Mining and Modeling for a Metropolitan Atlanta Ozone Pollution Decision-Making Framework

Zehua Yang
Abbott Laboratories, MS 2-30
1921 Hurd Drive, Irving, TX 75038

Victoria C. P. Chen
Department of Industrial & Manufacturing Systems Engineering
University of Texas at Arlington, Arlington, TX 76019

Michael E. Chang
School of Earth & Atmospheric Sciences
Georgia Institute of Technology, Atlanta, GA 30332

Terrence E. Murphy
6 Hunting Ridge, Hamden, CT 06518

Julia C. C. Tsai
Joy You Industrial Corporation
850 N. State Street, #5K, Chicago, IL 60610

COSMOS Technical Report 04-06

Abstract

In this paper, we present a decision-making framework (DMF) for reducing ozone pollution in metropolitan Atlanta. High concentrations of ozone at the ground-level continues to be a serious problem in several U.S. cities, and Atlanta is one of the most serious. In contrast to the "trial and error" approach utilized by state government decision makers, our DMF searches for *dynamic* and *focused* control strategies that require a lower total reduction of emissions than current control strategies. Our DMF utilizes a rigorous stochastic dynamic programming (SDP) formulation and includes an Atmospheric Chemistry Module to represent how ozone concentrations change over time. This paper focuses on the procedures within the Atmospheric Chemistry Module. Using the U.S. EPA's Urban Airshed Model for Atlanta, we employed mining and metamodeling tools to develop a computationally-efficient representation of the relevant ozone air chemistry. These effectively model the transitions over time within the SDP optimization.

Key words
Data Mining, Metamodeling, Decision-making, Ozone Pollution, Stochastic Dynamic Programming, Urban Airshed Model

1. Introduction

The increasing concentrations of ground-level ozone in the urban (and often rural) atmosphere continue to be one of the major environmental issues today. Ozone is considered a harmful pollutant because of its detrimental effects on humans and the natural ecosystem. Ozone is one of the U.S. EPA's six criteria pollutants and is regulated by a National Ambient Air Quality Standard. Under the former U.S. EPA standard which defines an *ozone episode* as an hourly-averaged concentration exceeding 0.12 parts per million (ppm), 77 nonattainment areas in the U.S. were defined, including Atlanta, which is categorized as a *serious* non-attainment area. (The new U.S. EPA standard, which defines an episode as an eight-hour average exceeding 0.08 ppm, is more stringent).

Ground-level ozone is not emitted directly into the air but is created by a complex series of reactions involving nitrogen oxides (**NOx** = $NO + NO_2$) and volatile organic compounds (**VOC**s) in the presence of sunlight (Sillman et al. 1995). The primary sources of NOx are power plants, automobiles and industry. VOCs have both anthropogenic sources (cars, industry) and natural sources (vegetation). Therefore, in order to control ground-level ozone it is necessary to control emissions of NOx and VOCs. Ozone does exist in the clean atmosphere, with surface concentrations of about 20 parts per billion parts of air (20 ppb). The natural source is due to mixing with the upper atmosphere, where ozone is chemically produced to form an ozone layer that protects the earth's surface from harmful UV-B radiation. However, ground-level ozone concentrations are often above 0.20 ppm, especially in urban areas, due to accumulations of NOx and VOCs. Sunlight and heat are the catalysts for the chemical reaction that creates ozone; thus, ozone concentrations build during the course of a day and fall as the sun goes down.

As the first of its kind of research, our objective is to develop a computationally-tractable and rigorous decision-making framework (**DMF**) that searches for *dynamic* and *focused* control strategies for reducing ozone pollution. Government decision-makers typically utilize "trial and error" to test control strategies, and typical control strategies involve an overall reduction in emissions and may require exorbitant (and impossible) reductions to maintain the U.S. EPA limits. Instead, it may be more practical to consider focused reductions that dynamically depend on the conditions of the day. Our DMF conducts an intelligent and comprehensive search through the possibilities using methods in statistics and optimization.

