

1 Overview and Summary

1.1 Purpose

The overall purpose of this *three* institution project was to apply both the Community Multiscale Air Quality model (CMAQ) [2] and the Comprehensive Air Quality Model with Extensions (CAMx) [1] to investigate the origins of ozone (O_3) in east Texas and, in particular, near the Dallas area. The fundamental idea was to improve the conceptual model for O_3 in all of East Texas, with particular concern for the quality of source apportionment, especially for O_3 transport. Previous modeling research by ENVIRON Corp. [11] using a CAMx Dalla/Fort Worth 1999 episode had shown sensitivity of O_3 to emissions of permitted or planned electric generation units (EGUs) to be located in East Texas. The main modeling evidence was the source apportionment techniques. Sensitivity and process analysis tools were employed to investigate the underlying factors behind source apportionment results by various techniques. The conceptual model for O_3 in east Texas was investigated through the development of a 2005 seasonal ozone model focused on Houston and Dallas, Texas. To provide a test bed for our process analysis investigation, while the seasonal episode was being developed, several Houston centered episodes were investigated by UNC. These included the TCEQ's 1-hour mid-course correction SIP 2000 episode and four post-2000 episodes performed by Alpine Geophysics for the industrial stakeholder group (8HCG).

ENVIRON and UH were charged with preparing CAMx inputs for an ozone seasonal model to be used for the conceptual model analysis. This model would also be used for the source apportionment inter-comparison.

UNC was tasked with using its Python-based Performance Analysis Support System (pyPASS) analysis software to help evaluate model performance of the Dallas-Fort Worth 2005 seasonal ozone episode. This tool was designed to implement the Protocol for Regulatory Ozone Modeling Performance Tests (PROMPT) [8,9]. Initially, ENVIRON developed a modeling episode that spanned several months in 2005 with grid cells at a 12 km resolution. The group decided to focus on the 5 groupings of days within the ozone season listed in Table 4.2. ENVIRON was tasked with de-

veloping 4 km resolution CAMx meteorological files that would then be sent to UNC. Based on our analysis of these wind fields, the group would then recommend one or two of these focus periods for further analysis.

Goals

The goals from the University of North Carolina-Chapel Hill (UNC) on this project were threefold:

- Develop Process Analysis tools to improve current O₃ source apportionment techniques;
- Provide performance evaluation data for Dallas and Houston 2005 simulations; and
- Present relevant findings from regulatory modeling efforts in Houston.

Objectives

Specific objectives from the University of North Carolina-Chapel Hill (UNC) on this project were:

- Serve as the external analysts for data pertaining to performance evaluations for the 2005 CAMx and CMAQ simulations integrating observations and model output using techniques and software developed in HARC project H12, e.g., the pyPASS system.
- Expand our process analysis system and apply it to all simulations conducted by the parties.
- Integrate the results and simulations performed by the industrial stakeholder's 8-Hour Ozone Group.
- Assist the University of Houston (UH) and ENVIRON Inc. in developing process analysis tools within the Anthropogenic Precursor Culpability Assessment (APCA).

1.2 Approach

The first step to incorporate our process analysis techniques into the APCA tool was a total re-design of the current process analysis tool [4,5,6]. The tool was repackaged into a cleaner more efficient formulation and tested with the Houston 2000

episode used for the State implementation plan (SIP). We then used the relevant processes from the newly re-formulated tool, pyPA (Python Process Analysis), and incorporated them into APCA to create a new tool called Kinetic Process Analysis Source Apportionment (KPASA). These new algorithms were expressed in a new scientific scripting language, Python (<http://www.python.org>), that has powerful data representation and extensive support libraries. The KPASA tool adopts the APCA tracer methodology of tracking the movement of emitted precursors and O₃, but also incorporates techniques used in the pyPA tool. By incorporating the pyPA methodology the allocation technique used to attribute chemically produced O₃ to its precursors was improved.

We fulfilled our role as "gatekeeper" by collecting model and observational data relevant to this project. External hard drives were shipped to ENVIRON and UH to collect modeling data. Observational data was collected from various state and private agencies. We then used our pyPASS tool to integrate and provide the data needed for a holistic day by day and monitor by monitor comparison. We also helped evaluate specific periods within the Dallas modeling episode that would warrant further detailed investigation.

1.3 Products

A major product from this project was the complete redesign of the UNC process analysis legacy code into a cleaner, more efficient, and more powerful Python based program. In this process, we were also able to add functionality that included Lagrangian capabilities and a variable vertical aggregation scheme. It is this work that formed the foundation for the design for the KPASA tool. For the KPASA tool, a design strategy was produced that highlighted the key chemical parameters needed for the improved ozone allocation. These formulations were tested and proven for various chemical environments.

Through the course of the project, we made improvements to the UNC model performance tool, pyPASS, and produced 1000+ plots integrating modeled and observational data. These plots allow for a holistic performance evaluation of the UH and ENVIRON modeling efforts. This database is maintained on a secure server at UNC.

