A comprehensive evaluation of the Eta-CMAQ forecast model performance for O_3, its related precursors, and meteorological parameters during the 2004 ICARTT study

Shaocai Yu[+$*], Rohit Mathur[+], Kenneth Schere[+], Daiwen Kang[+$],
Jonathan Pleim[+], and Tanya L. Otte[+]

Atmospheric Sciences Modeling Division, Air Resources Laboratory,
National Oceanic and Atmospheric Administration, Research Triangle Park, NC 27711

[+]In partnership with the U.S. Environmental Protection Agency,
Research Triangle Park, NC 27711
[$] On assignment from Science and Technology Corporation,
10 Basil Sawyer Drive, Hampton, VA 23666-1393

Submitted to

Journal of Geophysical Research
(*ICARTT special issue)*

[*] To whom the correspondence should be addressed: Phone: 919-541-0362
Fax: 919-541-1379, E-mail: yu.shaocai@epa.gov

ABSTRACT

The Eta-CMAQ model's forecast performance for ozone (O_3), its precursors, and meteorological parameters has been assessed over the eastern U.S. with the observations obtained by aircraft, ship, ozonesonde, and lidar and two surface networks (AIRNOW and AIRMAP) during the 2004 ICARTT study. The results at the AIRNOW sites show that the model was able to reproduce the day-to-day variations of observed daily maximum 8-hr O_3 and captured the majority (73%) of observed daily maximum 8-hr O_3 within a factor of 1.5 with normalized mean bias (NMB) of 22%. The model in general reproduced O_3 vertical distributions on most of the days at low altitudes but consistent overestimations above ~6km are evident due to a combination of effects related to the specifications of lateral boundary conditions from the Global Forecast System (GFS) as well as model's coarse vertical resolution in the upper free troposphere. The model captured the vertical variation patterns of the observed values for other parameters (HNO_3, SO_2, NO_2, HCHO, NO_y_sum (NO_y_sum=NO+NO_2+HNO_3+PAN)) with some exceptions, depending on the studied areas and air mass characteristics. The consistent underestimation of CO by ~30% from surface to high altitudes is attributed to the inadequate representation of the transport of pollution associated with Alaska forest fires from outside the domain. The model has good performance for marine or continental clear flows from the East/North/Northwest/South and southwest flows influenced only by Boston city plumes but overestimation for southeast flows influenced by the long-range transport of urban plumes from both New York City and Boston.

.

1. Introduction

Ozone (O_3) pollution is a major concern in the U.S. since it can adversely affect human and ecosystem health. Tropospheric O_3 is generated in the presence of solar ultraviolet radiation through a complex series of photochemical reactions involving many volatile organic compounds (VOC) and nitrogen oxides (NO_x), which originate either from anthropogenic sources (e.g., industry and vehicle emissions) or biogenic sources (e.g., forest and soil). Harmful levels of O_3 concentrations are typically observed during high pressure, hot, sunny and stagnant atmospheric conditions at the locations with substantial VOC and NO_x concentrations. According to the revised 8-hr National Ambient Air Quality Standard (NAAQS) for O_3 (0.08 ppm) promulgated by the U.S. EPA in 1997, *EPA* [2004] estimated that about 160 million Americans are exposed annually to the daily maximum 8-h O_3 concentrations that exceed this new NAAQS. Therefore, it is desirable for local air quality agencies to accurately forecast ozone concentrations to alert the public of the onset, severity and duration of unhealthy air and to encourage people to help limit outdoor activities and reduce emissions-producing activities (e.g., reduce automobile usage).

Real-time forecasting systems for O_3 with regional-scale air quality models have been developed and deployed for several years [*EPA*, 1999; *McHenry et al*., 2004; Cope et al., 2004; *McKeen et al*., 2005; *Kang et al*., 2005; *Otte et al*., 2005]. *McKeen et al*. [2005] statistically evaluated the real-time forecasts of O_3 from seven air-quality forecast models over the eastern U.S. and southern Canada during the summer of 2004, and concluded that relative to any individual model, the ensembles, which were based on the mean and median of the seven models, had higher correlation coefficients, lower root-mean-square errors (RMSE), and better threshold statistics, pointing to ensemble modeling as a better real-time O_3 forecast tool.

The regional air quality in New England was a focus of the 2004 International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) study. Two of the major goals of the 2004 ICARTT study were to link surface air quality with the important features of transport and chemistry that occur above the surface and to determine the relative importance of local pollution compared to long-range transport in shaping local air quality. The temporal and spatial coordination of the measurement platforms afforded by the 2004 ICARTT consortium provided comprehensive coverage that is invaluable to the evaluation and improvement air-quality models and model forecasts for understanding regional complex pollution events. In this

study, the National Weather Service's (NWS) operational mesoscale forecast Eta model is used to supply meteorological input to the Community Multiscale Air Quality (CMAQ) model (Eta-CMAQ model suite); the models are then used to provide predictions of O_3 and related chemical species in the forecast mode. The purposes of this paper are two-fold. First, the temporal and spatial performance of the Eta-CMAQ forecast model for O_3 is evaluated against the observations from the Air Quality System (AQS) network over the eastern U.S. Second, the ability of the Eta-CMAQ model to predict air quality and the meteorological conditions dictating episodes of high O_3 horizontally and vertically is comprehensively examined on the basis of the extensive measurements obtained by aircraft, ship, ozonesonde, and lidar during the 2004 ICARTT field experiment. Note that the summer of 2004 in the eastern U.S. exhibited very few O_3 "episodes" or exceedances due to unusually cool and wet conditions (i.e., temperatures either below or much below normal, and precipitation either above or much above normal [*http://www.ncdc.noaa.gov*]), associated with continental polar air masses during July, and the influence of several hurricanes during August. Therefore, the model performance presented here is probably not climatologically representative of summertime conditions, but is unique to the summer of 2004.

2. Description of the Eta-CMAQ Forecast Model and Observational Database

The Eta-CMAQ air quality forecasting system [Otte et al., 2005], created by linking the Eta model [*Rogers et al*., 1996] and the CMAQ Modeling System [*Byun and Ching*, 1999; *Byun and Schere*, 2006], was applied over a domain encompassing the eastern U.S. (Figure 1) during summer 2004. The linkage of the two modeling systems is described in detail by Otte et al. [2005]. A series of post-processors interpolates the Eta model output fields in the horizontal and in the vertical onto a coordinate structure and map projection that are compatible with CMAQ. The model domain has horizontal grid spacing of 12 km. Twenty-two layers of variable thickness are specified on a sigma vertical coordinate system to resolve the atmosphere between the surface and 100 hPa. The thickness of layer 1 is about 38 m. The lateral boundary conditions are a horizontally constant and are specified by continental "clean" O_3 and other trace gas profiles with some vertical variation based on climatology. To improve representation of O_3 in the free troposphere and possible effects related to stratospheric intrusion, the O_3 lateral boundary conditions above altitudes of 6 km were augmented using O_3 forecast results from

NCEP's Global Forecast System (GFS). The primary Eta-CMAQ model forecast for next-day's surface-layer O_3 is based on the current day's 12 UTC Eta simulation cycle. The target forecast period is local midnight through local midnight (04 UTC to next day's 03 UTC). The emissions are projected to 2004 from the 2001 U.S. EPA national emission inventory [Pouliot, 2005]. The Carbon Bond chemical mechanism (version 4.2) has been used to represent photochemical reaction pathways.

The hourly, near real-time O_3 data at 614 sites in the eastern U.S. are available from the U.S. EPA's Air Quality System (AQS) network (Figure 1), resulting in nearly 1.2 million total hourly O_3 observations for the study period (see Table 1). Four Atmospheric Investigation, Regional Modeling, Analysis, and Prediction (AIRMAP) [*DeBell et al*., 2004; *Mao and Talbot*, 2004] sites provided continuous measurements of O_3 and related photochemical species as well as meteorological parameters during the study; the sites include Castle Springs (CS) (43.73^{0}N, 71.33^{0}W) (New Hampshire (NH)), Isle of Schoals (IS) (42.99^{0}N, 69.33^{0}W) (Maine), Mount Washington Observatory (MWO) (44.27^{0}N, 71.30^{0}W) (NH), and Thompson Farm (TF) (43.11^{0}N, 70.95^{0}W) (NH). From July 1 to August 15, 2004, measurements of vertical profiles of O_3, its related chemical species (CO, NO, NO_2, H_2O_2, CH_2O, HNO_3, SO_2, PAN, isoprene, toluene), and meteorological parameters (liquid water content, water vapor, temperature, wind speed and direction and pressure) were carried out by instrumented aircraft (NOAA P-3 and NASA DC-8), ozonesonde and ship-based lidar deployed as part of the 2004 ICARTT field experiment. The detailed instrumentation and protocols for measurements are described in http://www.al.noaa.gov/ICARTT/FieldOperations/. The flight tracks of P-3, DC-8, ship and locations where daily ozonesondes were launched are presented in Figure 2. The model performance during July 1-August 15, 2004 is examined in this study based on the 12 UTC model run for the target forecast period.

3. Results and Discussion

3.1. Spatial and Temporal Evaluation of O_3 over the Eastern U.S. Domain at the AQS Sites

Table 1 summarizes the evaluation results for the hourly, daily maximum 1-hr and maximum 8-hr O_3 concentrations for two cases; one using all data and other only using data with O_3>40 ppbv. As can be seen, the Normalized Mean Bias (NMB) and Normalized Mean Error (NME) values for the data only with O_3>40 ppbv range from 6.1 to 11.9% and 18.2 to 21.5%,

respectively, much lower than those when all data are used, indicating that the overestimations in the low O3 concentration range significantly contribute to the overall overestimations. The recommended performance criteria for O_3 by U.S. EPA (1991) are: mean normalized bias ±5 to ±15%; mean normalized gross error 30% to 35%; unpaired peak estimation accuracy: ±15 to ±20%. Table 1 shows that for the case only using data with O_3>40 ppbv, the Normalized Mean Bias (NMB) and Normalized Mean Error (NME) values for maximum 1-hr (maximum 8-hr) O_3 are 7.0% (11.9%) and 18.2% (19.7%), respectively, close to the performance criteria for the unpaired peak O_3. Additional insight into the AQF modeling system's positive bias (over prediction) and error (scatter) can be gained from Figure 1a, which shows that the model reproduced the majority (73.1%) of the observed daily maximum 8-hr O_3 concentrations within a factor of 1.5 but generally over-estimated the observations in the low O_3 concentration range (<50 ppbv), in part, due to the vertical resolution in layer 1 assumed high background O_3 levels specified in these simulations. It is believed that the points in the low O_3 concentration ranges below the background level reflect depletion of O_3 from deposition or titration by NO [*Lin et al.*, 2000]. The overestimation of the observations in the low O_3 concentration ranges could be indicative of titration by NO in urban plumes that the model does not resolve because most of the AQS sites are located in urban areas. This is further supported by the fact that most of sites with MB (mean bias)>40 ppb in the low O_3 concentration range (<50 ppb) for the maximum 8-hr O_3 in Figure 1a are located within the big cities and along the Washington, D.C./New York City/Boston urban corridor (not shown). In order to investigate the AQF system's performance over time, the values of mean, NMB and correlation coefficients were calculated (domain wide averages) and plotted as a daily time series for the daily maximum 8-hr O3 concentrations (Figure 1b). Although the forecasts tracked the general temporal pattern well, the over estimation discussed above was prevalent throughout the four month period. The domain-wide mean values of NMB (Normalized Mean Bias) and NME (Normalized Mean Error) during the ICARTT period for maximum 8-hr O_3 are 22.6% and 28.8%, respectively. The model had the best performance on August 8 (NMB=1.9%, correlation coefficient (r)=0.73) and the worst performance on August 12 (NMB=42.4%, r=0.47). A close inspection of the synoptic-scale meteorological conditions (not shown) reveals that on August 8, the majority of the domain was dominated by high pressure and clear sky (conditions that are conducive to O_3 formation), whereas on August 12, an active cold front stretched from the north to south accompanied by

convective cloud cover and precipitation through the domain under low pressure. As shown by the diagnostic analysis [*Mathur et al.*, 2004], the significant overestimation in areas of cloud cover is mainly caused by the unrealistic vertical transport of excessive amounts of high O_3 concentrations near the tropopause to the ground associated with downward entrainment in CMAQ's convective cloud scheme. Spatially, the model performed better over the western region with NMB of ±25% than eastern coastal region of the domain with NMB>25% (Figure 1c and 1d). The largest overestimation of the observed daily maximum 8-hr O_3 concentrations was in the northeast (NMB>+50% and NME>50%), mainly caused by very low observed O_3 concentrations, which typically coincide with non-conducive meteorological conditions (i.e., cloud cover, precipitation and cool temperatures). Biases and errors associated with the maximum 1-hr O_3 (not shown) follow a similar pattern to the maximum 8-hr O_3.

3.2 Evaluation of Vertical Profiles for O_3, Its Related Species and Meteorological Parameters

Comparisons of modeled vertical profiles with aircraft, ozonesonde, and ship-based lidar observed vertical profiles provide an assessment of the ability of the model to simulate the vertical structure of air pollutants and meteorological fields. Following *Mathur et al*. [2005], modeled results were extracted by "flying" the aircraft through the 3-D modeling domain by mapping the locations of the aircraft to the model grid indices (column, row, and layer). Hourly resolved model outputs were linearly interpolated to the corresponding observational times. The flight tracks of aircraft show that measurements onboard the P-3 cover a regional area over the northeast around New York and Boston (Figure 2a) from 0 to ~5km altitudes, whereas the DC-8 aircraft covers a broader regional area over the eastern U.S. (see Figure 2b) between 0-12 km altitudes. All DC-8 measurements were conducted in the daytime (~7:00 to ~19:00 EST), and P-3 also conducted most of its measurements during the daytime except on 7/11, 7/31, 8/3, 8/7, 8/9, and 8/11 when the P-3 measurements were conducted at the night (~20:00 to ~6:00 EST). To compare the modeled and observed vertical profiles, the observed and modeled data were grouped according to the model layer for each day and each flight because different flights had different mission tasks for each day as summarized in Tables 2 and 3. Thus, these vertical profiles may be regarded as representing the mean conditions along the flight track for each day. Figures 3-7 present modeled and observed vertical daily profiles for O_3, its related species, and

meteorological parameters during the ICARTT period. The temporal variations of modeled and observed JNO_2 (photolysis rates of NO_2) along the flight tracks for the daytime only are shown in Figures 8 and 9.

As shown in Figures 3a and 3b, the model generally reproduced the observed O_3 vertical structure on most days with the best performance on 7/25, 7/27, 7/31, 8/6, 8/7, 8/9, and 8/11 for the P-3 measurements and 7/18, 7/22, 8/6, 8/7, 8/13, 8/14 for the DC-8 measurements, although it tended to overestimate in the upper layers (because of higher O_3 levels aloft from the GFS model), especially for DC-8 observations at altitudes >6 km. Noticeable among these are generally the better model performance during nighttime (7/11, 7/31, 8/3, 8/7, 8/9 and 8/11) relative to P-3 observations. A close inspection of the temporal variations of modeled and observed O_3 along the flight tracks (not shown) reveals that the modeled overestimations of observed O_3 at the low altitudes for most days in Figures 3a and 3b occurred over the ocean regions. Comparisons of time-height variations in O_3 structure along the ship tracks (Figure 3c) indicates that the model predicted more uniform vertical O_3 profiles than the observations, and the overestimations increase with altitude based on the lidar measurements over the ocean off the coast of New Hampshire (NH) and Maine (see Figure 2c). The poor model performance over the ocean may be tied to poor representation of coastal boundary layers and their interaction with land/sea breezes in the interpolated meteorological model fields. Comparisons of vertical profiles of median O_3 concentrations at five sites on the basis of ozonesonde observations reveal a consistent model overestimation above ~6 km (Figure 3d), although the model reproduces the O_3 vertical profile well at the low altitudes, especially at the Pellston site. As discussed before, this higher bias at the high altitude is attributed to the lateral boundary conditions derived by the global forecast system (GFS) model and coarse model resolution in the free troposphere, and is consistent with the DC-8 comparisons shown in Figure 3b.

