Atmospheric Environment 43 (2009) 355–363

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

# A study of the ozone formation by ensemble back trajectory-process analysis using the Eta–CMAQ forecast model over the northeastern U.S. during the 2004 ICARTT period

Shaocai Yu*[1,2], Rohit Mathur[1], Daiwen Kang[1,2], Kenneth Schere[1], Daniel Tong[1,2]

*Atmospheric Sciences Modeling Division, Air Resources Laboratory, National Oceanic and Atmospheric Administration, RTP, NC 27711, United States*

ARTICLE INFO

*Article history:*
Received 17 April 2008
Received in revised form
3 September 2008
Accepted 4 September 2008

*Keywords:*
Ozone formation
Eta–CMAQ model
Process analysis
Back trajectory

ABSTRACT

The integrated process rates (IPRs) estimated by the Eta–CMAQ model at grid cells along the trajectory of the air mass transport path were analyzed to quantitatively investigate the relative importance of physical and chemical processes for $O_3$ formation and evolution over the northeastern U.S. during the 2004 International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) period. The Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model is used to determine the back trajectory of air masses reaching the northeast by linking a downwind receptor to upwind source areas. The process analysis is applied to a high $O_3$ episode occurring on July 22, 2004 at three selected sites in the northeastern U.S. The process analysis at the location of the site shows that during the daytime, the $O_3$ concentrations in the surface layer are mainly enhanced by the vertical diffusion of $O_3$-rich air from aloft, followed by horizontal advection (HADV) and chemical production (CHEM), whereas dry deposition (DDEP) and vertical advection (ZADV) mainly deplete $O_3$ concentrations at the sites of Valley Central (VC), NY and Castle Spring (CS), NH. By integrating the effects of each process over the depth of the daytime planetary boundary layer (PBL), it was found that at the VC site, CHEM and HADV contributed about 53% and 41%, respectively, to $O_3$ levels within the PBL. This confirms the significance of regional transport of $O_3$ from the industrialized areas into the Northeast. On the other hand, the process analysis results for $O_3$ formation in moving air masses indicate that on July 22, large chemical production of $O_3$ along the transport path over the polluted urban regions leads to significant increase in $O_3$ in the air mass reaching the VC site, whereas the low chemical production of $O_3$ along the transport path over the low emission regions leads to the low $O_3$ concentration at the site of Belleayre Mountain (BM), NY. The dramatic buildup of $O_3$ concentration from 50 ppb to 102 ppb in the air masses before reaching the VC site after 12:00 EST on 7/22 indicates the significant impact of pollution from the northeastern urban corridor at this site. On the basis of the results at the CS site, it was found that high $NO_x$ emissions along the transport path led to large chemical production of $O_3$ in the air mass reaching the CS site on July 22. In contrast, the low chemical production of $O_3$ associated with low emission (relatively clean conditions) along the transport path over the northern portions of the domain is responsible for the low $O_3$ concentration at the CS site on July 26.



## 1. Introduction

Tropospheric $O_3$, a secondary pollutant, 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 ($NO_x$), originating from both anthropogenic (e.g., industry and vehicle emissions) and biogenic sources (e.g., forest, and soil). Harmful $O_3$ levels are typically observed during high pressure, hot, dry and stagnant atmospheric conditions as a polluted urban air mass moves downwind during the day (EPA, 2004). According to the revised maximum 8-h National Ambient Air Quality Standard (NAAQS) for $O_3$ (0.08 ppm) promulgated by the U.S. Environmental Protection Agency (EPA) in 1997, it is estimated that about 160 million Americans are exposed annually to maximum 8-h $O_3$ concentrations that exceed this NAAQS (EPA, 2004).

Lagrangian and Eulerian grid models have been used to investigate ozone exceedances on urban and regional scales. Back

* Corresponding author. Tel.: +1 919 541 0362; fax: +1 919 541 1379.
*E-mail address:* yu.shaocai@epa.gov (S. Yu).
[1] On assignment to National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, United States.
[2] On assignment from Science and Technology Corporation, 10 Basil Sawyer Drive, Hampton, VA 23666-1393, United States.

1352-2310/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2008.09.079

trajectories from the Lagrangian models have traditionally been used for identifying potential source areas of air pollutants and determining their respective contribution at receptor sites (Stohl, 1996; Draxler, 2003; Zhou et al., 2004; Song et al., 2007). Eulerian regional-scale air quality models, such as the U.S. EPA Models-3/Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006) with the process analysis option, are able to determine quantitatively the relative importance of different chemical–physical processes dictating $O_3$ distributions within selected grids. Jiang et al. (2003) applied process analysis from an Eulerian model (e.g., CALGRID) in a forward mode to the grids along the trajectory to investigate an $O_3$ episode occurring during 11–14 July 1996, in the Puget Sound region of Washington State. The combination of back trajectories and process analysis is becoming a useful diagnostic analysis tool to determine the relationship between sources and receptors with respect to $O_3$ formation within a moving air mass.

It has become evident that in the New England region of the United States, the summer $O_3$ levels exceeding the NAAQS had been observed frequently (Ray et al., 1996). These high $O_3$ episodes typically occur during periods when winds are from the southwest, placing the northeastern U.S. downwind of the concentrated $NO_x$ and VOC sources in the industrialized Midwest and Ohio River Valley (Eder et al., 1994; Mao and Talbot, 2004; Yu et al., 2006). The regional air quality in New England was the focus of the 2004 International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) study. Two of the major goals of the 2004 ICARTT were to connect surface air quality with the important features of transport and chemistry that are occurring above the surface and to determine the relative importance of local pollution compared to long-range transport in shaping local air quality. In this study, we employ a simple approach by using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT version 4.7, http://www.arl.noaa.gov/raedy/hysplit4.html, Draxler, 2003) model to determine the back trajectory which links a downwind receptor to upwind source areas and then examine detailed modeled process rates from the CMAQ along the trajectory to quantify $O_3$ formation pathways at selected sites over the northeastern U.S. during the 2004 ICARTT period. This approach can provide quantitative information about the relative importance of each process in changing $O_3$ concentrations along the trajectory. It can also determine the relationship between sources and receptors with respect to $O_3$ formation within the moving air mass.

## 2. Description of the Eta–CMAQ forecast model and calculation of back trajectories

Our study builds on previous work associated with the development and deployment of the Eta–CMAQ air quality forecasting system (Otte et al., 2005), created by linking the NOAA Eta model (Rogers et al., 1996) and the U.S. EPA's CMAQ modeling system (Byun and Schere, 2006), over a domain covering the eastern U.S. during summer 2004. Detailed description of the model configuration can be found in Yu et al. (2007). Here a brief summary relevant to the present study is presented. The horizontal domain is discretized using a grid cell size of 12 km resolution. Twenty-two layers of variable thickness set on a sigma-type coordinate are used to resolve the vertical extent from the surface to 100 hPa. The area source emissions are based on the 2001 national emission inventory (NEI). The point source emissions are based on the 2001 NEI with $SO_2$ and $NO_x$ projected to 2004 on a regional basis using the Department of Energy's 2004 Annual Energy Outlook issued in January of 2004 (Pouliot, 2005). The mobile source emissions are based on EPA'S MOBILE6 model using 1999 vehicle miles traveled (VMT) data and a fleet year of 2004. Daily temperatures from the Eta model are used to drive the inputs into the MOBILE6 model using a nonlinear least squares relationship described in Pouliot (2005). The Carbon Bond chemical mechanism (CB-IV, version 4.2) is used to represent the gas-phase reaction pathways.

