006.

640 Byun, D., Ngan, F., Li, X., Lee, D., Kim, S..: "Analysis of Air Pollution Events in Summer 2006

641 and Preparation of Model Input Data for the Assessment Study", Grant No. 582-5-64594- FY07-

642 02, Final Report: Evaluation of Retrospective MM5 and CMAQ Simulations of TexAQS-II

643 Period with CAMS Measurements, Texas Commission on Environmental Quality, February

644 2008, 25 pp

645 Cheng, F.Y., and Byun, D. : Application of high resolution land use and land cover data for

646 atmospheric modeling in the Houston-Galveston metropolitan area, Part I: Meteorological

647 simulation results, Atmos. Env., 42, 7795-7811, doi:10.1016/j.atmosenv.2008.04.055, 2008.

648 Cheung, V. T., and Wang, T.: Observational study of ozone pollution at a rural site in the

649 Yangtze Delta of China, Atmos. Env., 35, 4947-4958, 2001.

650   Choi, Y.: The impact of satellite-adjusted NO_x emissions on simulated NO_x and O_3 discrepancies

651   in the urban and outflow areas of the Pacific and Lower Middle US, Atmos. Chem. Phys., 14,

652   675-690, doi:10.5194/acp-14-675-2014, 2014.

653   Choi, Y. and Souri, A.: Chemical condition and surface ozone in large cities of Texas during the

654   last decade: observational evidence from OMI, CAMS, and Model Analysis, Remote Sensing of

655   Environ., 168:90-101, doi:10.1016/j.rse.2015.06.026, 2015

656   Choi, Y., Kim, H., Tong, D., and Lee, P.: Summertime weekly cycles of observed and modeled

657   NO_x and O_3 concentrations as a function of satellite-derived ozone production sensitivity and

658   land use types over the Continental United States, Atmos. Chem. Phys., 12, 6291-6307,

659   doi:10.5194/acp-12-6291-2012, 2012

660   Cuchiara, G. C., Li, X., Carvalho, J., and Rappenglück, B.: Intercomparison of planetary

661   boundary layer parameterization and its impacts on surface ozone concentration in the

662   WRF/chem model for a case study in Houston, Texas, Atmos. Environ., 175-185,

663   doi:10.1016/j.atmosenv.2014.07.013, 2014.

664   Czader, B. H., Li, X. S., and Rappenglueck, B.: CMAQ modeling and analysis of radicals,

665   radical precursors, and chemical transformations, J. Geophys. Res.-Atmos., 118, 11376-11387,

666   doi:10.1002/Jgrd.50807, 2013.

667   Czader, B.H., Choi, Y., Li, X., Alvarez, S., Lefer, B.: Impact of updated traffic emissions on

668   HONO mixing ratios simulated for urban site in Houston, Texas. Atmos. Chem. Phys., 15(3),

669   1253-1263, doi:10.5194/acp-15-1253-2015, 2015.

670   Darby, L. S.: Cluster analysis of surface winds in Houston, Texas, and the impact of wind

671   patterns on ozone, J. Appl. Meteorol., 44, 1788-1806, doi:10.1175/jam2320.1, 2005.

672   Daum, P. H., Kleinman, L. I., Springston, S. R., Nunnermacker, L. J., Lee, Y. N., Weinstein-

673   Lloyd, J., Zheng, J., and Berkowitz, C. M.: Origin and properties of plumes of high ozone

674   observed during the Texas 2000 Air Quality Study (TexAQS 2000), J. Geophys. Res.-Atmos.,

675   109, D17306, doi:10.1029/2003jd004311, 2004.

Deng, A., Stauffer, D., Gaudet, B., Dudhia, J., Hacker, J., Bruyere, C., Wu, W., Vandenberghe, F., Liu, Y., Bourgeois, A.: Update on WRF-ARW End-to-end Multi-scale FDDA System. 10th WRF Users' Workshop, Boulder, CO, NCAR, 2009

Diao, L., Roy, A., Czader, B., Pan, S., Jeon, W., Souri, A. H., and Choi, Y.: Modeling the effect of relative humidity on nitrous acid formation in the Houston area, Atmospheric Environment, 131, 78-82, 2016

Foley, K. M., Roselle, S. J., Appel, K. W., Bhave, P. V., Pleim, J. E., Otte, T. L., Mathur, R., Sarwar, G., Young, J. O., Gilliam, R. C., Nolte, C. G., Kelly, J. T., Gilliland, A. B., and Bash, J. O.: Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7, Geosc. Model Development, 3, 205-226, doi:10.5194/gmd-3-205-2010, 2010.