The model's ability to simulate the vertical profiles for other parameters (CO, HNO_3, SO_2, NO, NO_2, HCHO, NO_y_sum (NO_y_sum=NO+NO_2+HNO_3+PAN) and NO_y (NO_y=NO + NO_2 + NO_3 + 2*N_2O_5 + HONO + HNO_3 + PNA + PAN + NTR)), as measured by the P-3 and DC-8 aircrafts, is illustrated in Figures 4-6. In general, the model captured the vertical variation patterns of the observed values for various species, with some exceptions, depending on the studied regions and air mass characteristics. Noticeable among these are the consistent underestimations for CO vertical profiles on most days except 7/15, 7/31, 8/14 and 8/15 for P-3

measurements, and 7/18 and 7/31 for DC-8 measurements. The summary of Table 3 indicates that there was a widespread signature of biomass burning plume (i.e., the observed acetonitrile, the biomass burning plume tracer, was strongly enhanced) over the studied areas except these days, which were only significantly affected by the urban (New York, Boston or Washington and Baltimore) plumes. One of the reasons for this under-estimation of CO is attributed to the inadequate representation of the transport of pollution associated with biomass burning from outside the domain [*Mathur et al*., 2005; *McKeen et al*., 2002]. The significant underestimations of CO during July 20 and July 22, 2005, further support this explanation as the aerosol index images from the TOMS satellite observations [http://toms.gsfc.nasa.gov/] clearly show that the eastern U.S. was significantly influenced by pollutants from large Alaskan forest fires during these days. Tables 3 and 4 indicate that a progressively more aged NY city plume over the Bay of Fundy on 7/22 with a widespread signature of biomass burning was sampled by both P-3 and DC-8, confirming the conclusion.

Another noticeable discrepancy is the consistent underestimations of observed NO at altitudes greater than 6 km relative to DC-8 measurements (Figure 6b). This may be because the aircraft and lightning NO emissions are not included in the current model emission inventory. On 7/9, 7/21 7/27 and 7/28, the P-3 encountered the fresh city plume (Boston or New York) shortly after takeoff as summarized in Table 3 with very high NO concentration at low altitudes. The model estimations also missed these high NO concentrations at low altitudes as shown in Figure 6a.

As summarized in Table 3, the P-3 sampled the plume of Ohio Valley power plants at ~1000 m during 8/6 from 15:30 to 20:30 UTC and 8/10 from 0:30 to 3:30 UTC. Figures 3-6 show that the model reproduced the SO_2, NO, NO_2, HNO_3, O_3 and NOy_sum concentrations well relative to P-3 observations in the power plant plumes at this height for these two days. However, the model overestimated SO_2 in the NYC and Boston plumes at low altitudes <700m for these two days. The modeled SO_2 concentrations are generally higher than the observations at the low altitude (<200 m) most of the time when the P-3 sampled the urban plumes of New York and Boston except 7/21 and 8/7, indicating that the model may have overestimated some of emission sources of SO_2 from the New York and Boston areas.

The point source emissions from power plants are often rich in SO_2 and NO_x and mobile sources (or urban plumes) are rich in CO and NO_x. On 7/27, the surface weather map showed

convective activity associated with a surface cold front that stretched from the center of a surface low over the West Virginia-Pennsylvania state line to the Southwest along the Appalachian Mountains with thunderstorms. There was pollution accumulation ahead of the cold front. The pollution upwind and downwind of the Washington and Baltimore metropolitan area between 600 and 2000 m altitudes was sampled by the P-3 during 17:30 to 18:30 UTC with very high SO_2 (>5 ppb), CO (>180 ppb), and HNO_3 (>3 ppb) but low O_3 (~60 ppb) concentrations. The model reproduced the low O_3 concentrations, but underestimated all other species (CO, HNO_3, NO, NO_2, NOy_sum) below 2 km for this pollution accumulation event ahead of the cold front as shown in Figures 3-6.

The model shows good performance for HNO_3 most of the time except 7/9, 7/21, 7/22, 7/27, 7/28, and 8/11 relative to P-3 observations and 7/18, 7/20 and 7/22 relative to DC-8 observations as shown in Figures 4c and 4d. The model overestimated the HNO3 concentrations at the low altitudes in the air masses containing fresh plumes such as 7/9 and 7/21-7/22. The model performance for NO_y_sum is generally very good most of the time except 7/21, 7/27, 8/14 and 8/15 at the low altitudes as shown in Figure 6. NO_2 follows the same pattern of NO_y_sum for the model performance. The very good model performance of NOy_sum combined with consistent overestimations of NOy on 7/11 and 7/15 in Figure 6 reveals that the model overestimated the sum of NO_3, N_2O_5, HONO, PNA and NTR. Possible reasons for these overestimations, including (1) the sharp nocturnal gradients near the surface are not resolved in the model and (2) uncertainties associated with atmospheric sinks for the modeled terminal organic nitrate species represented by the lumped species called NTR in the CB IV chemical mechanism.

Figure 7 reveals that the Eta model reproduced the vertical profiles of observed water vapor and wind speed very well most of the time and is in better agreement with the DC-8 observations. Specifically, the model consistently overestimated water vapor at low altitudes relative to P-3 observations, especially on 7/27 and 7/28 when there was a surface cold front across the northeastern domain. The model also overestimated water vapor at low altitudes relative to the DC-8 observations on 7/28, indicating that the model did not reproduce moisture well for the cold front system. The model seems to consistently underestimate the wind speed slightly in layer 1 most of the time except on 7/31 and 8/3 relative to the P-3 observations. The

model also tracked the vertical variations of temperatures, pressures and wind directions very well most of the time (not shown).

Cloud and aerosol can significantly affect photolysis rates of NO_2 (JNO_2) by enhancing and reducing the UV actinic flux, depending on their optical properties, solar zenith angle, and the position of the layer of interest relative to the observation point. Figures 8 and 9 show that the model captured the temporal and vertical variations of the observed JNO_2 very well for some periods but was weighted too low or too high for other periods along the flight tracks for each day. Upon a closer inspection of visible satellite images and aircraft observations, it is noted that the model generally captured the observed JNO_2 very well during the cloud-free periods, but underestimated the JNO2 values by 20-90% (see Figure 9) when there was a solid cloud deck below the aircraft such as the period of 14:00 to 15:00 EST on 7/15 (see Figure 10) and overestimated JNO2 values significantly when solid cloud deck is above the aircraft such as the period of 14:00 to 15:00 EST on 8/6 (see Figure 9). Note that the cloud location information in Figure 10 was obtained according to the observational documents (See Table 3). As summarized in Table 3, the P-3 encountered the plume of Ohio valley power plants (e.g., cloud-free polluted conditions) at ~1000m during 10:30 and 15:30 EST on 8/6. A very large fluctuation of JNO_2 values varying from 6.9×10^{-4} to 1.1×10^{-1} s^{-1} at this altitude was observed due to the significant effects of the strongly scattering aerosols within the power plant plume (see Figure 8). The relatively constant modeled JNO_2 values of approximately 8.6×10^{-3} s^{-1} during the power plant plume sampling period indicate that the model generally overestimated the observed JNO_2 without capturing its fluctuations within the power plant plume.

3.3. Time-Series Comparisons over the Ocean surface with the Ron Brown Ship Observations

The cruise tracks of the NOAA ship Ronald H. Brown in Figure 2 indicate that most of ship's time was spent sampling along the coast of New Hampshire, Massachusetts and Maine. Anthropogenic sources from the Washington, D.C./New York City/Boston urban corridor and biogenic emissions in New Hampshire and Maine significantly impact the sampled air masses along the coast of New England. Driscoll et al. [2003] found that NO_x emissions in the Northeast U.S. are primarily from the transportation (54%), electric utilities (25%) and industrial sources (11%). The time-series and scatter plots of the model predictions and observations for

each parameter (O_3, CO, NO_y, NO_2, NO, SO_2, Isoprene, wind speed, wind direction, RH, photolysis rates for O_3 (JO_3) and NO_2 (JNO_2)) along the ship tracks during the ICARTT period are shown in Figure 11. The air mass flow patterns sampled in the Gulf of Maine can be divided into two groups as shown in Figure 12b. One is the offshore flow from the southwest and west, and another is the relatively clear marine and continental flow from east, south, north and northwest as summarized in Table 5. The air masses in the southwest offshore flows had passed over the urban New York/Boston corridor during the previous 2-24 hours before being sampled at the ship. *Angevine et al*. [2004] showed that transit times from Boston and NYC to the regions of ship measurements (Gulf of Maine) were approximately 2-3 hours and 12 hours, respectively. These southwest offshore flows led to high pollution episodes along the New England coast. The sampled air masses in the westerly flows were typically about 2-4 hours downwind of Boston and the surrounding forested areas. As indicated in Figure 11 and Table 5, the urban plumes from NYC and Boston in the southwest/west offshore flows were clearly seen above the background concentrations for each species on days 7/10, 7/15-7/17, 7/20-7/23, 7/29-8/1, 8/3-8/4, 8/8-8/12, and 8/16-8/17, whereas the clear marine or continental flows from the East/North/Northwest/South mainly impacted the ship observational areas on 7/11-13, 7/18, 7/25-28, and 8/5-7 days characterized by low concentrations for O_3, CO, NO_y, NO_2, NO, SO_2. Note that due to unusually cool and wet conditions with temperatures either below or much below normal and precipitation either above or much above normal over the eastern U. S. during the summer of 2004, O_3 concentrations are not very high, even during the pollution episodes. There was very good model performance for the clear marine or continental flows from the East/North/Northwest/South on days 7/11-13, 7/18, 7/25-28, and 8/5-7 for each species (O_3, NO_y, NO_2, NO, SO_2) as shown in Figure 11 and Table 5. This suggests that the model can simulate the background environments for clear marine or continental flows very well.

The model overestimated the observed O_3 in all southwest/west offshore flows except on days 7/16-7/17, 8/4 and 8/8-8/11. Two case studies for southwest offshore flows on 7/20 and 7/30 are shown in Figures 12 and 13, respectively. On these days, long-range transport of urban plumes from both NYC and Boston region significantly impacted the atmosphere over the Gulf of Maine during the late afternoon as illustrated in Figure 12c and 13b. Comparison with the ozonesonde observations in Figures 12d and 13c shows that on these two days, the model simulated O_3 at 18:40 UTC very well from ground to high altitudes (~2 km) at Narragansett, RI,

but consistently overestimated morning O_3 (by ~50 ppb) between the surface and ~2 km over the ocean as shown by the ship ozonesonde. The model reproduced surface O_3 concentrations at ~22 UTC at the Portsmouth site when the ship arrived back in Portsmouth as shown in Figures 12c and 13b for both 7/20 and 7/30. Figure 13 also reveals that the model actually simulated the surface high O_3 (>80 ppb) very well at the AQS coastal sites in NH and Maine such as Reid State Park (Maine), Cape Elizabeth (Maine), Portsmouth (NH), Odiorne State Park (NH) and Newbury (MA). On the other hand, a case study for the southwest offshore flow on 8/10 illustrated in Figure 14 reveals that the transport of only Boston urban plume impacted the ship observational region (Gulf of Maine). The relative weak wind speeds (see Figure 11) later during the day reveal that sea breeze carried polluted air from the coastal waters inland into New Hampshire (see Figure 14b); similar features were found during 2002 Northeast Air Quality Study (NEAQS) [*Angevine et al*., 2004]. The model reproduced this episode very well as shown in Figures 14a and 14b. The significant increases of observed CO, NO_y, NO_2, and NO in Figure 11 during this day also strongly indicates the fingerprint of the fresh urban plumes directly from the Boston city. The model captured the buildup of these species well although it tended to overestimate their concentrations. Compared to the ozonesonde profiles, Figure 14c shows that the model O_3 vertical profiles between surface and ~2 km are close to the observations both at Narragansett, RI, and Ron Brown ship although the model results are slightly lower than observations by ~10 ppb. The better model performance for O_3+NO_2 than for O_3 at low concentrations as shown in Figure 14a which reveals that the model exaggerated the effects of NO titration on O_3 as inferred from the O_3 observations during the nighttime over the ocean. Upon a closer inspection, it is noted that on other days with good model performance for southwest/west offshore flows, i.e., 7/16-7/17, 8/4 and 8/8-8/11, the ship observations in Gulf of Maine were significantly affected only by Boston city plumes according to the model simulations. As pointed out by *Angevine et al* [2004], pollutant concentrations in stable layers over coastal water surfaces are allowed to remain high due to the lack of deposition or deep vertical mixing on the over-water trajectories. Our analyses suggest that for conditions involving southwest offshore flows impacted by long-range transport of NYC and Boston urban plumes, the model overestimated the O_3 concentrations over the ocean regions, but simulated the O_3 concentrations well over the ocean regions under the conditions impacted only by the Boston plumes. The transport patterns determine the model performance, indicating that the model does not simulate well the transport

over land-ocean interface. This suggests additional investigation of the representations of boundary-layer mixing and dry deposition over the ocean in the model. The large discrepancies between the model and observations for coastal grid cells where the model results are too high are due to the fact that the model's boundary-layer mixing cannot resolve steep subgrid land-to-sea gradients.

Figure 11 indicates that the model captured, with a good deal of fidelity, the temporal variations and broad synoptic changes seen in the observed wind speed, wind direction and relative humidity (RH) along the ship track most of the time, especially for RH, although with some occasional major excursions. The model reproduced the diurnal variations in the observed JNO2 very well along the ship track most of the time, except on the peaks of 7/9, 7/18, 7/19, 7/27, 8/5, and 8/8 in which the model seriously over estimated the observations. Misplacements of cloud cover in the model results in the overestimations of the observed JNO_2 (not shown), which can also contribute to the higher O_3 bias. The model performance for JO_3 follows those of JNO_2 as shown in Figure 11.

3.4 Time-series comparison and diagnostic evaluation at the AIRMAP sites during the 2004 ICARTT

Figure 15 presents time-series comparisons and scatter plots of the model predictions and observations for O_3, CO, NO, NO_y, SO_2, JNO_2, temperature (T) and RH at the CS site. Following *Yu et al*. [2003; 2006], the percentages of the comparison points where the model results are within a factor of 1.5 and 2, respectively, of the observations are listed in Table 6. The model captured the hourly variations and broad synoptic changes seen in the observations of each parameter (O_3, CO, NO_y, JNO_2, T and RH) (correlation coefficient>0.49, see Table 6) except NO and SO_2 at CS, IS and TF sites. The serious underestimation of NO, CO, NOy and SO_2 at the MWO site (the highest mountain (1916 m) in the northeastern U.S.), in part, reflects the inherent subgrid variability in their emissions and concentrations that are not adequately captured by the model grid structure. This is also due to the fact that usually the models misrepresent mountain sites because they essentially sample free tropospheric air while models can't resolve the terrain. Relatively large discrepancies between modeled and measured concentrations are noted for primary species, such as NO and SO_2. These are likely related to the discrepancies between modeled and observed wind speed and direction, which cause modeled

plumes to be displaced leading to relatively larger error for primary species when the modeled and measured values are paired in space and time. The model underestimated CO by 20-50% consistently at each site, similar to those comparisons for the vertical profiles. The model overestimated NO_y at the CS and TF sites like those comparisons from the aircraft measurements. The model reproduced the observed temperatures with ~±5% errors and relative humidity (RH) with ~±10% at each site.

The analysis of photolysis rates of NO_2 focuses on daytime data by excluding data where $JNO_2 < 5\times10^{-5}$ s^{-1} following *Thornton et al*. [2002]. Table 6 indicates that the model reproduced 49.6%, 43.1% and 53.8% of observed JNO_2 values within a factor of 1.5 at the CS, MWO and TF sites, respectively. DeMore et al. [1997] suggest that about ±20% uncertainty in photolysis rates can be associated with uncertainty in the cross-section and quantum yield data used in the calculation of JNO_2 values. The sensitivity tests of Hanna et al. [2001] indicate that a 50% uncertainty in JNO_2 could cause about a 40 ppbv, or a 20% uncertainty in predicted maximum O_3 concentration in their cases. Additional uncertainties in the model simulations can also arise from uncertainties and errors associated with the spatial and temporal representation of cloud fields in the model and their subsequent effects on photolysis attenuation.