This study employs the HYSPLIT model to determine the back trajectory. The same meteorology applied in the Eta–CMAQ simulation is used to generate input data sets for use in the HYSPLIT back trajectory calculation. There are many uncertainties in the calculation of trajectories arising from the possible errors in input meteorological fields and the numerical methods. To reduce uncertainties associated with a single trajectory, the HYSPLIT is run in the ensemble mode to generate multiple trajectories from a single meteorological field (Draxler, 2003). Each ensemble member is computed from the same location, but during the initial calculation the meteorological grid is offset by $\pm 0.15$ factor (i.e., $0.15 \times 12$ km here) in the horizontal ($x$, $y$) direction and $\pm 0.005$ factor (i.e., $0.005 \times 250$ m) in the vertical direction (i.e., 250 m is the initial height of the trajectory). The configuration results in 27 ensemble members and each member are assumed to have an equal probability. The integrated process rate (IPR) analysis option implemented in the CMAQ model was used. The IPR analysis deals with the effects of all the physical and chemical processes on model predictions. The IPR analysis provides the hourly time series of vertical advection/diffusion (ZADV/VDIF), horizontal advection/diffusion (HADV/HDIF), dry deposition (DDEP), cloud process (CLD), and chemical process (CHEM) for each model grid cell.

## 3. Results and discussion

### 3.1. Model performance for $O_3$

A rigorous evaluation of the Eta–CMAQ forecast model performance for $O_3$, its related precursors, and meteorological parameters over the eastern U.S. by comparing the model results with the observations obtained during the 2004 ICARTT study is summarized in Yu et al. (2007). The performance relative to several other regional models is further summarized in McKeen et al. (2005). Here a brief summary is presented. The results at the AIRNow surface sites show that the model was able to reproduce the day-to-day variations of observed daily maximum 8-h $O_3$ and captured the majority (73%) of observed daily maximum 8-h $O_3$ within a factor of 1.5 with normalized mean bias of 22%. On the basis of results from aircraft, ozonesonde and ship-based lidar observations, the model generally reproduced $O_3$ vertical structures during most days at low altitudes with consistent overestimations above $\sim$6 km due to the lateral boundary conditions derived by the global forecast system (GFS) and coarse model resolution in the free troposphere. The model captured the vertical variation patterns of the observed values for other species ($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 regions and air mass characteristics. The model can generally capture the observed photolysis rates ($JNO_2$) very well during the cloud-free periods, but underestimated the $JNO_2$ values by 20–90% when there was a solid cloud deck below the aircraft and overestimated $JNO_2$ values significantly when solid cloud deck was above the aircraft. The model was able to reproduce the vertical profiles of observed water vapor and wind speed. Time-series comparisons at the Atmospheric Investigation, Regional Modeling, Analysis and Prediction (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 for NO and $SO_2$ at each site, although there were occasional major excursions.

### 3.2. Process analysis at the selected site locations

Analysis results of modeled process budgets for 2 sites: Castle Springs (CS), New Hampshire (NH), and Valley Central (VC), NY, on

July 22, 2004, are presented because there was an $O_3$ episode over the northeastern domain on this day and these two sites exhibited high $O_3$ concentrations as shown in Fig. 1a. The $NO_x$ emissions over the domain at 20:00 UTC July 22, 2004, are shown in Fig. 1c. $NO_x$ emissions in the northeastern U.S. are primarily from the transportation (54%), electric utilities (25%) and industrial source (11%) as shown in Driscoll et al. (2003). As shown in Fig. 1d, the model $NO_x$ concentration distributions are generally in agreement with the $NO_x$ emissions over the domain. Figs. 2 and 3 show the flight pathways of the aircraft P-3 and DC-8 and levels of atmospheric species along these flight tracks measured on July 22, 2004, over the eastern U.S. during the 2004 ICARTT period. Note that the measurements onboard the P-3 cover a regional area over the northeast around New York and Boston from 0 to ~5 km altitudes (see Fig. 2a), whereas the DC-8 aircraft covers a broader regional area over the eastern U.S. between 0 and ~12 km altitudes (see Fig. 3a). On July 22, a progressively more aged NY city plume over the Bay of Fundy reaching beyond Cape Breton and Prince Edward Island was sampled by the P-3 (Yu et al., 2007). Upon closer inspection, it is noted that as shown in Fig. 2, the P-3 measurements between 459 and 1165 m altitudes sampled highly polluted air masses from the impact of the northeastern urban corridor over the areas of the New York and Massachusetts during 20:33 and 21:33 UTC with very high $O_3$ (80–123 ppb), CO (208–294 ppb), $HNO_3$ (2–11 ppb), $SO_2$ (0.3–12 ppb) and $NO_x$ (0.3–3 ppb). On the other hand, on the basis of the DC-8 measurements shown in Fig. 3, during 12:07 and 13:18 UTC, the DC-8 measurements between 4305 and 9586 m altitudes sampled polluted offshore flows over the Atlantic coast of North America (latitude: 35.5–42.9°N; longitude: −70.4 to −72.7°W) with very high $O_3$ (80–267 ppb), CO (57–159 ppb) and $NO_2$ (0.01–0.64 ppb) under the impact of anthropogenic sources from the Washington, D.C./New York City/Boston urban corridor and biogenic emissions in New Hampshire and Maine. During 20:31 and 20:55 UTC between 40 and 1136 m altitudes, the DC-8 aircraft also encountered a high pollution episode over the western part of NH (latitude: 43.1–44.3°N; longitude: −70.4 to −71.15°W) with high $O_3$ (81–130 ppb), CO (214–330 ppb) and $NO_2$ (0.3–1.4 ppb). Both surface (see Fig. 1) and aloft (see Figs. 2 and 3) indicate that there was an $O_3$ pollution episode over the northeastern domain on July 22, 2004.