For the Atlanta application, we are studying the ozone episode from July 29 – August 1, 1987, which remains to be one of the worst on record. To date, no practical control strategy using typical approaches has been identified to handle this episode, and our ultimate objective is to use our DMF to see if a practical control strategy is at all achievable. A prototype for the Atlanta ozone pollution DMF is in progress, and in this paper we describe the difficult task of efficiently representing ozone pollution air chemistry within a rigorous optimization formulation, using mining and metamodeling tools.

In the next section, we describe the DMF formulation. Section 3 provides details on the Mining and Metamodeling Phases applied to the Atlanta ozone pollution problem. Section 4 presents our preliminary results for efficiently representing ozone pollution in our DMF, and Section 5 discusses refinement of this work. Section 6 presents a verification of the ozone transition metamodels. Finally, concluding remarks are given in Section 7.

2. DMF Formulation

The modules of our DMF are illustrated in Figure 1. To accommodate continuous variables, time dependencies, and uncertainty, the DMF is based on a rigorous continuous-state stochastic dynamic programming (**SDP**) approach (Chen et al. 1999, Chen 1999, Tsai et al. 2004). The *objective* of SDP is to minimize expected "cost" subject to certain *constraints* over several *stages* $t = 1,\ldots,T$, i.e., to solve:

$$\begin{aligned} &\min_{u_1,\ldots u_T} E\left\{\sum_{t=1}^{T} c_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{\varepsilon}_t)\right\} \\ &\text{s. t.} \quad \boldsymbol{x}_{t+1} = f_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{\varepsilon}_t), \text{ for } t = 1,\ldots,T-1 \text{ and} \\ &\qquad (\boldsymbol{x}_t, \boldsymbol{u}_t) \in \Gamma_t, \text{ for } t = 1,\ldots,T \end{aligned} \tag{1}$$

where $\boldsymbol{x}_t$ is the *state* vector, $\boldsymbol{u}_t$ is the *decision* vector, $\boldsymbol{\varepsilon}_t$ is the stochastic component, $\boldsymbol{x}_{t+1}$ is determined by the *transition functions* $f_t(\bullet)$, Γ_t represents capacity constraints, and the cost function $c_t(\bullet)$, contains both the economic and penalty costs. The SDP stages are based on pre-specified *time periods*. The *state variables* represent the current state of the system. Continuous-state SDP assumes the state variables are continuous. The *decision variables* are the variables one controls to minimize cost. The *transition functions* define how the state variables change over time, specifically from one SDP stage to the next. A computationally-tractable solution method for continuous-state SDP has been developed by Chen et al. (1999).

For the ozone pollution SDP, our objective is to minimize the cost of avoiding ozone episodes, where the U.S. EPA standard specifies a constraint on ozone concentrations. Because of ozone's daily cycle, our time horizon is one day, with time periods involving groups of hours (e.g., 7:00 AM to 10:00 AM). The state variables $\boldsymbol{x}_t$ at a given time potentially include concentrations of ozone, NOx, and VOC at various spatial locations across the metropolitan Atlanta region. Similarly, the decision variables $\boldsymbol{u}_t$ potentially include emissions of NOx and VOC at various locations and times over the course of the day.

SDP transition functions $f_t(\bullet)$ are constructed by the Atmospheric Chemistry Module in Figure 1. This is the most difficult task for the DMF. Within this module is the representation of ozone pollution air chemistry. An advanced three-dimensional, photochemical air quality grid model like the Urban Airshed Model (UAM, EPA 1990) may be used to calculate transitions; however, the UAM simulation is computationally intensive. Thus, a more efficient approach is needed for incorporation into SDP optimization. In the remainder of this paper, we describe a two-phase *mining* and *metamodeling* approach (Chen et al. 2003) for constructing these SDP transition functions.