Gilliam, R. C., and Pleim, J. E.: Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW, J. Appl. Meteorol. Climatol., 49, 760-774, doi:10.1175/2009jamc2126.1, 2010.

Haman, C.L., Lefer, B., Morris, G.A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment. J. Atmos. Ocean Tech., 29, 697-710, 2012.

Haman, C.L., Couzo, E., Flynn, J.H., Vizuete, W., Heffron, B., Lefer, B.L.: Relationship between boundary layer heights and growth rates with ground-level ozone in Houston, Texas. J Geophys Res-Atmos., 119, 6230-6245, 2014.

Houyoux, M., Vukovich, J., Brandmeyer, J., 2000. Sparse Matrix Kernel Emissions Modeling System: SMOKE User Manual. MCNC-North Carolina Supercomputing Center. Available at: https://cmascenter.org/smoke/.

Huang, X.-Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J., Huang, W., Henderson, T., Bray, J., Chen, Y., and Ma, Z.: Four-dimensional variational data assimilation for WRF: Formulation and preliminary results, Mon. Weather Rev., 137, 299-314, 2009.

701   Kleinman, L. I., Daum, P. H., Lee, Y. N., Nunnermacker, L. J., Springston, S. R., Weinstein-

702   Lloyd, J., and Rudolph, J.: Ozone production efficiency in an urban area, J Geophys Res.-Atmos,

703   107, 4733, doi:10.1029/2002jd002529, 2002.

704   Le Dimet, F.X. and Talagrand, O.: Variational algorithms for analysis and assimilation of

705   meteorological observations: theoretical aspects, Tellus, 38A, 97–110, 1986.

706   Li, X., and Rappenglück, B.: A WRF–CMAQ study on spring time vertical ozone structure in

707   Southeast Texas, Atmos. Environ., 97, 363-385, doi:10.1016/j.atmosenv.2014.08.036, 2014.

708   Li, X., Lee, D., Kim, S.-T., Kim, H., Ngan, F., Cheng, F., and Byun, D.: Performance

709   Evaluation of a Year-long Run of an Air Quality Forecasting System for Southeast Texas,

710   10th Conference on Atmospheric Chemistry, New Orleans, January 2008, 2008.

711   Liu, Y., Bourgeois, A., Warner, T., Swerdlin, S., and Hacker, J.: An implementation of

712   observation nudging-based FDDA into WRF for supporting ATEC test operations, 2005 WRF

713   user workshop, Boulder, CO, 2005.

714   Liu, Y., Bourgeois, A., Warner, T., Swerdlin, S., and Yu, W.: An update on "observation

715   nudging"-based FDDA for WRF-ARW: Verification using OSSE and performance of real-time

716   forecasts, 2006 WRF user workshop, Boulder, CO, 2006.

717   Ngan, F., Byun, D., Kim, H., Lee, D., Rappengluck, B., and Pour-Biazar, A.: Performance

718   assessment of retrospective meteorological inputs for use in air quality modeling during TexAQS

719   2006, Atmos. Environ., 54, 86-96, doi:10.1016/j.atmosenv.2012.01.035, 2012.

720   Mason, R.; Strum, M.; Houyoux, M. Technical Support Document (TSD) Preparation of

721   Emissions Inventories for the Version 4, 2005-based Platform; U.S. Environmental Protection

722   Agency, Office of Air and Radiation, Office of Air Quality Planning and Standards, Air Quality

723   Assessment Division, 2010

724   Olaguer, E. P., Rappengluck, B., Lefer, B., Stutz, J., Dibb, J., Griffin, R., Brune, W. H., Shauck,

725   M., Buhr, M., Jeffries, H., Vizuete, W., and Pinto, J. P.: Deciphering the Role of Radical

726   Precursors during the Second Texas Air Quality Study, J. Air & Waste Manag. Assoc., 59, 1258-

727   1277, doi:10.3155/1047-3289.59.11.1258, 2009.

728   Otte, T. L.: The impact of nudging in the meteorological model for retrospective air quality

729   simulations. Part I: Evaluation against national observation networks, J. Applied Meteorol.