Figs. 4 and 5 illustrate the results of the process analysis for the grid cells of the CS and VC sites, respectively. The hourly rates of change in $O_3$ due to each process at the surface layer shown in Figs. 4b and 5b indicate that during the daytime (7:00–19:00 LST), the $O_3$ concentrations in the surface layer are enhanced predominantly due to vertical diffusion of $O_3$-rich air from aloft, followed by horizontal advection and chemical production, whereas dry deposition and vertical advection mainly deplete $O_3$ from the surface layer. The effects of horizontal diffusion and aqueous processes on $O_3$ formation are negligible in this case. Similar features are also noted for the modeled results at the VC site except that chemical production is also a significant contributor to $O_3$ enhancements at this site. At the CS site, the mean rates of change in $O_3$ from 10:00 EST to 16:00 EST for CHEM, DDEP, HADV, HDIF, VDIF, and ZADV processes at the surface layer are 2.88, −45.70, 15.75, −0.008, 41.09, and −9.55 ppbv $h^{-}$, respectively, whereas at the VC site, they are 12.41, −46.87, 28.29, −0.028, 33.67, and −20.63 ppbv $h^{-}$, respectively (see Figs. 4a and 5a). Obviously, the higher positive rates of change of $O_3$ at the VC site due to CHEM and HADV processes relative to those at the CS site result in the larger increase in $O_3$ concentration from 70 ppb at 10:00 EST to 114 ppb at 16:00 EST, whereas the $O_3$ concentrations only change from 59 ppb at 10:00 EST to 88 ppb at 16:00 EST at the CS site as shown in Figs. 4c and 5c. It is of interest to note that from layer 2 (~74 m) to layer 8 (1000 m), chemistry and horizontal advection are still major sources, whereas the vertical diffusion and advection contributions become negative at both sites. Note that the negative and positive values indicate the reduction and enhancement of $O_3$

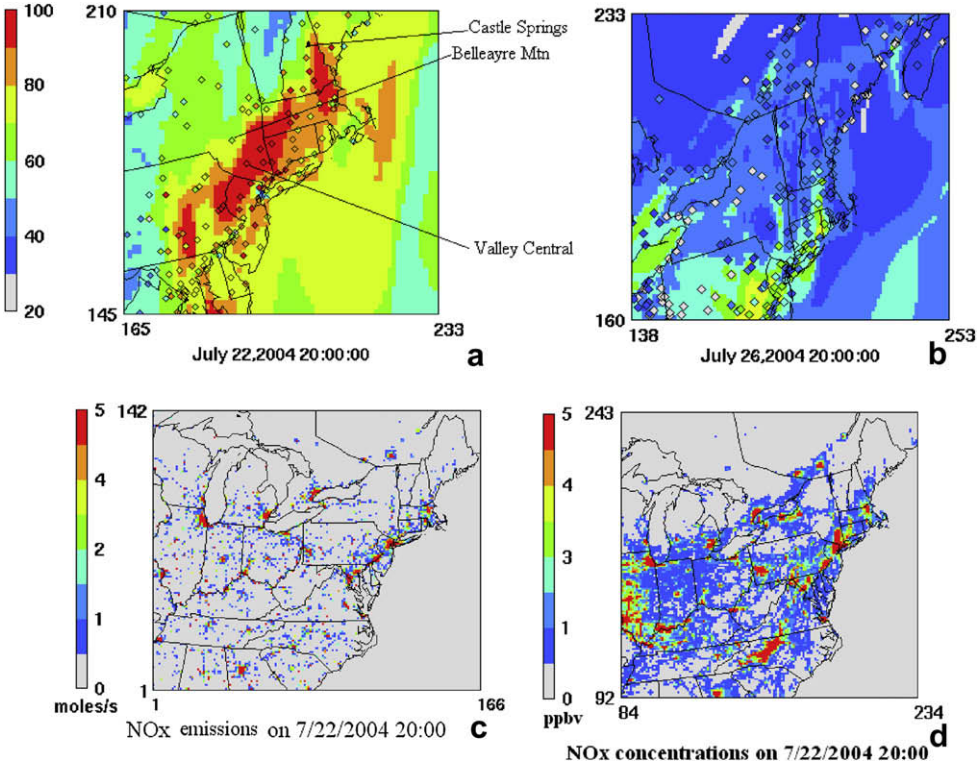


**Fig. 1.** (a) Eta–CMAQ simulation results for $O_3$ concentration (ppb) with AQS observed data overlaid (◇) at 20:00 UTC (3:00 p.m.) on July 22, 2004 and (b) same as (a) but on July 26, 2004, and (c) $NO_x$ emission and (d) $NO_x$ concentrations at 20:00 UTC on July 22, 2004.

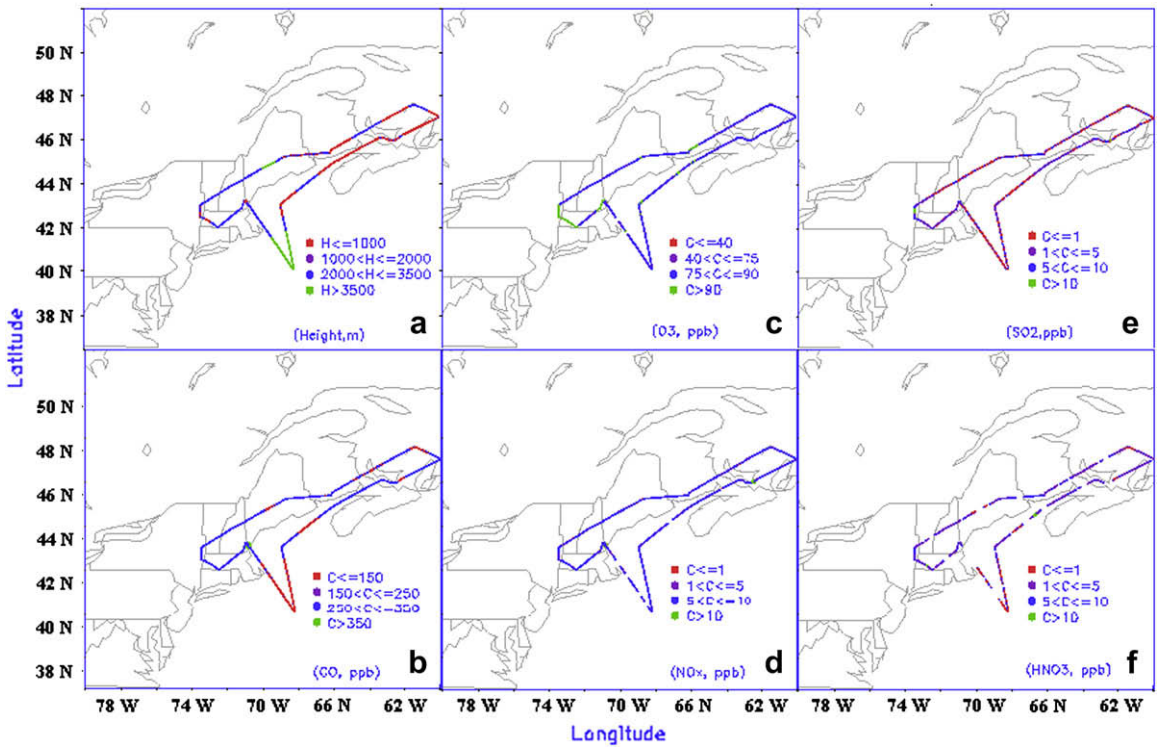


**Fig. 2.** The observed track height and species concentrations along the aircraft P-3 track from 13:48 to 21:33 UTC on July 22, 2004: (a) height, (b) CO, (c) $O_3$, (d) $NO_x$, (e) $SO_2$ and (f) $HNO_3$.