The upper limits of the ozone production efficiencies (ϵ_N) value can be estimated by the O_3-NO_z ($NO_z = NO_y - NO_x$) slope. *Jacob et al* [1995] estimated NO_2 concentrations for daytime conditions during the Shenandoah Cloud and Photochemistry Experiment (SCAPE) by assuming the $NO/NO_2/O_3$ photo-stationary steady state (PSS) in order to obtain NO_x and NO_z concentrations. Following *Jacob et al.* [1995], *Griffin et al*. [2004] and *Kleinman et al*. [2004], NO_2 concentrations at the CS and TF sites were estimated based on the PSS assumption in this study. On the basis of comparisons between observed and calculated NO_2 from the field program, *Kleinman et al*. [2004] estimated an accuracy of ±25% for the calculated NO_2 values by the PSS assumption for NO_x to NO_y ratios in fresh plumes being typically in the range of 75-110%. Based on the ship data in which NO_2 concentrations were observed (see Section 3.3) in this study, it was found that the mean NO_2 concentrations for the observations and estimations by the PSS assumption are 2.41±2.87 and 3.36±4.36 ppbv, respectively, with correlation coefficient of 0.932 between them. The PSS assumption overestimated the observed NO_2 by 28% in this case. In this study, the PSS assumption is only used to estimate NO_2 concentrations at the CS and TF sites. The $[O_3]/[NO_x]$ values can be used to determine NO_x-sensitive and VOC-sensitive

chemical regimes. *Arnold et al*. [2003] showed that $[O_3]/[NO_x]$ values >46 indicate strong NO_x-sensitive conditions, whereas values <14 indicate VOC-sensitive conditions. Table 7 summarizes the variations in the $[O_3]/[NO_x]$ ratio at the CS and TF sites. The results along the Ron Brown ship tracks are also listed in Table 7 for comparison. As can be seen, the model generally reproduced the temporal variations in the observed $[O_3]/[NO_x]$ ratios across the different conditions represented at the three locations. Both model and observations show that the CS site is mainly under strongly NO_x-sensitive conditions (>66%), whereas the TF site and ship over the ocean are under neither strongly NO_x-sensitive nor VOC-sensitive conditions. Following *Arnold et al*. [2003], both modeled and observed O_3-NO_z slopes are obtained for only observational data with $[O_3]/[NO_x]$>46 at the CS and TF sites and on ship. There is significant correlation between O_3 and NO_z for both model predictions and observations (r>0.61) at the three locations (see Fig. 16 and Table 8). While the ε_N values of the model (5.2 to 6.4) and observation (8.5 to 10.7) at the CS and TF sites are close to the lower and higher bounds of the estimated ranges (5 to 10) of other investigators [*Olszyna et al.*, 1994; *Fiore et al*., 2002] at rural sites in the eastern US, respectively, the modeled ε_N value is about 40 % lower than the observations. The modeled intercepts (background O_3) are consistently higher than the observations at each location. The results along the ship tracks over the ocean in Table 8 and Figure 16 reveal that the modeled ε_N value (3.6) is much lower than the corresponding observation (11.7) and both modeled and observed ε_N values are outside of estimation range (5 to 10) of other investigators at rural sites in the eastern US. As suggested by Chin et al. [1994], the ε_N values estimated by the O_3-NO_z slopes are upper limits because NO_z species (primarily HNO_3) are removed from the atmosphere more rapidly than O_3. Figure 16 shows that compared to the observations, the model produced less O_3 at the high NO_z regime. The scatter plots of Figure 16 also reveal that the modeled NO_z concentrations were higher than the observations, indicating that the model chemistry produces more terminal oxidized nitrogen products than inferred from observations, thereby contributing in part to the noted under estimation of ε_N.

4. Summary

A rigorous evaluation of the Eta-CMAQ forecast model performance for O_3, its related precursors, and meteorological parameters has been carried out over the eastern U.S. by

comparing the model results with the observations using extensive measurements obtained during the 2004 ICARTT study. The results at the AIRNOW surface sites shows that the model was able to reproduce the day-to-day variations of observed daily maximum 8-hr O_3 and captured the majority (73%) of observed daily maximum 8-hr O_3 within a factor of 1.5 with NMB=22%. The model significantly overestimated the O_3 concentrations in areas of cloud cover mainly caused by the unrealistic vertical transport in CMAQ's convective cloud scheme. On the basis of results from aircraft, ozonesonde and ship-based lidar observations, the model generally reproduced O_3 vertical structures most of the days at low altitudes with consistent overestimations above ~6km due to the lateral boundary conditions derived by the GFS and coarse model resolution in the free troposphere. The model consistently underestimates CO by ~30% from surface to high altitudes because of the inadequate representation of the transport of pollution associated with biomass burning from outside the domain. The model captured the vertical variation patterns of the observed values for other parameters (HNO_3, SO_2, NO_2, HCHO, NOy_sum) with some exceptions, depending on the studied regions and air mass characteristics. The consistent underestimation of observed NO at altitudes>6 km relative to DC-8 measurements is attributed, in part, to the exclusion of aircraft and lightning NO emissions in the real-time model emission inventory. The very good model performance of NO_y_sum relative to consistent overestimation of NOy reveals that the model overestimated sum of NO_3, N_2O_5, HONO, PNA and NTR. The model can generally capture the observed JNO_2 very well during the cloud-free periods, but underestimated the JNO2 values by 20-90% when there was a solid cloud deck below the aircraft and overestimated JNO2 values significantly when solid cloud deck is above the aircraft. On the other hand, the model has reproduced the vertical profiles of observed water vapor and wind speed.

The capability of the model to reproduce the observed pollutants over the ocean areas (Gulf of Maine) differed from day-to-day, depending on the offshore flow types and transport patterns, i.e., good performance for marine or continental clear flows from the East/North/Northwest/South and southwest flows influenced only by Boston city plumes but overestimation for southeast flows influenced by the long-range transport plumes including both NYC and Boston. Time-series comparisons at the AIRMAP sites indicate that the model captured the hourly variations and broad synoptic changes in the observations of different gas species (O_3, NO_2, CO, NO_y, PAN) except NO and SO_2 at each site, although there were

occasional major excursions. The ε_N values of the model (5.2 to 6.4) and observation (8.5 to 10.7) at the CS and TF sites are close to the lower and higher bounds of the estimated ranges (5 to 10) of other investigators at rural sites in the eastern US, respectively. However, the modeled ε_N value is about 40 % lower than the observations. Since the majority of the eastern U. S. during the summer 2004 experienced unusually cool and wet conditions, the model performance presented here is probably not climatologically representative summertime, but is unique to the summer of 2004.

Acknowledgements

The authors would like to thank Drs. S.T Rao, R. Pinder and J. Godowitch for the constructive and very helpful comments that led to a substantial strengthening of the content of the paper. We are grateful to the 2004 ICARTT investigators for making their measurement data available. NOAA Environmental Technology Laboratory provided the Lidar O_3 and the High Resolution Doppler Lidar (HRDL) wind vertical profiles. The research presented here was performed under the Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number DW13921548. This work constitutes a contribution to the NOAA Air Quality Program. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

References

Angevine, W.M., C.J. Senff, A.B. White, E.J. Williams, J. Koermer, S.T.K. Miller, R. Talbot, P.E. Johnston, S.A. McKeen, and T. Downs (2004), Coastal Boundary Layer influence on pollutant transport in New England, *J. Appl. Meteor*., 43, 1425-1437.

Arnold, J.R., R.L. Dennis, and G.S. Tonnesen (2003), Diagnostic evaluation of numerical air quality models with specialized ambient observations: testing the Community Multiscale Air Quality modeling system (CMAQ) at selected SOS 95 ground sites. *Atmos. Environ*., 37, 1185-1198.

Byun, D.W. and J.K.S. Ching, (1999), Science algorithms of the EPA Models-3 Community Multi-scale Air Quality (CMAQ) modeling system, EPA/600/R-99/030, Office of Research and Development, U.S. Environmental Protection Agency.

Byun, D.W., and K.L. Schere (2006), Review of the governing equations, computational algorithms, and other components of the models-3 Community Multi-scale Air Quality (CMAQ) modeling system, *Applied Mechanics Reviews*, 59, 51-77.

Brewer, A. (2005), Doppler lidar measurements on the Ronald H. Brown. 2005 NOAA Research Vessel *Ronald H. Brown* Data Workshop, March 8-9 Boulder, CO.

Chin, M., D.J. Jacob, J.W. Munger, D.D. Parrish, and B.G. Doddridge (2004), Relationship Of Ozone And Carbon Monoxide Over North America. *J. Geophys. Res.*, 99, 14565-14573.

Cope, M.E., et al., (2004), The Australian air quality forecasting system. Part I: Project description and early outcomes. *J. Appl. Meteor*., 43, 649-662.

DeBell, L.J., M. Vozzella, R.W. Talbot, and J.E. Dobb (2004), Asian dust storm events of spring 2001 and associated pollutants observed in New England by the Atmospheric Investigation, Regional Modeling, Analysis and Prediction (AIRMAP) monitoring network. *J. Geophys. Res*., 109, D01304, doi:10.1029/2003JD003733.

DeMore, W.B., S.P. Sander, C.J. Howard, A.R. Ravishankara, D.M. Golden, C.E. Kolb, R.F. Hampson, M.J. Kurylo, and M.J. Molina (1997), Chemical kinetics and photochemical data for use in stratospheric modeling, Eval. 12, NASA Jet Propul. Lab., Calif. Inst. Of Technol., Pasadena.

Driscoll, C.T., et al. (2003), Nitrogen pollution in the Northeastern United States: Sources, effects, and management options, *BioScience*, 53, 357-374.

EPA (1991), Guideline for regulatory application of the urban airshed model. USEPA Report No. EPA-450/4-91-013. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

EPA (1999), Guideline for Developing an Ozone Forecasting Program, EPA-454/R-99-009; Research Triangle Park, NC.

Fiore, A.M., D.J. Jacob, I. Bey, R.M. Yantosca, B.D. Field, A.C. Fusco, and J.G. Wilkinson (2002), Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes, *J. Geophys. Res*., 107, NO. D15, 4275, 10.1029/2001JD000982,

Griffin, R.., C.A. Johnson, R.W. Talbot, H. Mao, R.S. Russo, Y. Zhou, and B.C. Sive (2004), Quantification of ozone formation metrics at Thompson Farm during the New England Air Quality Study (NEAQS) 2002, *J. Geophys. Res*., 109, D24302, doi:10.1029/2004JD005344.

Hanna, S.R., Z. Lu, H.C. Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, and D.A. Hansen (2001), Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmos. Environ*., 35, 891-903.

Jacob, D.J., L.W. Horowitz, J.W. Munger, B.G. Heikes, R.R. Dickerson, R.S. Artz, and W.C. Keene (1995), Seasonal transition for NO_x- to hydrocarbon-limited conditions for ozone production over the eastern United States in September. *J. Geophys. Res*., 100, D5, 9315-9324..

Kang, K., Eder, B.K., Stein, A.F., Grell, G.A., Peckham, S.E., and J. McHenry (2005), The new England air quality forecasting pilot program: development of an evaluation protocol and performance benchmark. *J. Air & Waste Manage. Assoc*., 55, 1782–1796

Kleinman, L. I., W.F. Ryan, P.H. Daum, S.R. Springston, Y.-N. Lee, L.J. Nunnermacker, and J. Weinstein-Lloyd (2004), An ozone episode in the Philadelphia metropolitan area, *J. Geophys. Res*., 109, D20302, doi:10.1029/2004JD004563.

Lin, C.Y., D.J. Jacob, J.W. Munger, and A.M. Fiore (2000), Increasing back-ground ozone in surface air over the United States, *Geophys. Res. Lett*., 27, 3465-3468.

Mao, H., and R. Talbot (2004), O3 and CO in New England: Temporal variations and relationships; *J. Geophys. Res.*, 109(D21304), doi:10.1029/2004JD004913.

Mathur et. al. (2004), Adaptation and application of the Community Multiscale Air Quality (CMAQ) modeling system for real-time air quality forecasting during the summer of 2004, Proc. 2004 Models-3 Conference, October18-20, 2004,

Mathur, R., et al. (2005), Multiscale air quality simulation platform (MAQSIP): Initial applications and performance for tropospheric ozone and particulate matter, *J. Geophys. Res*., 110, D13308, doi:10.1029/2004JD004918.

McHenry, J.N., W. F. Ryan, N. L. Seaman, C. J. Coats, J. Pudykeiwicz, S. Arunachalum and J.M. Vukovich (2004), A real-time Eulerian photochemical model forecast system, *Bull. Am Meteorol. Soc.*, 525-548.

Mckeen, S.A., et al. (2002), Ozone production from Canadian wildfires during June and July of 1995, *J. Geophys. Res.*, 107(D14), 4192, doi:10.1029/2001JD000697.

Mckeen, S.A., et al. (2005), Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004, *J. Geophys. Res.*, 110 (D21307), doi:10.1029/2005JD005858.

Olszyna, K.J., E.M. Bailey, R. Simonaitis, and J.F. Meagher (1994), O_3 and NO_y relationships at a rural site. *J. Geophys. Res.*, 99, 14,557-14,563.

Otte, T.L.,et al. (2005), Linking the Eta model with the Community Multiscale Air Quality (CMAQ) modeling system to build a national air quality forecasting system. *Wea. Forecasting*, 20, 367-384, 2005.

Pouliot, G.A. (2005), The emissions processing system for the Eta/CMAQ air quality forecast system, Proc. 7th Conf. on Atmospheric Chemistry, The 85th AMS Annual Meeting, Paper 4.5, Amer. Meteor. Soc., San Diego, CA.

Rogers, E., T. Black, D. Deaven, G. DiMego, Q. Zhao, M. Baldwin, N. Junker, and Y. Lin (1996), Changes to the operational "early" Eta Analysis/Forecast System at the National Centers for Environmental Prediction. *Wea. Forecasting*, 11, 391-413.

Thornton, J.A., P.J. Wooldridge, R.C. Cohen, M. Martinez, H. Harder, W.H. Brune, E.J. Williams, J.M. Roberts, F.C. Fehsenfeld, S.R. Hall, R.E. Shetter, B.P. Wert, and A. Fried (2002), Ozone productions rates as a function of NO_x abundances and HO_x production rates in the Nashville urban plume. *J. Geophys. Res.*, 107 (D12), doi:10.1029/2001JD000932.

Yu, S.C., P.S. Kasibhatla, D.L. Wright, S.E. Schwartz, R. McGraw, and A. Deng (2003) Moment-Based Simulation Of Microphysical Properties Of Sulfate Aerosols In The Eastern United States: Model Description, Evaluation And Regional Analysis, *J. Geophys. Res.*, 108(D12), 4353, Doi:10.1029/2002JD002890.

Yu, S.C., R. Mathur, D. Kang, K. Schere, B. Eder, and J. Pleim (2006), Performance and diagnostic evaluations of ozone estimations by the Eta-CMAQ air quality forecast system during the 2002 New England Air Quality Study (NEAQS). *J. Air & Waste Manage. Assoc.*, (in press)

Figure Captions

Fig. 1. Comparison of the modeled and observed daily maximum 8-hr O_3 concentrations at the AIRNow monitoring sites (a) Scatter plot (ppbv) (The 1:1, 1.5:1 and 1:1.5 lines are shown for reference) (b) Daily variation of domain-wide mean, residuals, NME, NMB and correlation (r), and spatial distributions of (c) NME and (d) NMB during July 1 and August 15, 2004.

Fig. 2. Tracks of (a) P-3, (b) DC-8 and (c) ship tracks and ozonsonde locations

Fig. 3. Comparison of vertical O_3 (ppbv) profiles for the model and observations from (a) P-3, (b) DC-8, (c) ship-Lidar and (d) ozonesonde during the ICARTT period.