730   Climatol., 47, 1853-1867, doi:10.1175/2007jamc1790.1, 2008.

731   Pan, S., Choi, Y., Roy, A., Li, X., Jeon, W., and Souri, A.: Modeling the uncertainty of several

732   VOC and ints impact on simulated VOC and ozone in Houston, Texas, Atmos. Environ., 120,

733   404-416, 2015

734   Parrish, D. D., Allen, D. T., Bates, T. S., Estes, M., Fehsenfeld, F. C., Feingold, G., Ferrare, R.,

735   Hardesty, R. M., Meagher, J. F., Nielsen-Gammon, J. W., Pierce, R. B., Ryerson, T. B., Seinfeld,

736   J. H., and Williams, E. J.: Overview of the Second Texas Air Quality Study (TexAQS II) and the

737   Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS), J. Geophys. Res.-

738   Atmos., 114, doi:10.1029/2009jd011842, 2009.

739   Pour-Biazar, A., McNider, R. T., Roselle, S. J., Suggs, R., Jedlovec, G., Byun, D. W., Kim, S.,

740   Lin, C. J., Ho, T. C., Haines, S., Dornblaser, B., and Cameron, R.: Correcting photolysis rates on

741   the basis of satellite observed clouds, J. Geophys. Res.-Atmos., 112, D10302,

742   doi:10.1029/2006jd007422, 2007.

743   Rappengluck, B., Perna, R., Zhong, S. Y., and Morris, G. A.: An analysis of the vertical structure

744   of the atmosphere and the upper-level meteorology and their impact on surface ozone levels in

745   Houston, Texas, J. Geophys. Res.-Atmos., 113, doi:10.1029/2007jd009745, 2008.

746   Rappenglück, B., Lefer, B., Mellqvist, J., Czader, B., Golovko, J., Li, X., Alvarez, S., Haman,

747   C., and Johansson, J., 2011: University of Houston Study of Houston Atmospheric Radical

748   Precursors (SHARP), Report to the Texas Commission on Environmental Quality, August 2011,

749   145 pp

750   Souri, A. H., Choi, Y., Li, X., Kotsakis, A., and Jiang, X.: A 15-year climatology of wind pattern

751   impacts on surface ozone in Houston, Texas, Atmospheric Research,

752   doi:10.1016/j.atmosres.2016.02.007, 2016.

753   Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, M., Duda, K. G., Huang, Y.,

754   Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, 1-113,

755   2008.

756   Stauffer, D. R., and Seaman, N. L.: Use of 4-dimensional data assimilation in a limited-area

757   mesoscale model .1. Experiments with synoptic-scale data, Mon. Weather Rev., 118, 1250-1277,

758   1990.

759   Stauffer, D. R., and Seaman, N. L.: Multiscale 4-dimensional data assimilation, J. Appl.

760   Meteorol., 33, 416-434, 1994.

761   Tucker, S. C., R. M. Banta, A. O. Langford, C. J. Senff, W. A. Brewer, EJ. Williams, B. M.

762   Lerner, H. D. Osthoff, and R. M. Hardesty. : Relationships of coastal nocturnal boundary layer

763   winds and turbulence to Houston ozone concentrations during TexAQS 2006. Journal of

764   Geophysical Research: Atmospheres, 115, no. D10 (2010).

765   Willmott, C. J.: On the Validation of Models, Physical Geography, 2, 184-194,

766   doi:10.1080/02723646.1981.10642213, 1981.

767   Zhong, S. Y., In, H. J., and Clements, C.: Impact of turbulence, land surface, and radiation

768   parameterizations on simulated boundary layer properties in a coastal environment, J. Geophys.

769   Res.-Atmos., 112, doi:10.1029/2006jd008274, 2007.

770

771

772   Table 1. Major WRF physics and FDDA options, the numbers in the parentheses are the related

773   settings in WRF namelist file.

WRF Version	V3.5.1
Microphysics	Lin et al Scheme (1983)
Long-wave Radiation	RRTMG (Rapid Radiative Transfer Model for General Circulation Models), Iacono et al. (2008)
Short-wave Radiation	New Goddard scheme, Chou and Suarez (1999)
Surface Layer Option	Monin-Obukhov with Carlson-Boland viscous sublayer scheme (1978)
Land-Surface Option	Unified Noah LSM, Chen and Dudhia (2001)
Urban Physics	None
Boundary Layer Scheme	YSU, Hong et al. (2006)
Cumulus Cloud Option	Kain-Fritsch (2004)
FDDA	Grid nudging on for all; Observation-nudging on for the OA case

774
775   Table 2. Major CMAQ options, the text in the parentheses are the related settings in CMAQ

776   build script.