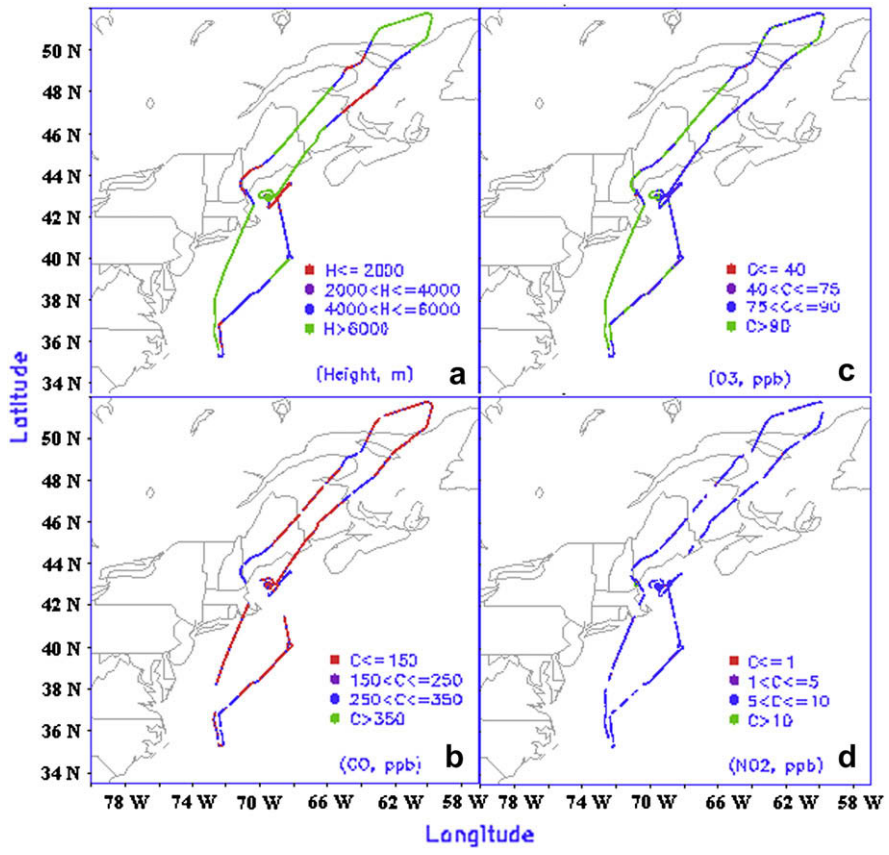


**Fig. 3.** The observed track height and species concentrations along the aircraft DC-8 track from 12:00 to 20:55 UTC on July 22, 2004: (a) height, (b) CO, (c) $O_3$, (d) $NO_2$.

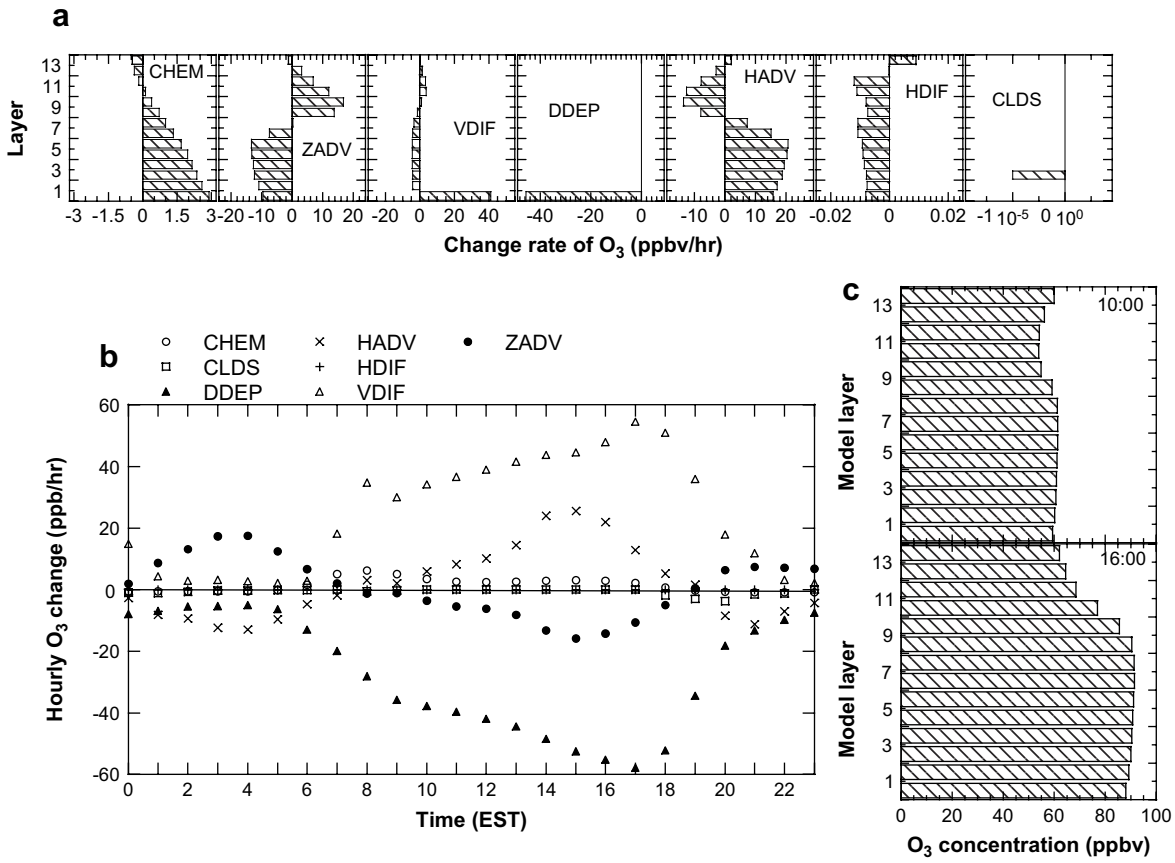


**Fig. 4.** (a) The mean $O_3$ change rates due to different processes for each layer from 10 to 16 EST, (b) the hourly $O_3$ change rate due to different processes at the surface layer, (c) evolution of $O_3$ vertical profiles from 10:00 to 16:00 EST within the PBL. All results are at Castle Springs (CS) site on July 22, 2004.