Fig. 4. Comparison of vertical CO and HNO_3 (ppbv) profiles for the models and observations from P-3 (a, c), and DC-8 (b, d) during the ICARTT period.

Fig. 5. The same as Figure 4 but for SO_2 (ppbv), and HCHO (pptv).

Fig. 6. Comparison of daily vertical NO, NO_2, NO_y and NO_{y_sum} profiles for the models and observations from the aircrafts P-3 (a, c, d), and DC-8 (b) during the ICARTT period

Fig. 7. The same as Figure 4 but for water vapor and wind speed (WS).

Fig. 8. Time-series comparison of the modeled and observed JNO_2 along the P-3 tracks for each day.

Fig. 9. The same as Figure 8 but along the DC-8 tracks.

Fig. **10**. Visible satellite image (GOES-12) for 16:30 UTC, 15 July, 2004, with DC-8 track overlaid, and the height along the flight track is shown in the embedded figure.

Fig. **11.** Time-series and scatter plots (the 2:1, 1:1 and 1:2 lines are shown for reference) of model predictions and observations for different species and meteorological parameters on the basis of ship measurements.

Fig. **12**. On 7/20/2004 for ship, (a) Time-series of modeled and observed O_3, (b) ship tracks on 7/20, 7/30 and 8/10, and three flow directions, (c) The model simulation results for O_3 concentration (ppb) with AQS observed data overlaid (diamond) on 15:00 and 22:00 UTC (7/20/2004), and (d) Vertical profiles of model and ozonesonde for O_3 at Ron Brown ship (15:30 UTC) and Narragansett, RI (18:40 UTC)

Fig. **13.** On 7/30/2004 for ship, (a) Time-series of modeled and observed O_3, (b) The model simulation results for O_3 concentration (ppb) with AQS observed data overlaid (diamond) on 15:00 and 22:00 UTC, and (c) Vertical profiles of model and ozonesonde for O_3 at Ron Brown ship (15:30 UTC) and Narragansett, RI (18:40 UTC)

Fig. **14**. The same as Figure 13 but for 8/10/2004.

Fig. 15. Time-series and scatter plots of model predictions and observations for each parameter at the Castle Springs site

Fig. 16. O_3 as a function of NO_z for the NO_x-limited conditions indicated by the observational data with $[O_3]/[NO_x]>46$ at (a) Castle Springs (CS), (b) Thompson Farm (TF), and (c) along ship tracks. Right panels are scatter plots of modeled and observed NO_z.

Table 1. Operational evaluation for O_3 concentrations on the basis of the AQS data over the eastern U.S.

	Number	Domain Mean (ppb) Obs	Domain Mean (ppb) Model	RMSE (ppbv)	MB (ppbv)	NMB (%)	NME (%)	R
All data								
Hourly	1170000	28.1	39.7	19.4	11.5	40.9	54.8	0.64
Max 1-hr	40189	51.9	60.4	16.9	8.5	16.4	25.3	0.61
Max 8-hr	40189	45.7	56.1	16.6	10.4	22.6	28.8	0.60
For >40 ppbv								
Hourly	297981	52.8	56.0	14.9	3.2	6.1	21.5	0.42
Max 1-hr	24943	61.9	66.2	14.8	4.3	7.0	18.2	0.47
Max 8-hr	24943	54.9	61.5	13.9	6.5	11.9	19.7	0.45

Table 2. Flight Observation Summary for WP-3 aircraft

Date	Observation summary*
7/9	Shortly after takeoff at 15:35 UTC, the P-3 encountered the fresh Boston plume, which was not well mixed and confined to low altitudes, above the ocean close to Cape Ann with the CO and SO_2 concentrations as high as 200ppbv and 4 ppbv, respectively.
7/11	On 7/11 night, after takeoff at 23:00 UTC, the P-3 encountered the Boston plume as the aircraft turned to the south flying to the west of Boston (see Figure 2) and intercepted a significant biomass-burning layer at ~2700 meter in the later (~24:00 UTC) flight to southeast of Boston during a spiral climb.
7/15	A very well defined NY city plume was sampled extensively by flight by making cross-sections and spirals through the plume after ~18:30 UTC for understanding of ozone evolution in an urban plume.
7/20	The aircraft encountered a fresh NY city plume immediately downwind of NY, which reached up to ~1800 m, and the biomass-burning plume at ~3000 m between 17:00 and 18:00 UTC. The aging NY plume was intercepted over the Gulf of Maine with CO reaching 270 ppbv.
7/21	An aged (1.5 to 2.5 days old) NY city plume over the Gulf of Maine was intercepted in the lower troposphere (~270 to ~1000 m) in the outbound northeasterly flight between 14:30 and 15:30 UTC. A biomass-burning plume above Cape Cod at ~3000 m was encountered at 20:00 UTC.
7/22	A progressively more aged NY city plume from 7/21 over the Bay of Fundy reaching beyond Cape Breton and Prince Edward Island was sampled.
7/25	The flight looked at the outflow of Sunday emissions from Boston and NY cities and the downwind plume with high SO_2 (>4 ppb) for Montour power plant in central Pennsylvania in the NW of the power plant during 16:00 and 17:30 UTC. There was a widespread signature of biomass burning.
7/27	The flight aimed to look at the pollution accumulation ahead of the cold front. The pollution upwind and downwind of the Washington and Baltimore metropolitan area was sampled by P-3 during 17:30 to 18:30 UTC with very high SO2 (>5 ppb), CO (>180 ppb) concentrations.
7/28	The conveyor belt outflow without a clear signature of the anthropogenic pollution export was sampled with low concentrations for all species as there was a stationary front. There was a biomass burning plume over Quebec beyond of the model domain.
7/31	On 7/31 night, the P-3 encountered the NYC plume at ~ 8/1 1:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/3	On 8/3 night, the P-3 encountered the NYC plume over southwestern Connecticut at ~700m during 4:00 to 4:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/6	The P-3 encountered the plume of Ohio valley power plants at ~1000m during 15:30 and 20:30 UTC with high SO_2 (>5 ppbv) and low O_3 (<60 ppbv) concentrations.
8/7	On 8/7 night, the P-3 encountered the NYC and Boston plumes at ~700m during 8/8 1:00 to 4:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/9	On 8/9 night, the P-3 encountered the plume of Ohio valley power plants at ~1000m during 8/10 0:30 and 3:30 UTC the NYC and Boston plumes at ~700m during 8/10 4:30 to 6:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/11	On 8/11 night, the P-3 encountered NYC plume at ~700m during 2:30 to 10:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/14	It was a cloudy day across the whole eastern U.S under the influence of Hurricane Charley. The P-3 encountered NYC plume at ~200m during 16:30 to 17:30 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).
8/15	It was still cloudy along eastern coast. The P-3 encountered Atlanta plume at ~700m during 18:20 to 20:00 UTC with high SO_2 (>5 ppb), and CO (>180 ppb).

* based on flight information presented at www.al.noaa.gov/ICARTT/fieldoperations/fomp.shtml)

Table 3. Flight Observation summary for DC-8 aircraft

Date	Observation summary*
7/15/2004	Characterization of Asian pollution, Alaskan fires, and anthropogenic pollution, The flight occurred above the cloud at ~8 km during 17:30 and 20:00 UTC in the NW of Boston city.
7/18/2004	Characterization of North American pollution outflow, possible characterization of Alaskan fires, and a flyby over the NOAA ship Ron Brown in the NE of Boston city.
7/20/2004	Characterization of smoke from Alaskan fires transported over the US, boundary layer pollution over the southeast and mid- west. There were some scattered clouds.
7/22/2004	Sampling polluted boundary layer outflow along the eastern seaboard both to the north and south of Pease. Intercomparison between the NASA DC-8 and NOAA WP-3D aircraft.
7/25/2004	Convective outflow from southeast US, map Ohio River Valley emissions in northerly flow under flight. The DC-8 flew above the clouds at 8 km during18:30 to 19:30 UTC.
7/28/2004	Sample the structure and chemical evolution of the US continental outflow out over the Atlantic Ocean. Most of time was beyond the model domain
7/31/2004	Aged air sampling/recirculation, low level outflow, P-3 intercomparison, and possible Asian influences
8/02/2004	Sample low level North American outflow and aged air pollution aloft, conduct a flyby over the ground Appledore Island air quality station.
8/06/2004	Flew over the Ohio River Valley. The DC-8 flew above the clouds at 10 km during12:30 to 13:30 UTC but below the cloud at 200m during 19:30 and 20:00 UTC.
8/07/2004	Sample North American outflow, a stratospheric intrusion, and perform P-3 intercomparison.
8/11/2004	NA outflow and WCB lifting, frontal crossing and low level pollution
8/13/2004	Outflow from major industrial cities (Houston, New Orleans) with clear skies for most of time except the period of 21:30 to 22:00 UTC.
8/14/2004	Flight above the cloud over Missouri-Kansas during 19:00 and 20:00 UTC.

* based on flight information presented at www.al.noaa.gov/ICARTT/fieldoperations/fomp.shtml)

Table 4. Summary of wind fields observed by High Resolution Doppler Lidar (HRDL) on the Ron Brown Ship and model performance for O_3 during the 2004 ICARTT period.

Day (2004)	Flow types**	Model performance for O_3
	Days with offshore flows (South-Westerly/Westerly):	
July 10*	Wly-NWly all levels, LLJ (4-6 Z), BL wave (12-21 Z)	Over prediction
July 15*	NEly winds shifting to Sly, UL Easterly in BL, SWly, LLJ (22-24 Z)	Over prediction
July 20	SSW to Sly flow in BL, Wly above	Over prediction
July 21*	SWly winds shifting to Nly to Ely to Sly in BL, LLJ (0-10 Z), BL wave (14-17 Z)	Over prediction
July 22*	SSWly for low level winds, shifts from Wly to Ely to SWly >1km, LLJ (21-24 Z)	Over prediction
July 23*	SSW to Sly flow, LLJ events (0-15 Z)	Over prediction
July 29*	Wly all levels then shift to Nly to Sly at surface, LLJ (23 Z), BL waves	Over prediction
July 30*	Wly and SWly flow all levels, LLJ (0-3 & 20-24 Z)	Over prediction
July 31*	WSWly flow all levels, LLJ throughout, vels up to 20 m/s	Over prediction
Aug 1*	W-SWly flow 15-20 m/s, LLJ (6-9Z)	Over prediction
Aug 3	SWly winds shifting to NEly, BL wave	Over prediction
Aug 12	SWly-SSWly flow all levels	Over prediction
July 16*, 17*	sustained WSWly flow for 2.5 days, BL & LLJ events throughout	Very good
Aug 4	Wly in BL, NW above	Good
Aug 8*, 9*, 10*, 11*	W-SWly flow for 4 days, LLJ's throughout & hi velocity shear	Very good
	Days with clear marine and continental flows (Easterly/Northerly/Northwesterly):	
July 11*	Wly winds shifting to Nly to Ely, LLJ (3-6 Z)	Very good
July 12*	Ely winds shifting to Sly to SWly, LLJ (6 Z & 23 Z)	Very good
July 13*	SSWly winds shifting to Ely, LLJ (3-4 Z, 9-11 Z & 23 Z)	Very good
July 18	SWly winds shifting to SEly	good
July 25, 26, 27	Light predominantly Ely flow <4 m/s throughout	Very good
July 28*	Nly winds shifting to Ely at 20 Z and then to WNWly, LLJ (11-Z)	Very good
Aug 5	Wly & NWly 0-100 m, Nly 100-1000 m	Very good
Aug 6, 7*	Wly winds shifting to Ely then back via N	good

* Nocturnal Low-level jet (LLJ), **obtained from Brewer (2005).

Table 5. Statistical summaries of the comparisons of the model results with the observations at the different sites during the 2004 ICARTT period (7/1 to 8/15, 2004). r is correlation coefficient between the model predictions and observations.

Parameters	<C>* Obs	<C>* Model	r	% within a factor of 1.5**	% within a factor of 2**
Castle Springs (N=1047)					
O_3	35.17	43.63	0.493	66.6	90.1
NO	0.14	0.05	0.222	12.1	22.5
CO	188.84	108.78	0.706	19.3	74.7
NO_Y	2.27	3.14	0.587	43.6	67.7
SO_2	1.16	0.87	0.388	29.6	45.8
JNO_2 (1/s)	3.18×10^{-3}	4.07×10^{-3}	0.820	49.6	63.4
Temperature (C)	19.65	19.78	0.867	100.0	100.0
RH (%)	78.69	71.64	0.781	97.7	100.0
Isle of Schoals (N=1078)					
O_3	36.68	52.31	0.541	56.9	80.2
CO	171.70	121.15	0.610	60.9	90.3
NO	0.76	0.18	0.448	0.8	3.5
Mount Washington (N=1076)					
O_3	45.87	45.85	0.554	87.7	98.8
NO	3.64	0.01	-0.054	8.9	13.8
CO	152.43	95.19	0.301	46.7	84.3
NO_Y	4.04	2.23	-0.060	20.6	38.4
SO_2	0.74	0.30	-0.001	19.0	32.6
JNO_2 (1/s)	3.59×10^{-3}	4.43×10^{-3}	0.768	43.1	61.9
Thompson Farm (N=1067)					
O_3	28.80	41.68	0.751	48.1	73.8
NO	0.33	0.29	0.436	31.3	51.3
CO	173.07	154.66	0.593	77.7	98.5
NO_Y	3.93	7.26	0.321	28.8	51.6
SO_2	1.22	1.63	0.084	14.3	25.3
JNO_2 (1/s)	3.19×10^{-3}	3.90×10^{-3}	0.865	53.8	68.1
Temperature (C)	20.33	20.44	0.887	99.9	100.0
RH (%)	80.97	75.18	0.829	98.5	100.0

* <C> is the mean concentration (ppb)

** Percentages (%): are the percentages of the comparison points at which model results are within a factor of 1.5 and 2 of the observations. N is number of samples.

Table 6. Statistical summary of number of hours for response surface indicator ratios (O_3/NO_x) for model and observations at the CS, WMO and TF sites for all days (obs-limited hours) during the period of July 1 to August 15, 2004. The values in parentheses are the percentages (%).

	Castle Springs		Thompson Farm		Ron Brown ship	
O_3/NO_x	Obs	Model	Obs	Model	Obs	Model
0-14	32 (7)	18 (4)	181 (38)	105 (22)	106 (35)	49 (16)
15-25	34 (7)	19 (4)	51 (11)	72 (15)	58 (19)	21 (7)
26-45	94 (20)	18 (4)	59 (12)	125 (26)	46 (15)	30 (10)
>46	312 (66)	417 (88)	188 (39)	177 (37)	93 (31)	103 (34)
Total hours	472 (100)	472 (100)	479 (100)	479 (100)	303 (100)	303 (100)

Table 7. Correlations between O_3 and NO_z for the NO_x-limited conditions indicated by the observational data with [O3]/[NOx]>46 (aged air masses) at the CS, WMO and TF sites during the period of July 1 to August 15, 2004. N is number of points and r is correlation coefficient.