CMAQ version	V5.0.1
Chemical Mechanism	CB05 gas-phase mechanism with active chlorine chemistry, updated toluene mechanism, fifth-generation CMAQ aerosol mechanism with sea salt, aqueous/cloud chemistry
Lightning NOx emission	Included inline
Horizontal advection	YAMO (Yamartino)
Vertical advection	WRF omega formula
Horizontal mixing/diffusion	Multiscale
Vertical mixing/diffusion	Asymmetric Convective Model (ACM) version 2
Chemistry solver	EBI (Euler Backward Iterative)
Aerosol	AERO5 for sea salt and thermodynamics
Cloud Option	ACM cloud processor for AERO5
Boundary conditions	Default static profiles

777
778   Table 3 Statistics of surface T, U-wind and V-wind for three WRF simulations: N – data points;
779   Corr – Correlation; IOA – Index of Agreement; RMSE – Root Mean Square Error; MAE – Mean
780   Absolute Error; MB – Mean Bias; O – Observation; M - Model; O_M – Observed Mean; M_M –
781   Model Mean; SD – Standard Deviation; Units for RMSE/MAE/MB/O_M/M_M/O_SD/M_SD:
782   temperature, degree C; winds (U and V), m/s
783
784

Surface temperature T									
Case	N	Corr	IOA	RMSE	MAE	MB	O_M	M_M	O_SD	M_SD
No-OA	41058	0.83	0.89	2.0	1.5	0.9	27.4	28.3	3.1	2.8
1Hr-OA	41058	0.94	0.97	1.0	0.8	0.0	27.4	27.4	3.1	3.1

Surface U wind									
Case	N	Corr	IOA	RMSE	MAE	MB	O_M	M_M	O_SD	M_SD
No-OA	43246	0.76	0.84	1.4	1.1	-0.6	-1.3	-1.9	1.6	1.9
1Hr-OA	43246	0.81	0.89	1.0	0.8	-0.3	-1.3	-1.6	1.6	1.6

Surface V wind									
Case	N	Corr	IOA	RMSE	MAE	MB	O_M	M_M	O_SD	M_SD
No-OA	43246	0.76	0.8	2.1	1.7	1.2	0.4	1.7	2.0	2.6
1Hr-OA	43246	0.80	0.89	1.2	0.9	-0.1	0.4	0.4	2.0	2.0

785

786   Table 4 Statistics of ozone for CMAQ simulations, see table 3 for column header information

Case	N	Corr	IOA	RMSE	MAE	MB	O_M	M_M	O_SD	M_SD
No-OA	33308	0.72	0.78	14.9	12.3	9.3	24.4	33.7	16.5	14.1
1Hr-OA	33308	0.73	0.83	13.8	11.0	6.1	24.4	30.6	16.5	17.4

787

788 Table 5 Statistics of ozone on 09/25/2013, all day and hour 0 to 13. Both correlation and index of
789 agreement are unitless. The red numbers indicate the three hours (07 CST to 09 CST) when the
790 ozone in 1Hr-OA case is significantly better than the No-OA case due to much improved winds.

		No-OA		1Hr-OA	
	N	Corr	IOA	Corr	IOA
Hr|All	1150	0.79	0.86	0.81	0.88
0	48	0.04	0.30	0.40	0.46
1	43	0.20	0.24	0.36	0.30
2	48	0.14	0.25	0.35	0.35
3	48	0.19	0.30	0.32	0.35
4	48	0.27	0.36	0.31	0.35
5	47	0.24	0.36	0.28	0.37
6	47	0.33	0.38	0.35	0.37
7	48	0.06	0.39	0.29	0.47
8	48	0.09	0.43	0.53	0.63
9	47	0.05	0.41	0.55	0.74
10	47	-0.10	0.29	0.30	0.51
11	47	0.13	0.39	-0.07	0.36
12	49	0.09	0.38	0.25	0.40
13	49	-0.09	0.37	0.36	0.46

791