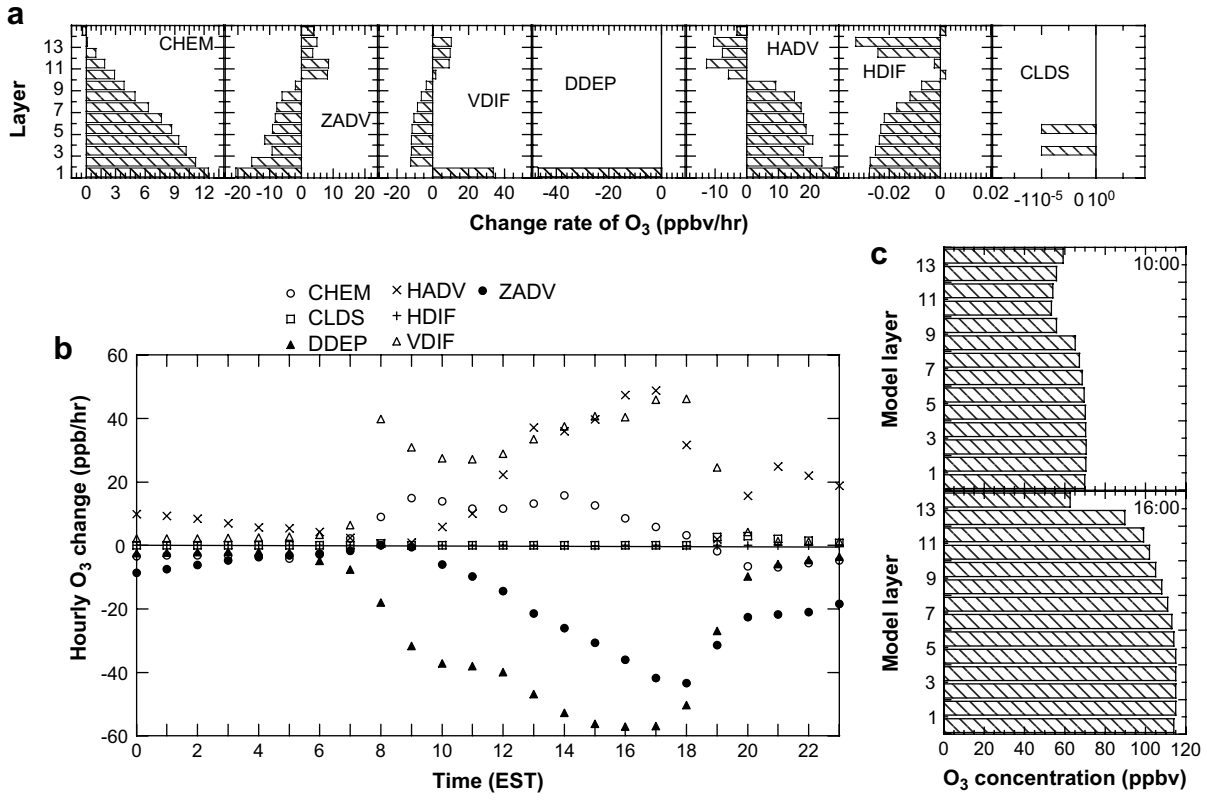


**Fig. 5.** The same as Fig. 4 but at Valley Central site, NY, on July 22, 2004.

**Table 1**
The mean values of the integrated effects of each process contribution to $O_3$ formation over the planetary boundary layer (PBL) from layer 1 to layer 14 for the period of 10:00 to 16:00 EST, July 22, 2004, at the CS and VC sites. The values in parentheses are percentages.

| Process (ppb $h^{-}$) | CS | VC |
|---|---|---|
| Source | | |
| Chemical process | 0.52 (13.6) | 3.34 (53.2) |
| Horizontal advection | 1.31 (34.2) | 2.56 (40.7) |
| Vertical diffusion | 0.07 (1.9) | 0.36 (5.8) |
| Vertical advection | 1.92 (50.3) | 0.03 (0.4) |
| Sink | | |
| Cloud process | 0.000 (0.0) | 0.000 (0.0) |
| Dry deposition | −0.66 (99.2) | −0.68 (98.2) |
| Horizontal diffusion | $-5.0 \times 10^{-3}$ (0.8) | $-1.3 \times 10^{-2}$ (1.9) |

concentrations, respectively. These processes contribute to the enhancement of $O_3$ vertical profiles from 10:00 to 16:00 EST as shown in Figs. 4c and 5c.

Since during daytime conditions, pollutants are generally well mixed through the planetary boundary layer (PBL), we examine the process budgets integrated over the depth of the daytime PBL. We use 2 km as being representative of typical daytime PBL in the region; this corresponds to layer 14 of the current model configuration. Table 1 shows that at the VC site, photochemical production and horizontal advection contributed about 53% and 41% of net sources of $O_3$ within the PBL, respectively, whereas at the CS site, photochemical production, horizontal advection and vertical advection contributed about 14%, 34% and 50% of net sources of $O_3$ within the PBL, respectively. This confirms the significance of regional transport of $O_3$ from the industrialized areas into the Northeast.

### 3.3. Process analysis of $O_3$ formation in a moving air mass

To gain further insight into the $O_3$ formation mechanisms over the northeastern U.S., we examine process analysis rates in the CMAQ grid cells along the 3-day back trajectories of the air mass reaching three sites (CS, VC, and Belleayre Mtn (BM), NY, see Fig. 1a). Since during the daytime pollutants are well mixed vertically through the PBL, we examine vertically integrated process