Sites	Regression equations
Castle Springs (N=312)	Obs: $[O_3]=10.7[NO_z]+22.8$, r = 0.838
	Model: $[O_3]=6.4[NO_z]+30.1$, r = 0.784
Thompson Farm (N=188)	Obs: $[O_3]=8.5[NO_z]+26.4$, r = 0.896
	Model: $[O_3]=5.2[NO_z]+34.0$, r = 0.911
Ship over the ocean (N=93)	Obs: $[O_3]=11.7[NO_z]+35.4$, r = 0.608
	Model: $[O_3]=3.6[NO_z]+38.7$, r = 0.833

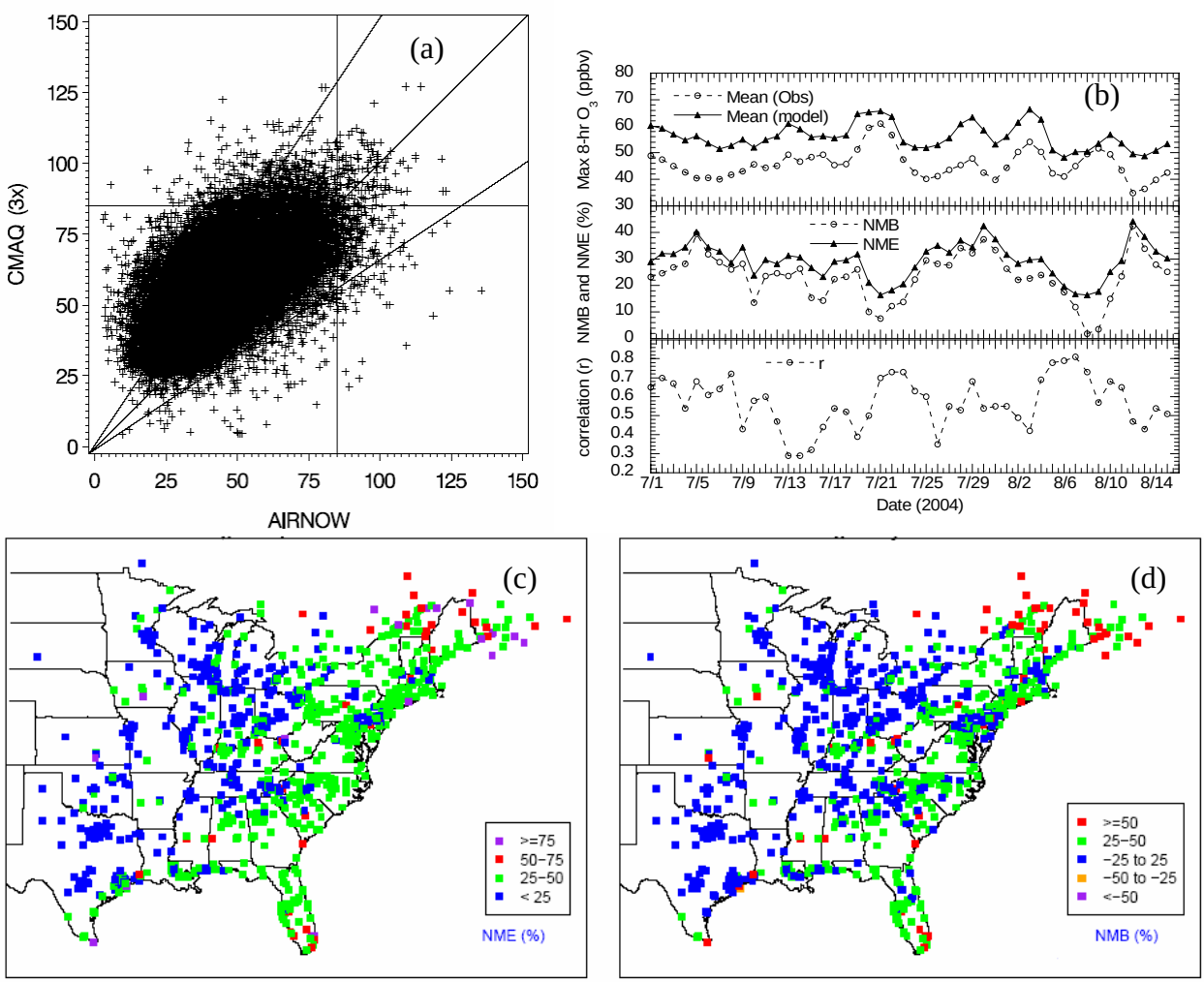
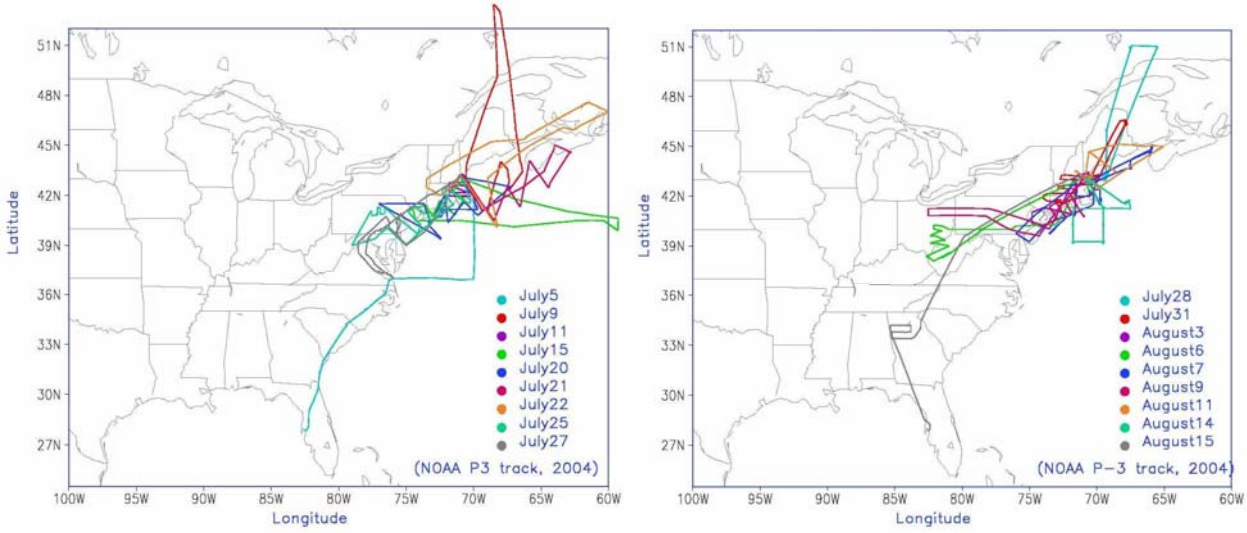
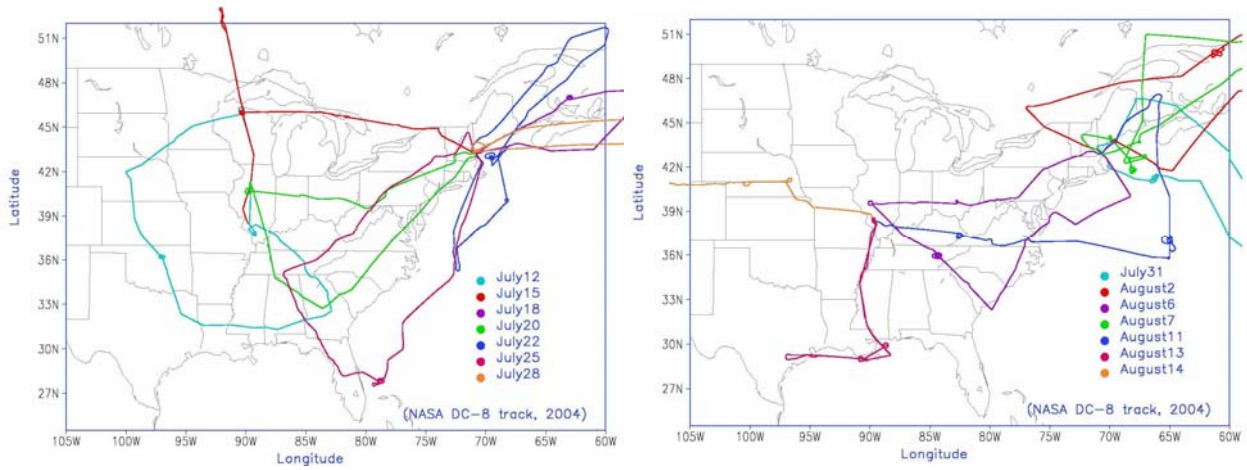


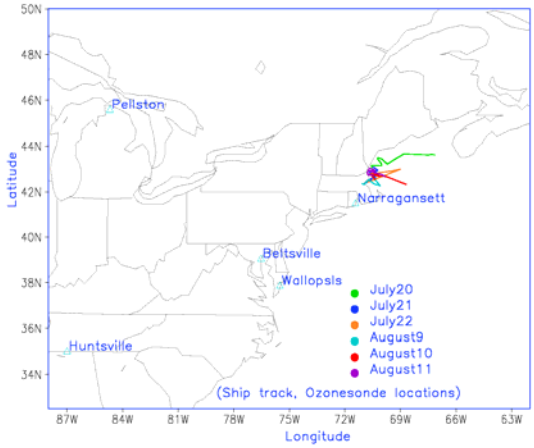
Figure 1. Comparison of the modeled and observed daily maximum 8-hr O_3 concentrations at the AIRNow monitoring sites (a) Scatter plot (ppbv) (The 1:1, 1.5:1 and 1:1.5 lines are shown for reference) (b) Daily variation of domain-wide mean, residuals, NME, NMB and correlation (r), and spatial distributions of (c) NME and (d) NMB during July 1 and August 15, 2004.



(a)



(b)



(c)

Fig. 2. Tracks of (a) P-3, (b) DC-8 and (c) ship tracks and ozonesonde locations

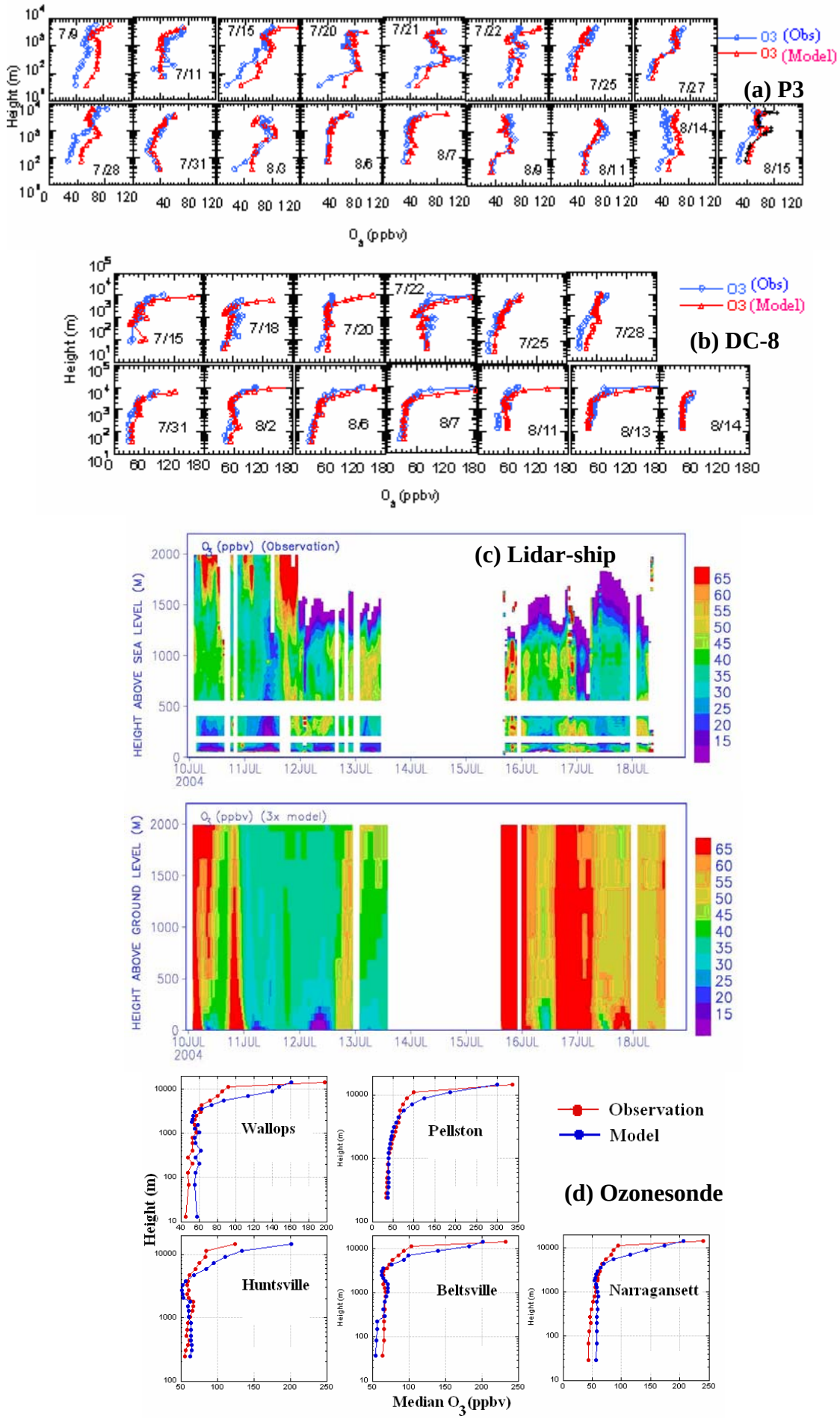


Fig. 3. Comparisons of daily vertical O_3 profiles for the models and observations from the aircraft (a) P-3, (b) DC-8, (c) ship-Lidar (Time is UTC), (d) Ozonesonde during the ICARTT period.

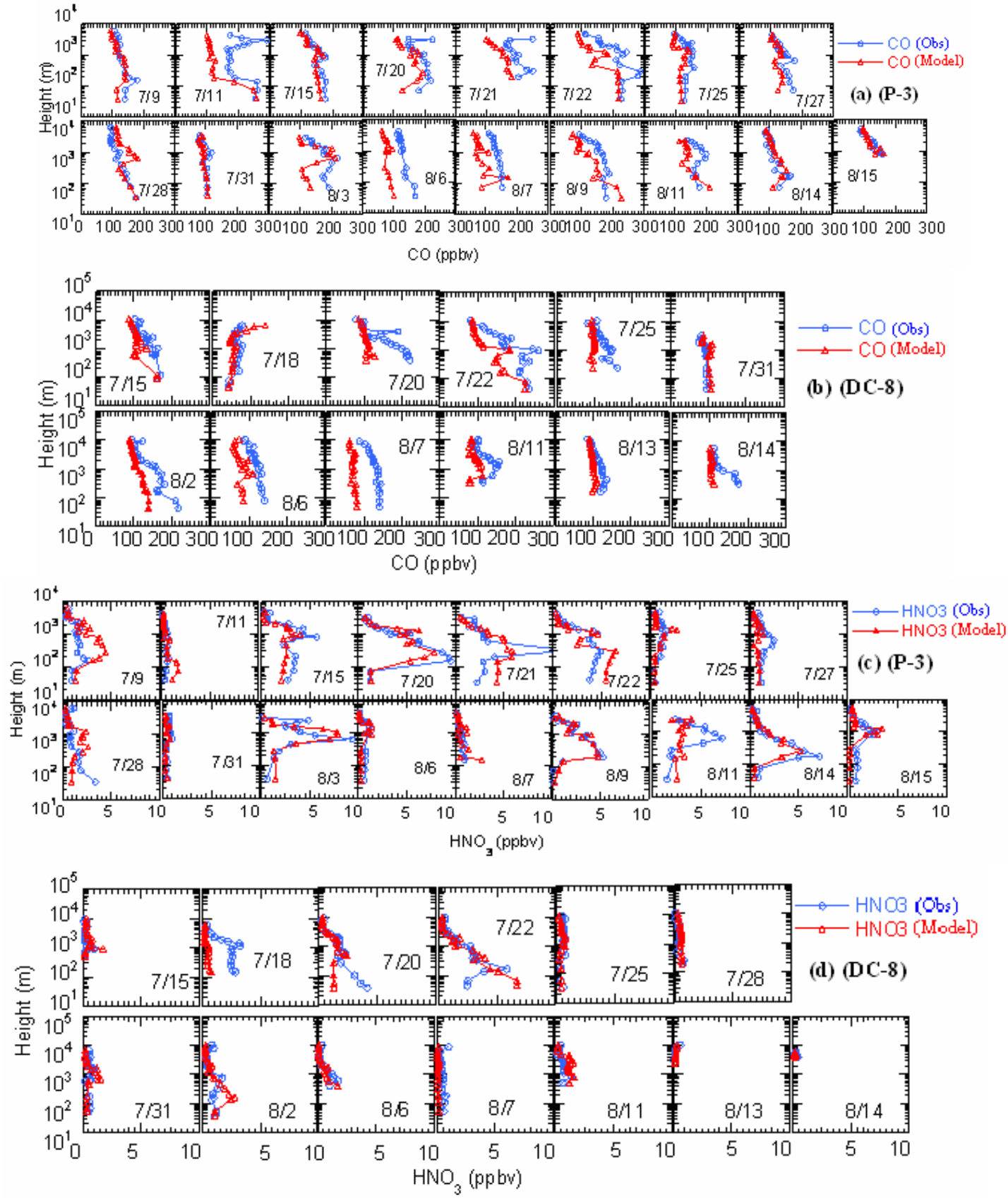


Fig. 4. Comparison of daily vertical CO and HNO_3 profiles for the models and observations from the aircrafts P-3 (a, c), and DC-8 (b, d) during the ICARTT period.