1.4 Findings

Process analysis Tools Incorporation into APCA

Equations 1.1 and 1.2 represent the method by which the chemical parameters calculated by our process analysis tool would be used in the KPASA ozone allocation

algorithms [5]. In these formulas the terms, f_i and j_i , represent the fractions of the NO_x or VOC tracer present in the grid cell. The calculations for these new terms are shown in equations 1.3 and 1.4. Once the allocation is determined, the tracers are then used to partition the formation of ozone over the source regions for the precursors. With this formula, we are able to describe the local formation of ozone in terms of the NO_x and the VOCs from the different sources regions that were involved in its formation. Each chemical parameter in this equation has been verified with our process analysis tool to yield the correct values for ozone produced in various chemical environments.

$$(\text{new O}_3)_i = (f_i)(\text{NO}_2 \text{ prod.} + \text{NO}_2 \text{ src} + \text{NO}_2 \text{ ph. losses}) \left[\left(\frac{\text{NO}_2 + h\nu}{\text{NO}_2 \text{ avail}} \right) \left(\frac{\text{O}_3 \text{ prod}}{\text{NO}_2 + h\nu} \right) \right] \quad (1.1)$$

$$(\text{new O}_3)_i = (j_i)(\text{VOC reacted}) \left(\frac{\text{NO}_2 \text{ prod.}}{\text{VOC reacted}} \right) \left(\frac{\text{NO}_2 \text{ avail}}{\text{NO}_2 \text{ prod.}} \right) \left[\left(\frac{\text{NO}_2 + h\nu}{\text{NO}_2 \text{ avail}} \right) \left(\frac{\text{O}_3 \text{ prod}}{\text{NO}_2 + h\nu} \right) \right] \quad (1.2)$$

$$f_i = \left(\frac{\text{NO}_{xi}}{\sum \text{NO}_{xi}} \right) \quad (1.3)$$

$$j_i = \left(\frac{k[\text{OH}_i][\text{VOC}_i]}{\sum k[\text{OH}_i][\text{VOC}_i]} \right) \quad (1.4)$$

where i represents the source region for emission of the NO and VOC.

The implementation in of this method in CAMx would require the addition of only one FORTRAN routine that computes the six parameters in the two formulas from values in the IRR data structures for a given grid cell. These do not have to be stored or remembered. They would be allocated across the two ozone tracers based on the NO_x and VOC tracers that are in the grid cell from each source region.

Performance Evaluation Data

Dallas

The model tends to overpredict wind speeds, especially at night. The model does a reasonable job capturing wind direction, although it does struggle to reproduce conditions when wind speeds are less than $4 \frac{\text{km}}{\text{hr}}$. After careful analysis of the 850 plots generated by pyPASS we concluded that two episodes, June 18 - 23 and August 30 - September 9, had the highest ozone concentrations. These two focus periods also represented the early and late part of the ozone season.

Houston

The pyPASS tool was used to integrate the model and observed data and then generate time series plots, scatter plots, wind hodograms, bar plots, and tile plots for each simulation. The data that was generated by pyPASS is presented in the appendix of this document and available at <http://ftpozzone.sph.unc.edu/>.

Relevant findings from other Houston Modeling Work

UNC tested our process analysis algorithms on several modeling attempts of the Houston, TX non-attainment area. The Houston modeling attempts include efforts by the Texas state environmental agency and private consultants resulting in multiple simulations of the same modeling episode and domain. The pyPA tool provided a framework for an in-depth analysis of modeling data by quantifying several key chemical parameters such as radical budgets, source and fate of ozone precursors, and the physical processes that affect each species. Using CAMx4.2 with process analysis (PA) extensions, UNC has re-run a 2000 Houston/Galveston simulation using basecase inputs (base1b) prepared by TCEQ. The base1b inputs were generated as part of the TCEQ's regulatory modeling work on the 8-hour ozone standard attainment study for Houston Texas. The model's predictions are often bias high for NO₂ and VOC and yet it underpredicts observed O₃ peaks. Analysis of the PA outputs, especially the radical source budgets, revealed that much of the Houston area is radical limited, especially for organically-derived radicals. These results prompted the creation of a "radical source" sensitivity simulation studies. In this study, the model's emissions inputs were adjusted to determine the model's response to various types of increases in the radical pool.

The basecase model underpredicted observed HCHO concentrations, even when it over-predicted HRVOCs. The incomplete combustion of flares and mobile emissions were tested as two possible sources of formaldehyde that are either underrepresented or omitted from the current inventory. UNC added a total of 172.4 US tons across a three day period to 13 flares located mainly in the eastern part of Houston; this run will be referred to as 8hr.FORMeqFLVOC. To represent a missing formaldehyde source from mobile emissions, formaldehyde was scaled to 4% of CO emissions; this run will be referred to as 8hr.FORMeq0.04CO. The flare imputation had the greatest effect on peak ozone concentrations causing increases of more than 30 ppb. The 8hr.FORMeq0.04CO scenario increased ground level ozone by as much as 20 ppb. Although there is ample VOC reactivity available, the lack of model ·OH radicals results in a sizable portion of VOCs unreacted and unable to contribute to ozone formation. The NO_x bias further hinders ozone formation by terminating nearly a quarter of ·OH radicals. The model is capable of responding to changes in the radical bud-

get. The most significant impacts on ozone formation rates resulted from increases in sources of new $\cdot\text{OH}$ radicals.