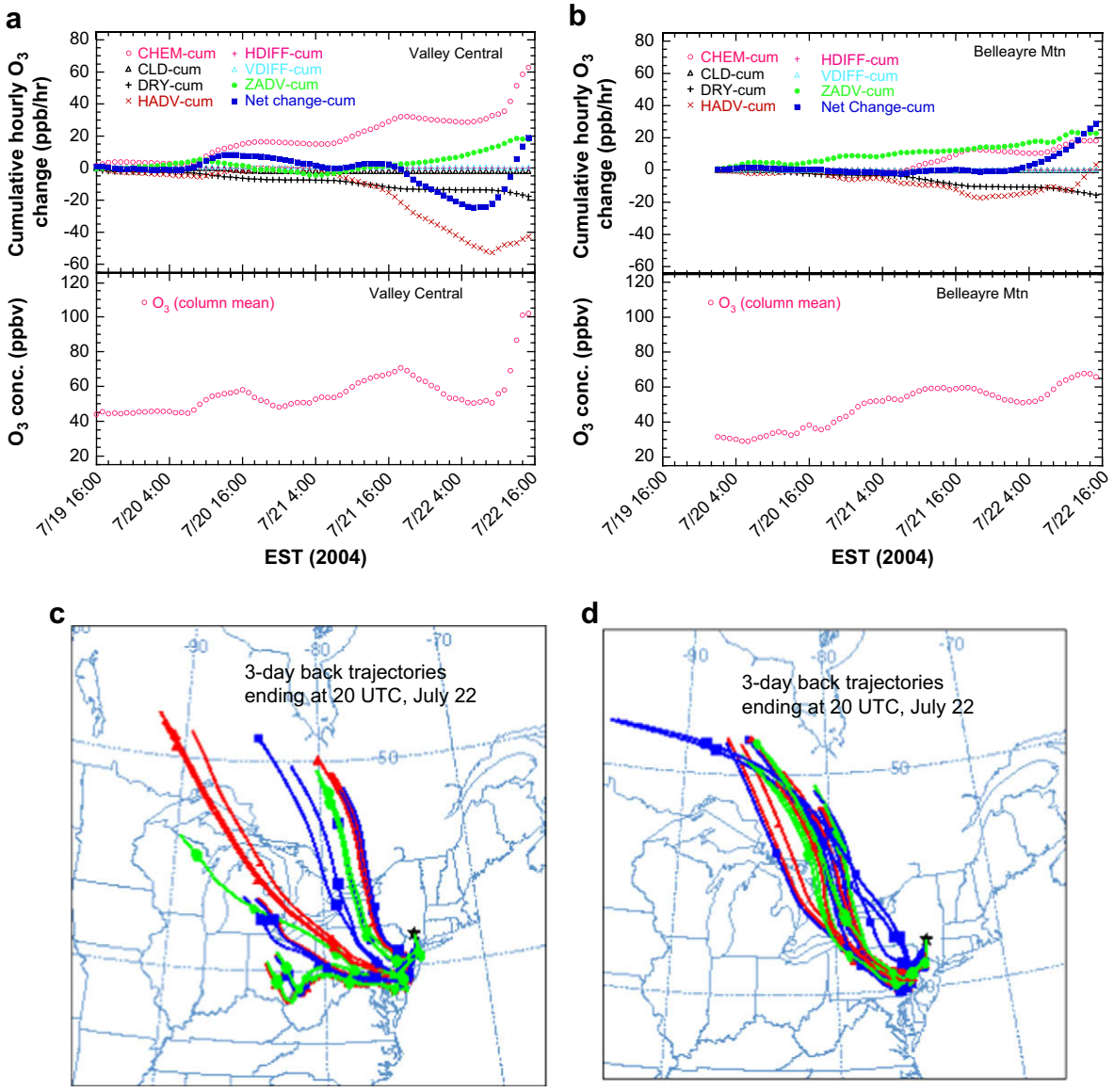


**Fig. 6.** Total accumulated contribution of each process to $O_3$ formation, column mean $O_3$ concentrations along the back trajectories and 3-day back trajectories at Valley Central (a, c) and Belleayre Mountain (b, d) on July 22, 2004.

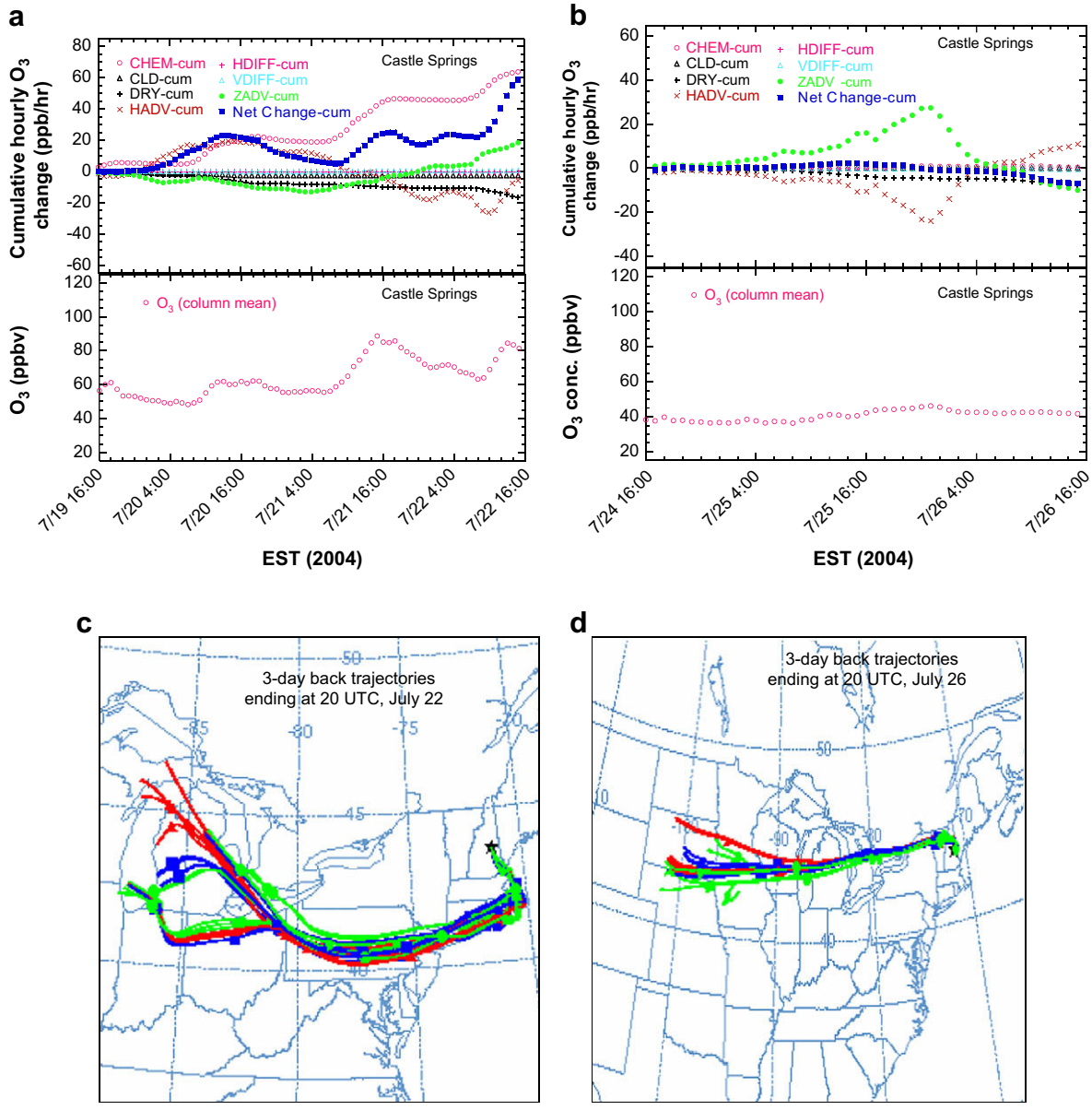


**Fig. 7.** The same as Fig. 6 but at Castle Springs site for July 22 and 26, 2004.

tendencies; we choose 2 km (approximately corresponding to layer 14 in the model) as being representative of the mean daytime PBL height for this analysis. In a study of the summertime atmospheric boundary layer over the eastern U.S., Rao et al. (2003) found that the PBL heights can vary from <200 m (nighttime) to ~2.5 km (the afternoon). Yu et al. (2007) indicated that average daytime PBL heights during the ICARTT period in the Eta model at Concord, NH, can be ~2 km. Compared to observed PBL heights derived from a radar wind profile at Concord, NH, averaged over the ICARTT period, the Eta model tended to overestimate the observations at all times (Yu et al., 2007). Additionally, since efficient long-range transport occurs above the nocturnal PBL, we use a height of 2 km to integrate the process tendencies along the back trajectories. Also, since dispersion is irreversible, the 2 km layer should be maintained for the full duration of the trajectory. By running the HYSPLIT model in the ensemble mode, 27 ensemble back trajectories at each site are generated as shown in Figs. 6c, d, 7c and d. The process analysis results for each ensemble back trajectory are retrieved. Then, total accumulated contributions of each process to $O_3$ formation along the back trajectory are calculated by averaging the corresponding values of each ensemble back trajectory for each site. The column mean results of accumulated contributions between layer 1 and layer 14 along the 3-day back trajectory are shown in Figs. 6a, b, 7a and b. Table 2 summarizes the total accumulated contributions of each process to the $O_3$ formation along the 3-day back-trajectories at the three sites.