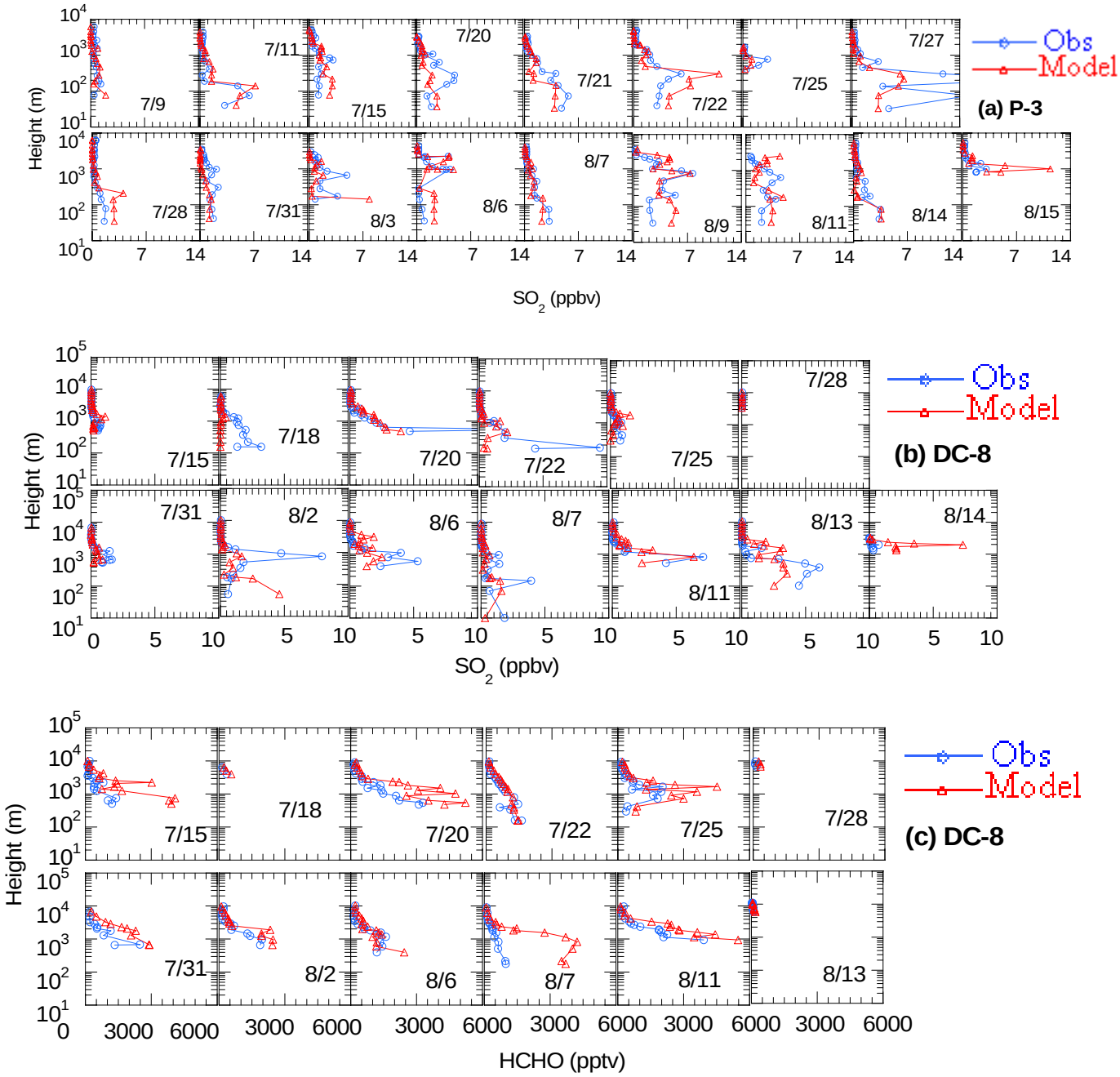


Fig. 5. Comparison of daily vertical SO_2 and HCHO profiles for the models and observations from the aircrafts P-3 (a), and DC-8 (b, and c) during the ICARTT period.

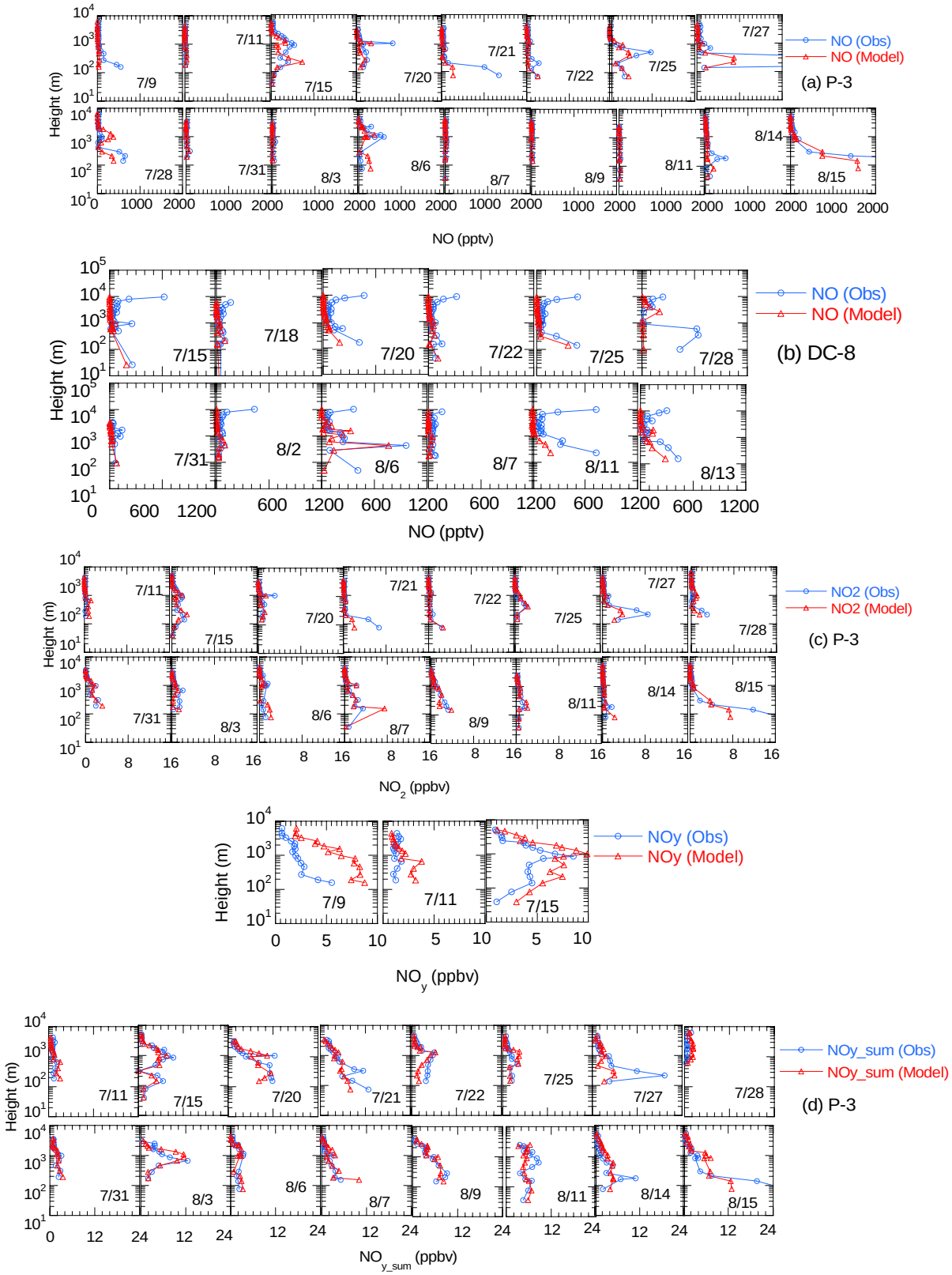


Fig. 6. Comparison of daily vertical NO, NO_2, NO_y and NO_{y_sum} profiles for the models and observations from the aircrafts P-3 (a, c, d), and DC-8 (b) during the ICARTT period.

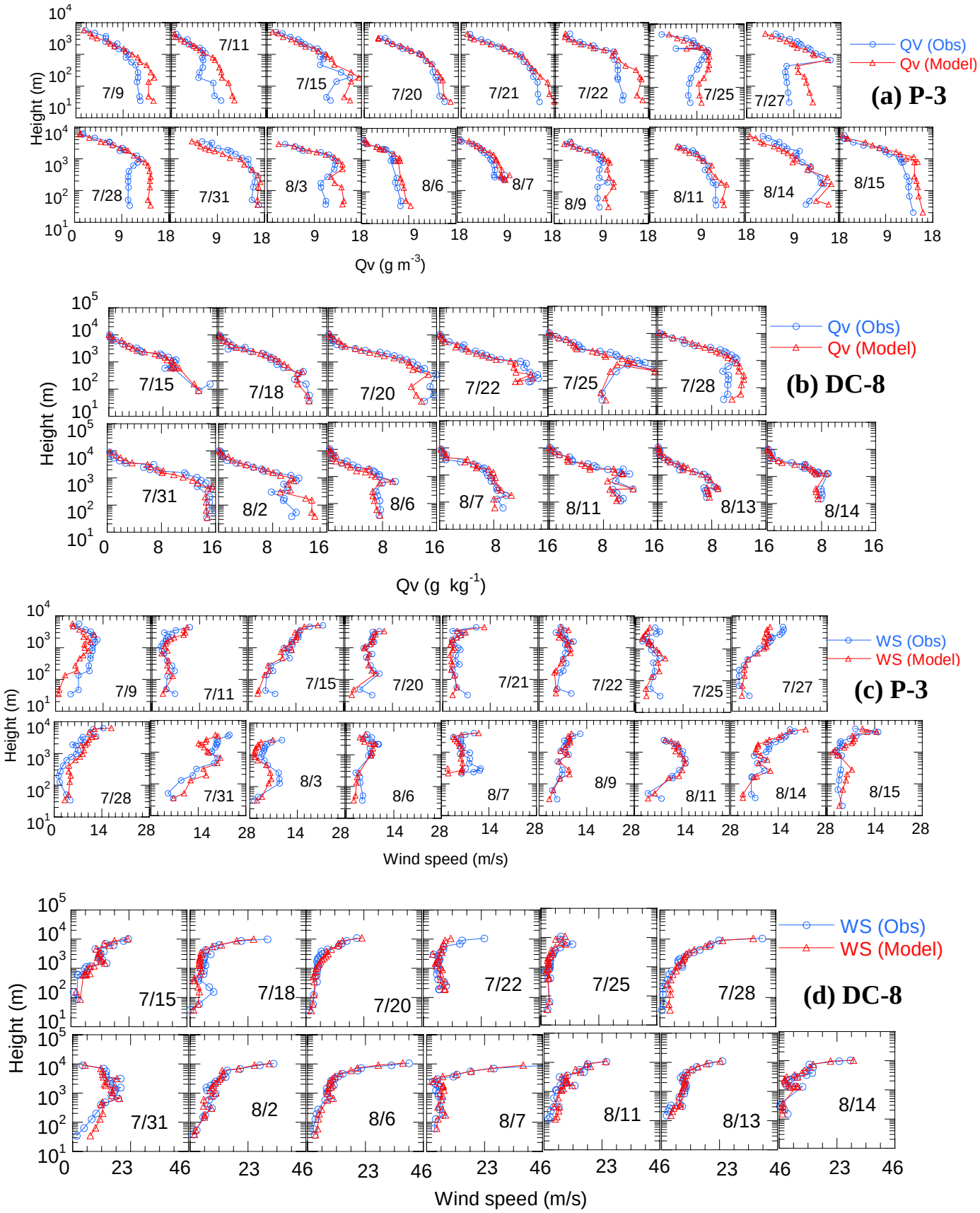


Fig. 7. Comparison of daily vertical water vapor and wind speed profiles for the models and observations from the aircrafts P-3 (a), and DC-8 (b, c, d) during the ICARTT period.

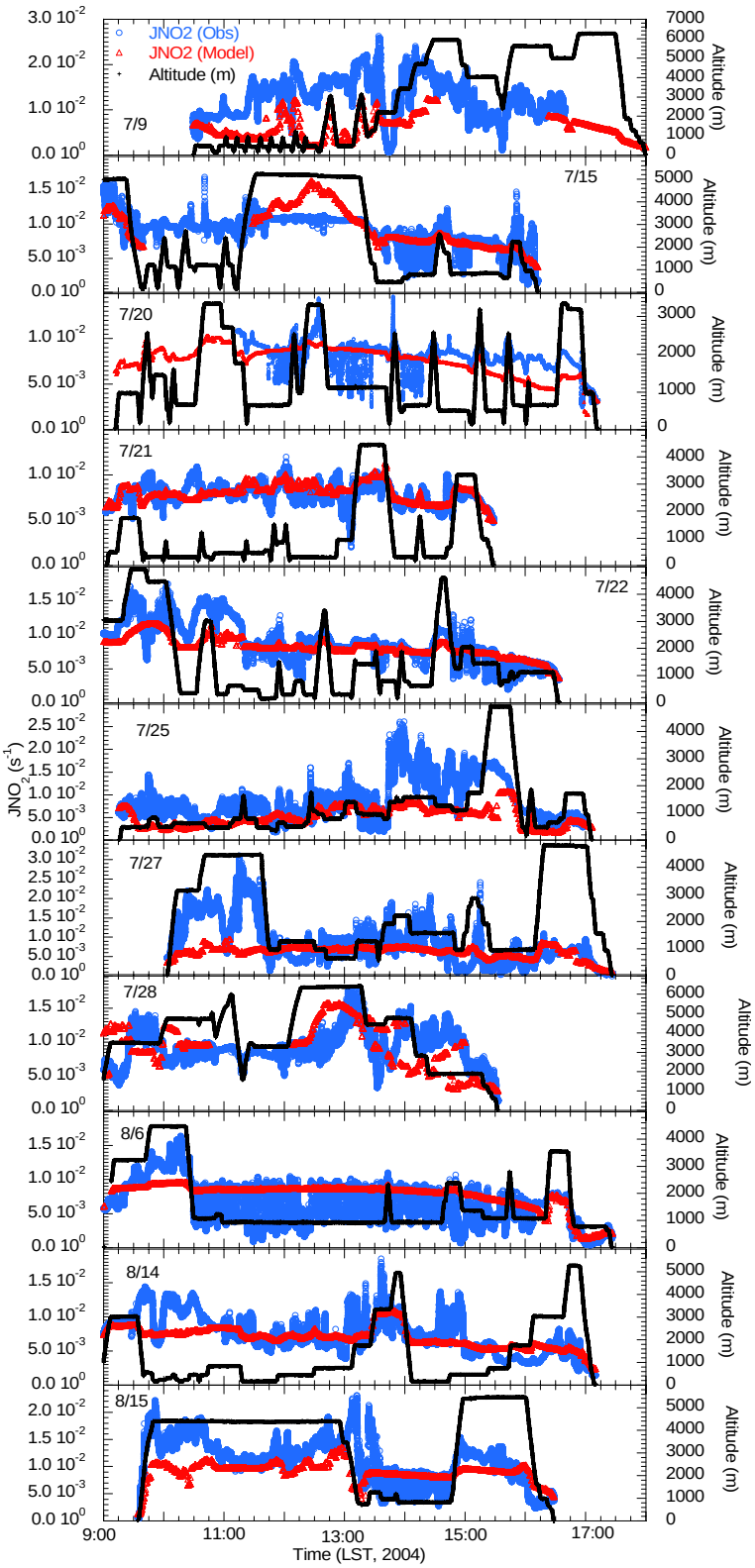

JNO2 (Obs)
JNO2 (Model)
Altitude (m)
7/9
7/15
7/20
7/21
7/22
7/25
7/27
7/28
8/6
8/14
8/15
JNO2 (s-1)
Altitude (m)
Time (LST, 2004)
9:00
11:00
13:00
15:00
17:00