As illustrated in Fig. 1, there are contrasting regional $O_3$ conditions occurred over the northeastern U.S. on July 22 and July 26. To determine the reasons contributing to these differing $O_3$ distributions over the northeastern U.S., we examine process analysis results at the three representative selected sites. For the high $O_3$ episode on July 22, the modeled $O_3$ concentrations at the VC and BM sites at 15:00 EST were 102 and 68 ppbv, respectively, close to the corresponding observed concentrations (97 and 64 ppbv, respectively) as shown in Fig. 1. Comparison of modeled $O_3$ fields with observations (shown as overlaid diamonds in Fig. 1) also shows that the model actually simulated the spatial distribution of high $O_3$ concentrations very well over the northeastern

**Table 2**
Summary of the total accumulated contributions of each process to $O_3$ formation for 3-day back-trajectories (see Figs. 6 and 7) on the basis of column means from layer 1 to layer 14 (PBL height) at the three sites. The total accumulated $NO_x$ emission along 3-day back-trajectories is also listed.

| Processes for $O_3$ | Sites (2004) | | | |
|---|---|---|---|---|
| | CS, NH | CS, NH | VC, NY | Belleayre Mtn, NY |
| ppb | 7/22 | 7/26 | 7/22 | 7/22 |
| Chemical process (CHEM) | 63.9 | 0.4 | 62.8 | 18.3 |
| Cloud process (CLD) | −2.6 | 0.0 | −1.9 | −0.2 |
| Dry deposition (DDEP) | −16.2 | −7.6 | −17.7 | −15.6 |
| Horizontal advection (HADV) | −5.6 | 10.9 | −42.7 | 3.2 |
| Horizontal diffusion (HDIF) | −0.1 | 0.0 | 0.0 | 0.0 |
| Vertical diffusion (VDIF) | 0.9 | −0.7 | 0.7 | 0.2 |
| Vertical advection (ZADV) | 18.6 | −10.0 | 17.8 | 22.9 |
| Total source | 83.4 | 11.3 | 81.3 | 44.7 |
| Total sink | −24.5 | −18.4 | −62.3 | −15.8 |
| Net change | 58.9 | −7.0 | 19.0 | 28.9 |
| Column mean $O_3$ | 81.3 | 41.7 | 102.5 | 65.8 |
| Total accumulated $NO_x$ emission along the back trajectory (mol $s^{-}$) | 39.0 | 8.1 | 142.5 | 23.4 |

U.S. for this time. Although the similar meteorological synaptic patterns impact the two sites, the obvious differences in the $O_3$ chemical production resulting from different air mass paths explain why $O_3$ concentrations at the VC site are much higher than those at the BM site. The ensemble back trajectories of Fig. 6c indicate that air masses transporting from the industrialized Midwest, Ohio River Valley, and northeastern urban corridor significantly affected the VC site. The dramatic buildup of $O_3$ concentration from 50 ppb to 102 ppb in the air masses before reaching the VC site after 12:00 EST on July 22 (see Fig. 6a) indicates the significant impact of pollutant emissions from the northeastern urban corridor (see Fig. 1c for the $NO_x$ emissions and see Fig. 6c for back trajectories) on the site, which resulted in the noted large increase of chemical production (see Fig. 6a). The air masses in the southwest offshore flows had passed over the northeastern urban corridor during the previous 2–24 h before being sampled at the VC site as shown in Fig. 6c. The large negative horizontal advection made a big contribution to the decrease of $O_3$ concentration in the air masses reaching the VC site early before 12:00 EST on 7/22. On the other hand, the air masses influencing the BM site, which originated from the north and passed the north Pennsylvania as shown in Fig. 6d, are not impacted by anthropogenic emission sources in the northeastern U.S. The $O_3$ concentrations in the air masses reaching the BM site only increase slightly from the 30 ppb to 65 ppb along the air mass trajectory as shown in Fig. 6b. The total $O_3$ chemical productions along the 3-day back trajectories (see Fig. 6d) were 62.8 ppb at the VC site and 18.3 ppb at the BM site as summarized in Table 2. The total of the $NO_x$ emission rates along the 3-day back trajectories (see Fig. 6d) was 142 mol $s^{-}$ at the VC site and 23.4 mol $s^{-}$ at the BM site (see Table 2). This indicates that relatively large chemical production of $O_3$ along the transport path over polluted urban regions leads to significant increase in $O_3$ in the air mass reaching the VC site. In contrast, considerably lower chemical production of $O_3$ along the transport path over the low emission regions leads to the noted low $O_3$ concentration at the BM site. Obviously, the meteorological conditions also need to be conducive to the formation of $O_3$ in addition to the high $NO_x$ emission for the VC site.

For the low $O_3$ episode on July 26, clear continental flow from the north mainly impacted the northeastern U.S. The modeled $O_3$ concentration at the CS site at 15:00 EST on July 26 is 38 ppb, close to the observed value of 35 ppbv, as indicated in Fig. 1. The model simulation results overlaid with the observations (represented by diamonds) in Fig. 1 also shows that the model actually reproduced the regional low $O_3$ concentrations very well over the northeastern U.S. The ensemble back trajectories of Fig. 7d indicate that transport of clean air masses from the northwestern and northern continent affected the CS site predominantly on July 26. As expected, these northerly air masses are not impacted by large anthropogenic emission source (see Fig. 1c). This results in significantly lower $O_3$ concentration over the northeastern U.S. on this day. Fig. 7b shows that the $O_3$ concentrations are consistently low (~40 ppb) along the back trajectories of air masses reaching the CS site (see Fig. 7d) in this day. On the contrary, on July 22, the transport of air masses from the northeastern urban corridor affected the CS site as shown in Fig. 7c. This results in the increase of $O_3$ concentrations from 60 ppb at 15:00 EST, July 19, to 81 ppb at 15:00 EST, July 22, along the back trajectories of air masses reaching the CS site as shown in Fig. 7a. The total of the $NO_x$ emission rates of sources along the 3-day back trajectories (see Fig. 6d) for the CS site were 39.0 mol $s^{-}$ on July 22 and 8.1 mol $s^{-}$ on July 26 (see Table 2). Previous analysis of $[O_3]/[NO_x]$ ratios (Yu et al., 2006, 2007) suggests that $O_3$ formation at the CS site is primarily $NO_x$-sensitive during the 2004 ICARTT period. This is also in agreement with other investigators (Olszyna et al., 1994), who also found that the eastern U.S. is primarily $NO_x$-sensitive. The total $O_3$ chemical production along the 3-day back trajectories (see Fig. 7c and d) was 63.9 ppb on July 22 and 0.4 ppb on July 26 as summarized in Table 2. This suggests that chemical production of $O_3$ associated with injection of $NO_x$ emissions in the air mass along the transport path leads to higher $O_3$ in the air mass reaching the CS site on July 22. In contrast the trajectory of air mass reaching the CS site at 15:00 EST on July 26 passed over comparatively lower $NO_x$ emission regions in the northwestern part of the domain, resulting in the lower modeled and observed $O_3$ levels.