Figure 8. (P-3)

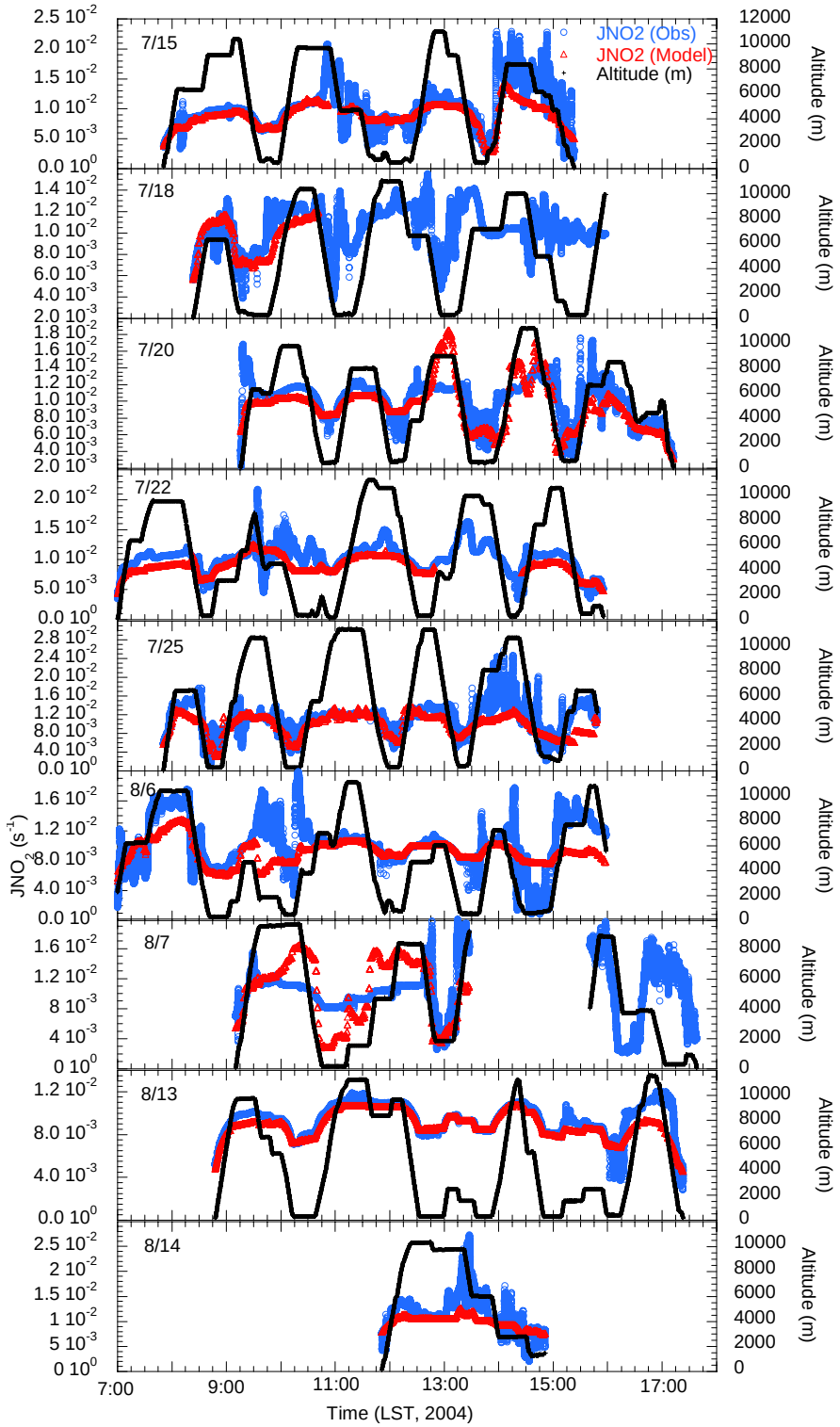

7/15
7/18
7/20
7/22
7/25
8/6
8/7
8/13
8/14
JNO2 (Obs)
JNO2 (Model)
Altitude (m)
JNO2 (s-1)
Time (LST, 2004)
7:00
9:00
11:00
13:00
15:00
17:00

Figure 9. (DC-8)

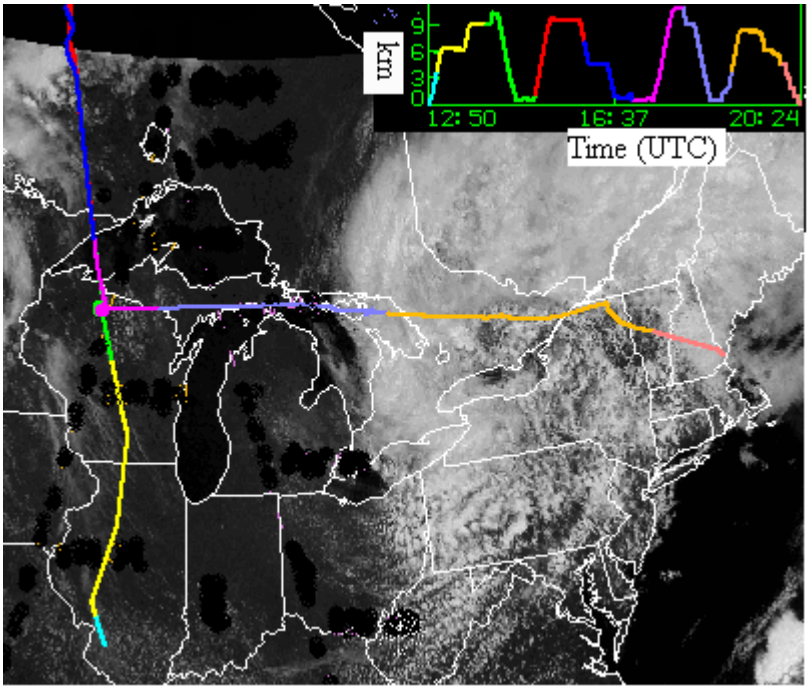


Figure 10. Visible satellite image (GOES-12) for 16:30 Z, 15 July, 2004, with DC-8 track overlaid, and the height along the flight track is shown in the embedded figure.

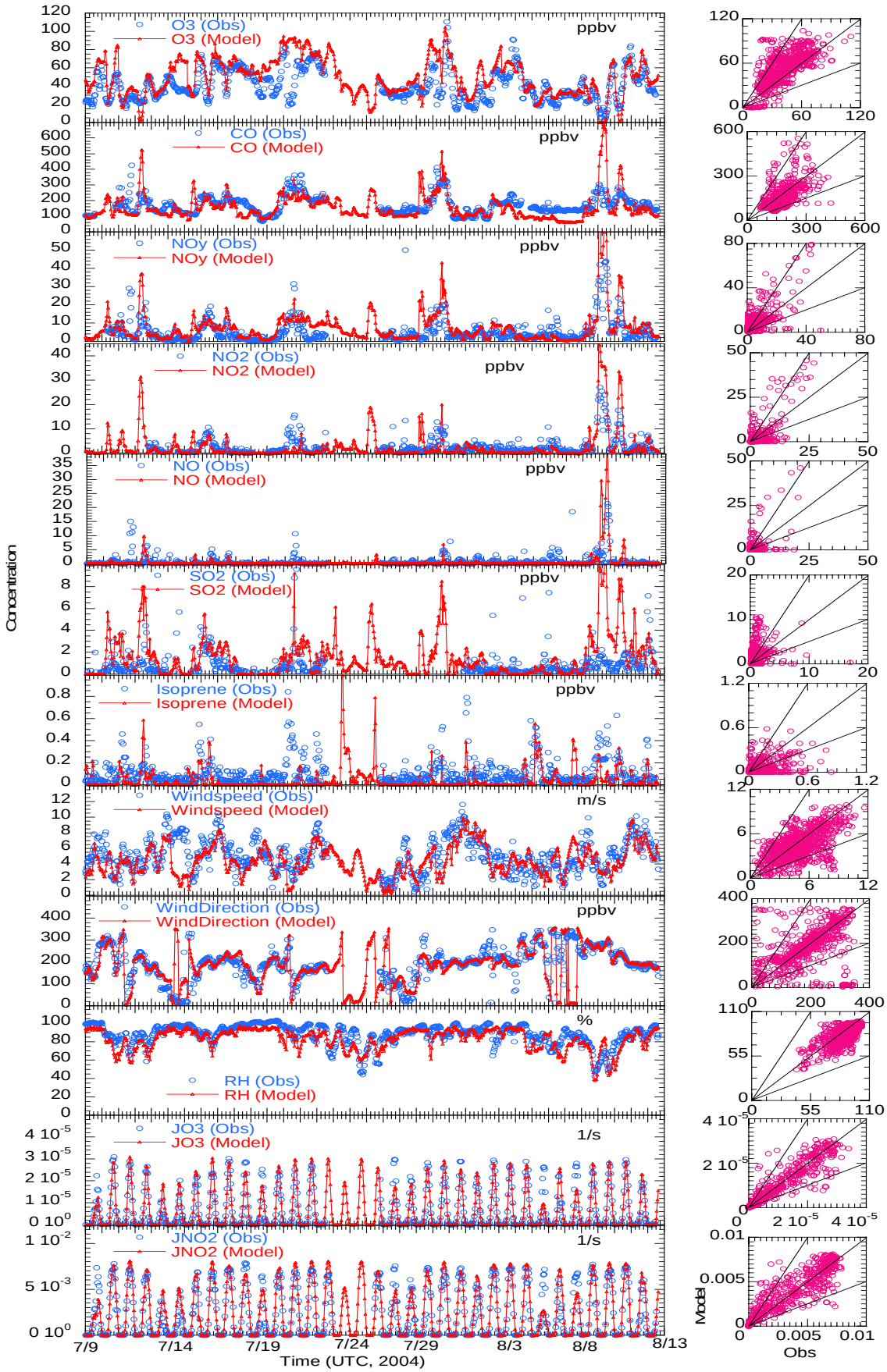


Figure 11. Time-series and scatter plots (the 2:1, 1:1 and 1:2 lines are shown for reference) of model predictions and observations for different species and meteorological parameters on the basis of ship measurements.

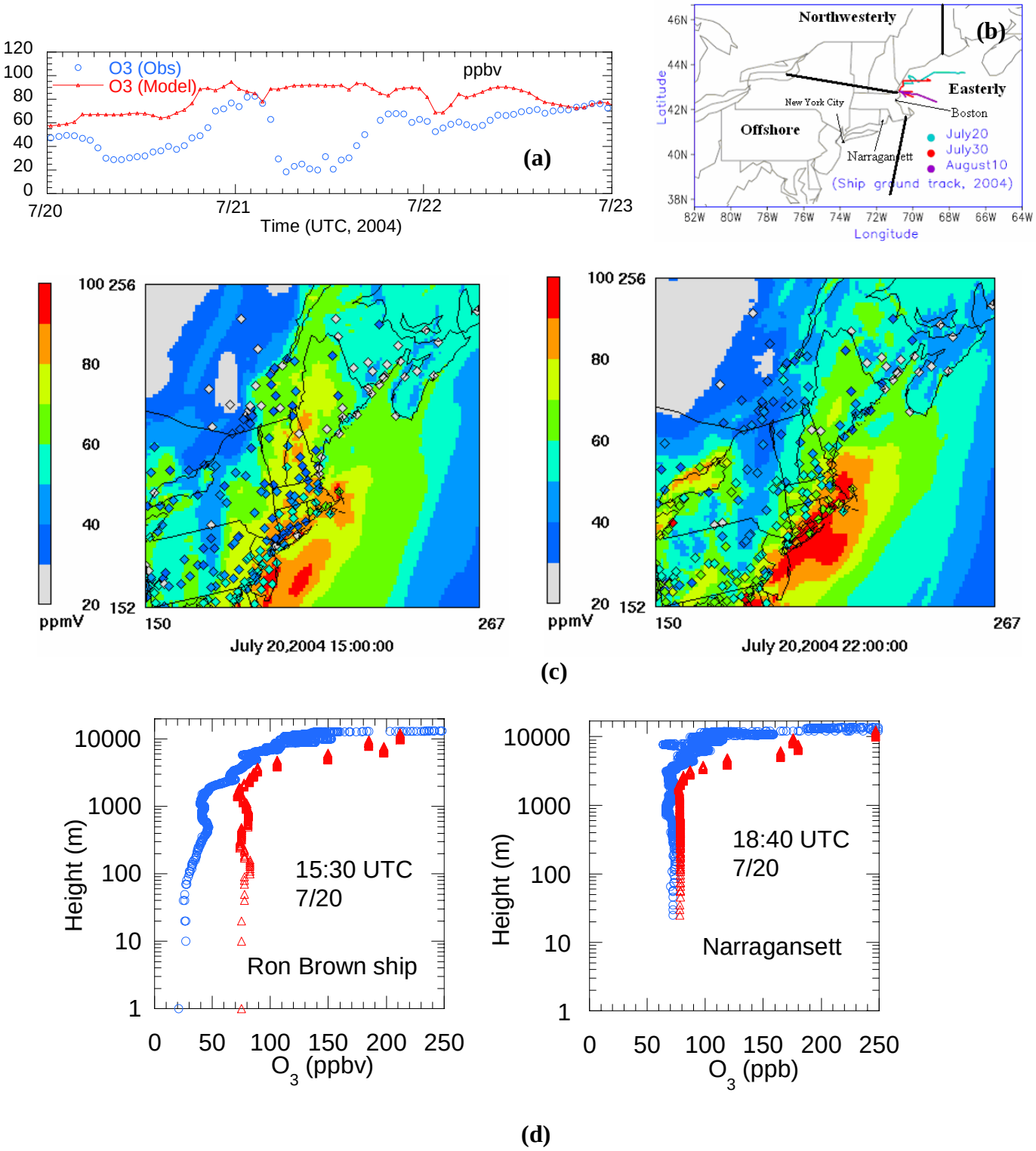

O3 (Obs)
O3 (Model)
ppbv
(a)
Time (UTC, 2004)
7/20
7/21
7/22
7/23
Northwesterly
(b)
Easterly
Offshore
New York City
Boston
Narragansett
July20
July30
August10
(Ship ground track, 2004)
Latitude
Longitude
ppmV
July 20,2004 15:00:00
July 20,2004 22:00:00
(c)
Height (m)
15:30 UTC
7/20
Ron Brown ship
18:40 UTC
7/20
Narragansett
O3 (ppbv)
O3 (ppb)
(d)