## 4. Summary

To investigate the details of $O_3$ formation and evolution over the northeastern U.S. during the 2004 ICARTT period, the HYSPLIT model is used to determine the back trajectories of air masses reaching the northeastern U.S. by linking a downwind receptor to upwind source areas. The role of atmospheric and chemical processes on $O_3$ concentrations is then quantitatively analyzed by integrating process budgets along these trajectories simulated by the CMAQ atmospheric chemical transport model. This approach provides the quantitative information on the relative importance of each process in changing $O_3$ concentration both locally as well as along air mass trajectory history. The process analysis results on a local basis for a single grid cell of the site show that during the daytime, the $O_3$ concentrations in the surface layer are mainly enhanced by the vertical diffusion of $O_3$-rich air from aloft, followed by horizontal advection and chemical production, whereas dry deposition and vertical advection mainly deplete $O_3$ concentrations, on the basis of the results at the VC and CS sites. By integrating the effects of each process over the planetary boundary layer (PBL) from layer 1 to layer 14, it was found that at the VC site, photochemical production and horizontal advection contribute about 53% and 41% of net sources of $O_3$ within the PBL, respectively, whereas at the CS site, photochemical production, horizontal advection and vertical advection contribute about 14%, 34% and 50% of net sources of $O_3$ within the PBL, respectively. This confirms the significance of regional transport of $O_3$ from the industrialized areas into the Northeast. To determine the reasons contributing to contrasting regional $O_3$ conditions occurred over the northeastern U.S. on July 22 and July 26, we applied the process analysis to the CMAQ grid cells along the trajectory of the air mass transport path at the three representative selected sites. The results indicate that on July 22, relatively large chemical production of $O_3$ along the

transport path over polluted urban regions leads to significant increase in $O_3$ in the air mass reaching the VC site. In contrast, considerably lower chemical production of $O_3$ along the transport path over the low emission regions leads to the noted low $O_3$ concentration at the BM site. The dramatic buildup of $O_3$ concentration from 50 ppb to 102 ppb in the air masses before reaching the VC site after 12:00 EST on 7/22 indicates the significant impact of pollution from the northeastern urban corridor on the site. The results at the CS site show that chemical production of $O_3$ associated with injection of $NO_x$ emissions in the air mass along the transport path leads to higher $O_3$ in the air mass reaching the CS site on July 22. In contrast the trajectory of air mass reaching the CS site at 15:00 EST on July 26, passed over comparatively lower emission regions in the northwestern part of the domain with considerably lower chemical production of $O_3$, resulting in the lower modeled and observed $O_3$ levels.

## Acknowledgements

The authors would like to thank Drs. S.T. Rao, Winston Luke, J. Godowitch and two anonymous reviewers 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. 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

Byun, D.W., Schere, K.L., 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.

Draxler, R.R., 2003. Evaluation of an ensemble dispersion calculation. Journal of Applied Meteorology 42, 308–317.

Driscoll, C.T., Whitall, D., Aber, J.D., Boyer, E., Castro, M., Cronan, C., Goodale, C., Groffman, P., Hopkinson, C., Lambert, K., Lawrence, G., Ollinger, S.V., 2003. Nitrogen pollution in the northeastern United States: sources, effects, and management options. Bioscience 53, 357–374.

Eder, B.K., Davis, J.M., Bloomfield, P., 1994. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. Journal of Applied Meteorology 33, 1182–1199.

Jiang, G., Lamb, B., Westberg, H., 2003. Using back trajectories and process analysis to investigate photochemical ozone production in the Puget Sound region. Atmospheric Environment 37, 1489–1502.

Mao, H., Talbot, R., 2004. $O_3$ and CO in New England: temporal variations and relationships. Journal of Geophysical Research 109, D21304. doi:10.1029/2004JD004913.

McKeen, S., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Hsie, E.-Y., Gong, W., Bouchet, V., Menard, S., Moffet, R., McHenry, J., McQueen, J., Tang, Y., Carmichael, G.R., Pagowski, M., Chan, A., Dye, T., Frost, G., Lee, P., Mathur, R., 2005. Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004. Journal of Geophysical Research 110, D21307. doi:10.1029/2005JD005858.

Otte, T.L., Pouliot, G., Pleim, J.E., Young, J.O., Schere, K.L., Wong, D.C., Lee, P., Tsidulko, M., McQueen, J.T., Davidson, P., Mathur, R., Chuang, H.-Y., DiMego, G., Seaman, N.L., 2005. Linking the Eta model with the community multiscale air quality (CMAQ) modeling system to build a national air quality forecasting system. Weather and Forecasting 20, 367–384.

Pouliot, G.A., 2005. The emissions processing system for the Eta/CMAQ air quality forecast system. Proceedings of seventh Conference on Atmospheric Chemistry. The 85th AMS Annual Meeting, Paper 4.5, Amer. Meteor. Soc., San Diego, CA, 2005.

Rao, S.T., Ku, J.Y., Berman, S., Zhang, K., Mao, H., 2003. Summertime characteristics of the planetary boundary layer over the eastern United States and relationships to ambient ozone levels. Pure and Applied Geophysics 160, 21–55.

Ray, J.D., Heavner, R.L., Flores, M., Michaelsen, C.W., 1996. Surface level measurements of ozone and precursors at coastal and offshore locations in the Gulf of Maine. Journal of Geophysical Research 101 (D22), 29,005–29,012.

Rogers, E., Black, T., Deaven, D., DiMego, G., Zhao, Q., Baldwin, M., Junker, N., Lin, Y., 1996. Changes to the operational "early" Eta analysis/forecast system at the national centers for environmental prediction. Weather Forecasting 11, 391–413.

Song, C.H., Han, K.M., Cho, H.J., Kim, J., Carmichael, G.R., Kurata, G., Thongboonchoo, N., He, Z., Kim, H.S., 2007. A Lagrangian model investigation of chemico-microphysical evolution of northeast Asian pollution plumes within the MBL during TRACE-P. Atmospheric Environment 41, 8932–8951.

Stohl, A., 1996. Trajectory statistics – a new method to establish source–receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. Atmospheric Environment 30, 579–587.

U.S. EPA, April, 2004. The Ozone Report: Measuring Progress Through 2003. Office of Air Quality Planning and Standards, the U.S. Environmental Protection Agency, EPA-454-K-04-001, Research Triangle Park, NC.

Yu, S.C., Mathur, R., Kang, D., Schere, K., Eder, B., Pleim, J., 2006. Performance and diagnostic evaluations of a real-time ozone forecast by the Eta–CMAQ model suite during the 2002 New England Air Quality Study (NEAQS). Journal of the Air & Waste Management Association 56, 1459–1471.

Yu, S.C., Mathur, R., Schere, K., Kang, D., Pleim, J., Otte, T.L., 2007. A detailed evaluation of the Eta–CMAQ forecast model performance for $O_3$, its related precursors, and meteorological parameters during the 2004 ICARTT study. Journal of Geophysical Research 112, D12S14. doi:10.1029/2006JD007715.

Zhou, L., Hopke, P.K., Liu, W., 2004. Comparison of two trajectory based models for locating particle sources for two rural New York sites. Atmospheric Environment 38, 1955–1963.