Figure 12

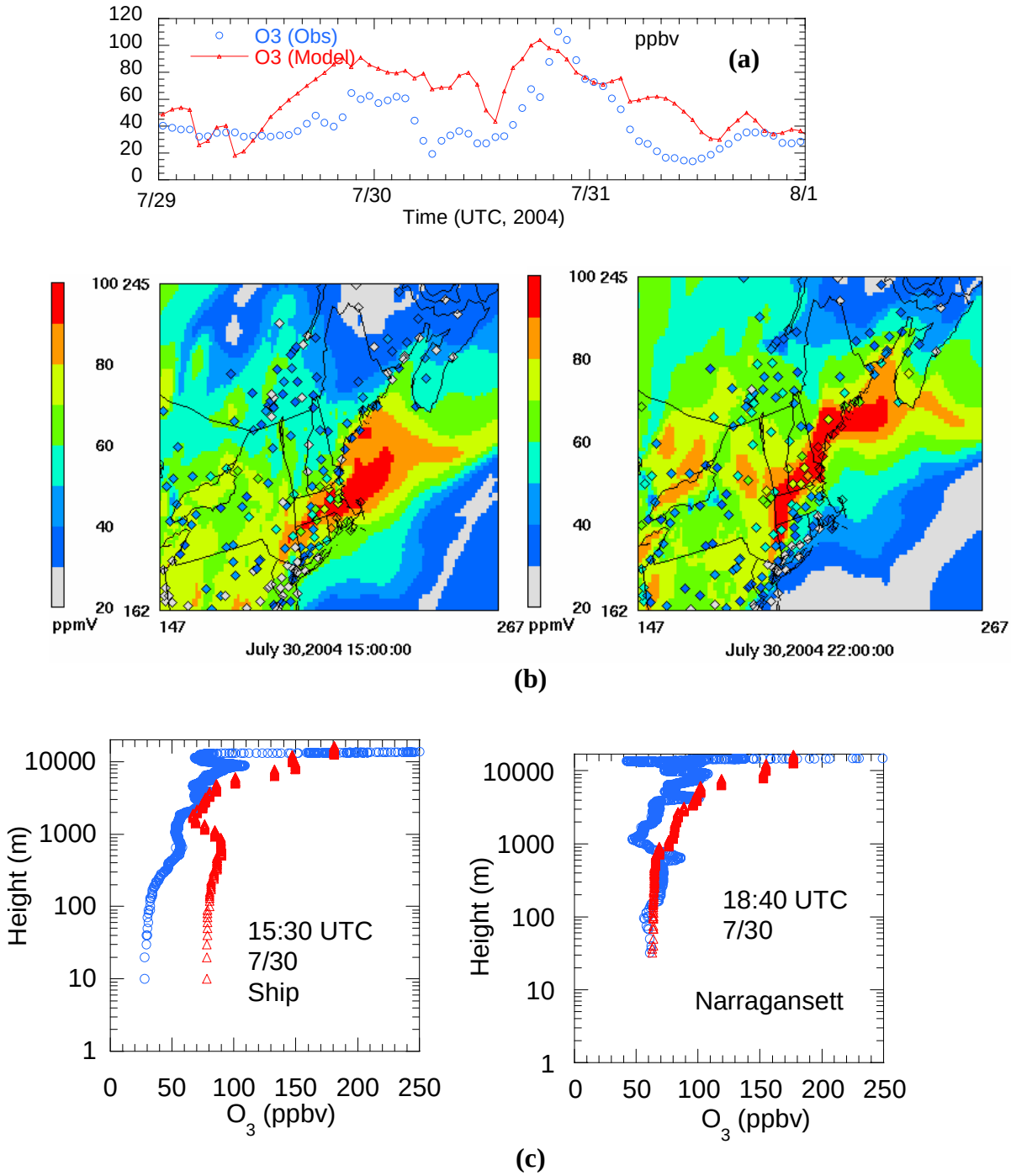
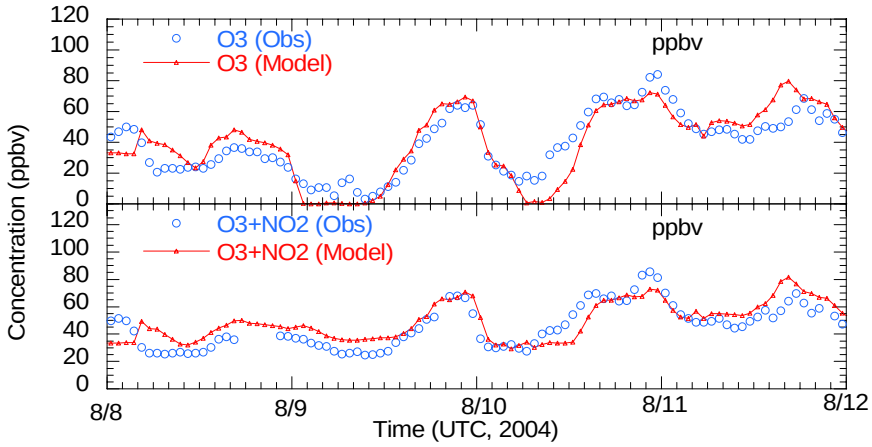
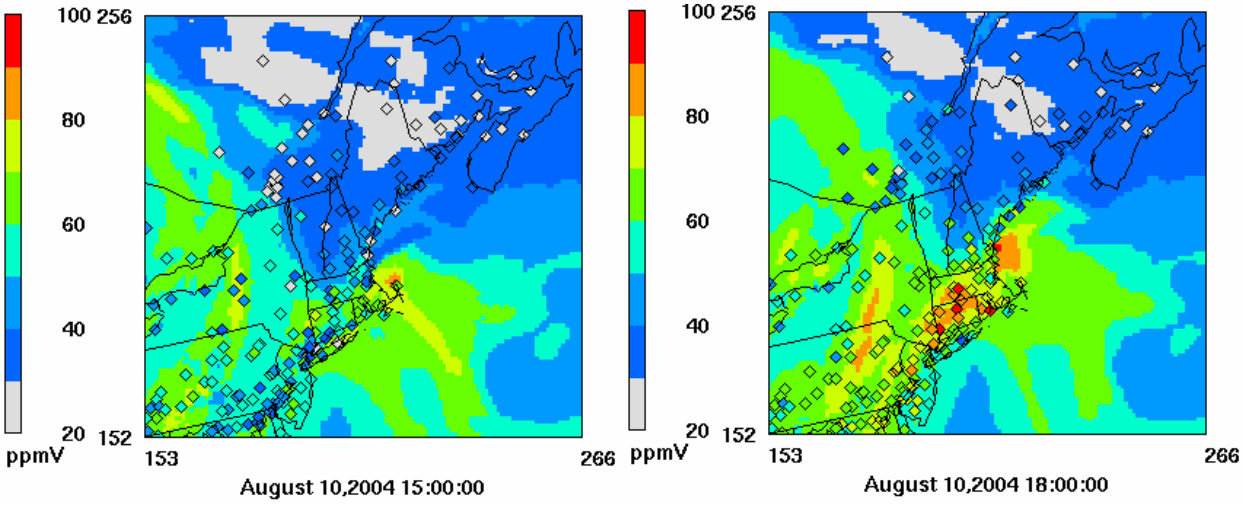

O3 (Obs)
O3 (Model)
ppbv
(a)
Time (UTC, 2004)
7/29
7/30
7/31
8/1
ppmV
July 30,2004 15:00:00
July 30,2004 22:00:00
(b)
Height (m)
15:30 UTC
7/30
Ship
18:40 UTC
7/30
Narragansett
O3 (ppbv)
(c)

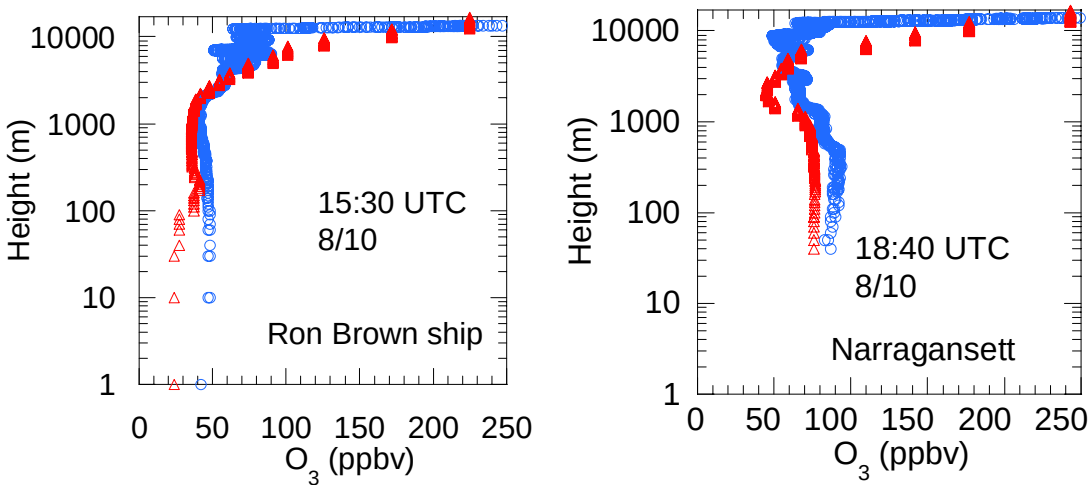
Figure 13.



(a)



(b)



(c)

Figure 14.

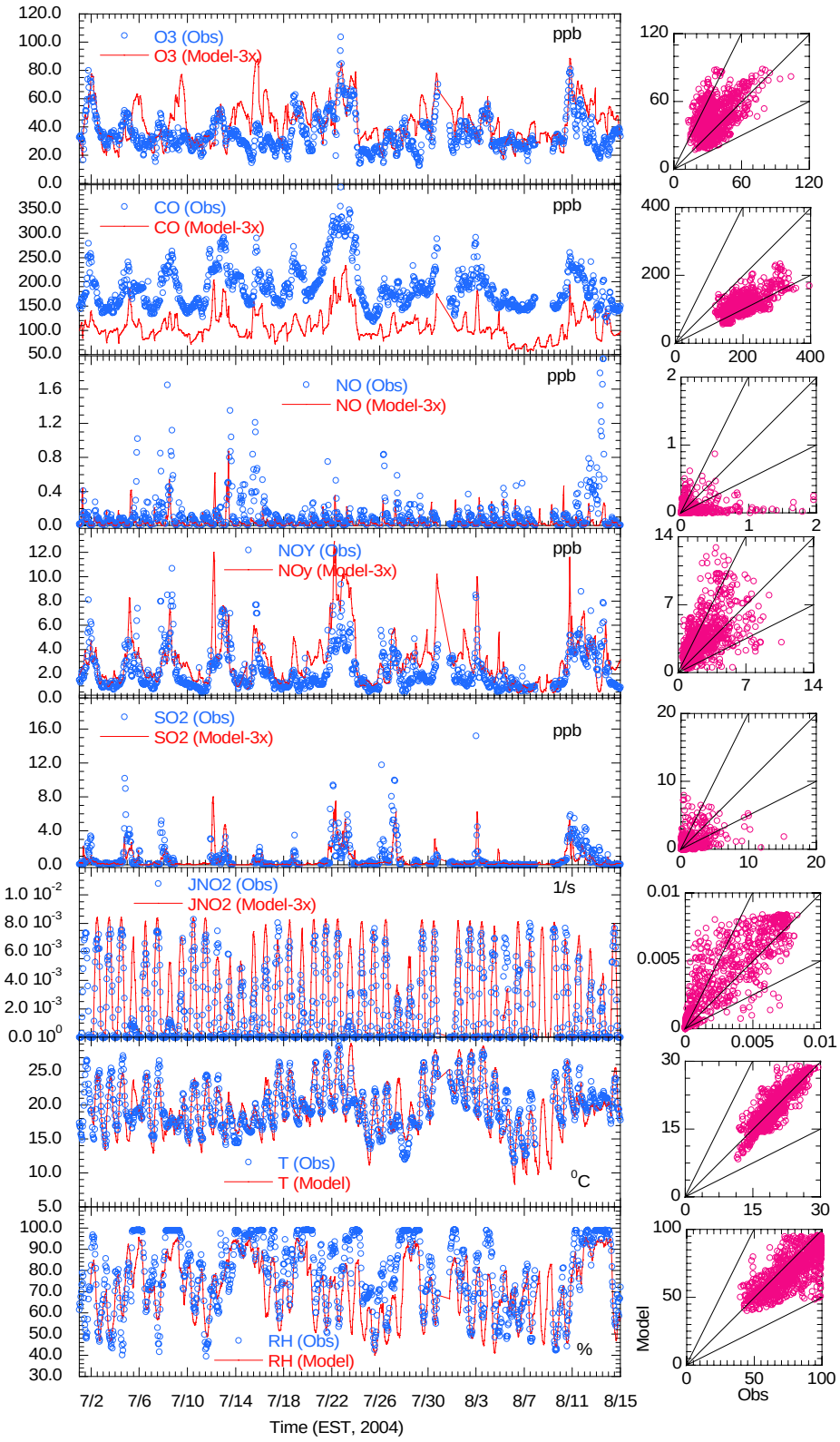


Fig. 15. Time-series and scatter plots of model predictions and observations for each parameter at the Castle Springs site.

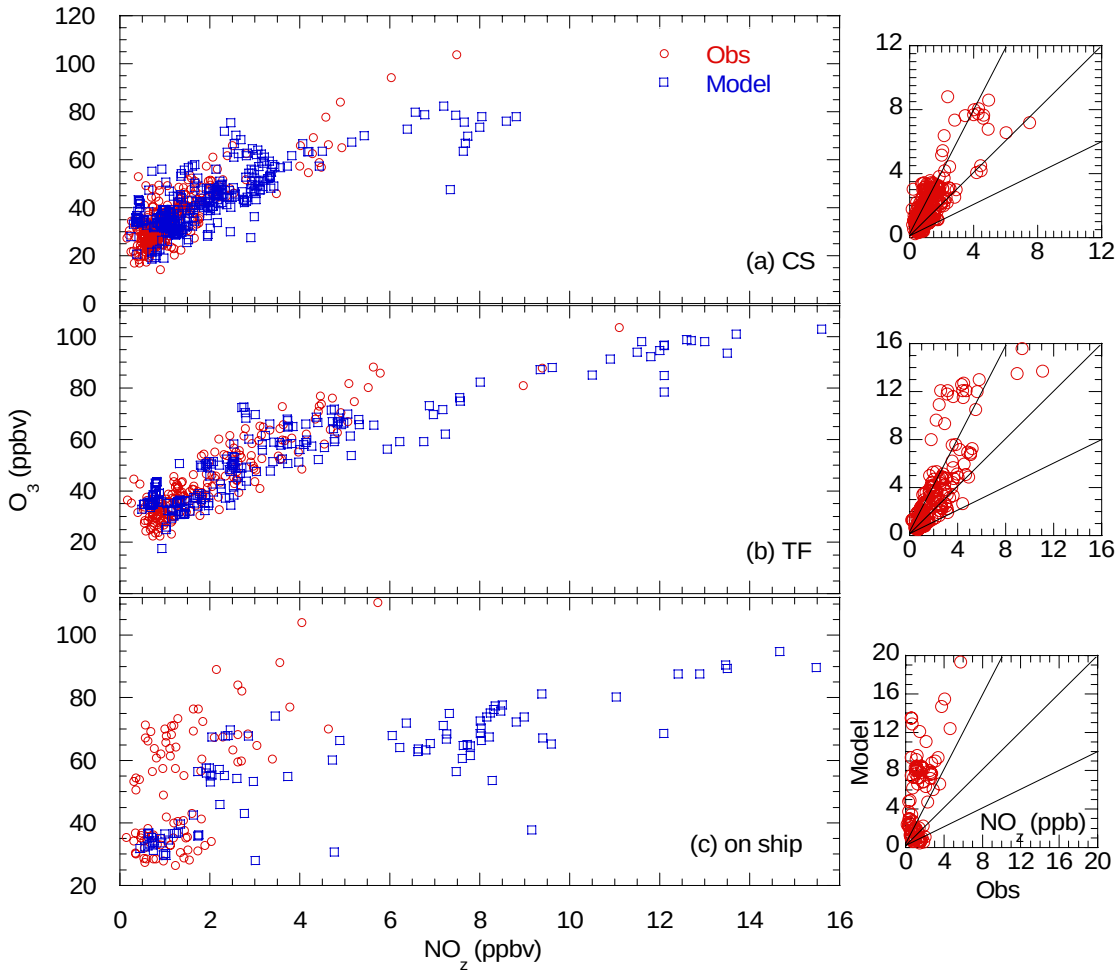


Figure 16. O_3 as a function of NO_z for the NO_x-limited conditions indicated by the observational data with $[O_3]/[NO_x]>46$ at (a) Castle Springs (CS), (b) Thompson Farm (TF), and (c) along ship tracks. Right panels are scatter plots of modeled and observed NO_z.