3. Atmospheric Chemistry Module for Atlanta

The UAM modeling domain for Atlanta encompasses a 160×160 kilometer square region containing the metropolitan area, shown in Figure 2. UAM ozone computations are hourly and occur on a 40×40 grid over the modeling domain. Emissions are represented in the Atlanta UAM by 102 *point sources* (e.g., power plant smoke stacks) and other sources within each 40×40 grid region. A 24-hour UAM simulation run takes more than one hour. The first two days of the given Atlanta ozone episode are "ramp-up" days to initialize the UAM. The actual ozone exceedances occurred on the third and fourth days. We consider here only the third day,

i.e., July 31, 1987, because if we cannot control ozone on the third day, then the fourth day is a lost cause.

To reduce ozone concentrations, NOx and VOC emissions must be reduced. Atlanta, in particular, is "NOx-limited," which means that targeting VOC emissions is not effective (Chameides et al. 1988). For our Atlanta DMF, we chose to ignore VOC emissions, so our state variables included only ozone concentrations and NOx emissions. However, if we included ozone and NOx at every point source (102) and every grid region (40×40) and each hour (24), the number of state variables would still be exceedingly large. Thus, in order to achieve a computationally-tractable DMF, a critical component of the Atmospheric Chemistry Module is dimension reduction, conducted by three phases in order: Initialization, Mining, and Metamodeling. In the Initialization Phase, we devise the following setup:

a) Aggregate the UAM 40×40 grid into a 5×5 grid (25 grid squares). We refer to these grid squares with the notation (i, j), for $i = 1, 2, 3, 4, 5$ and $j = 1, 2, 3, 4, 5$, where grid square (1, 1) is at the bottom left, and grid square (2, 1) is one square to the east, etc.
b) Reduce the 24-hour time horizon to the critical time horizon of 4:00 AM to 7:00 PM. Define five time periods: (0) 4:00 AM – 7:00 AM, (1) 7:00 AM – 10:00 AM, (2) 10:00 AM – 1:00 PM, (3) 1:00 PM – 4:00 PM, (4) 4:00 PM – 7:00 PM. Time period 0 is an initialization period.
c) Control maximum ozone concentrations only at the four Photochemical Assessment Monitoring Stations (PAMS). Only these four stations are monitored by the U.S. EPA.

NOx emissions are considered separately in the DMF for each of the grid squares and point sources during each of the time periods 1–4. The SDP state vector must include any variables that are related to maximum ozone at the four stations. Because ozone and emissions occurring as early as time period 0 (perhaps even earlier) could affect maximum ozone in the last time period, the number of possible state variables entering the last period is

$$(4 \text{ time periods}) \times (4 \text{ stations} + 25 \text{ regions} + 102 \text{ point sources}) = 524.$$

In the Mining Phase, data collected from the UAM was used to identify those point sources and grid squares that had the most influence on maximum ozone concentrations at the four PAMS sites. In the Metamodeling Phase, the SDP transition functions were constructed, and in the process further dimension reduction would be achieved. These two phases are described below.

3.1 Mining Phase

We used a 149-point Latin hypercube experimental design to scale NOx emissions in different regions and at different point sources from zero up to the nominal level. Different time periods were not explored. These were input into the UAM and resulting ozone concentrations across the 40×40 UAM modeling grid were collected. Using our aggregated 5×5 grid, the maximum ozone concentration was determined for each grid square containing a PAMS site. Mining via regression analysis was conducted separately for each PAMS site.

Of the 25 grid regions, 16 were statistically significant. Of the original 102 point sources only 15 were statistically significant. In Figure 3, all point sources (some points overlap) and the four PAMS sites (Conyers, S. Dekalb, Tucker, and Yorkville) are shown on the aggregated 5×5 grid. The shaded regions indicate those grid squares that were <u>not</u> significant. Note that

these are all located along the edges of the modeling domain, where one would expect to find the least influence. The labeled triangles mark the significant point sources (# 1-6, 9, 12, 15, 21, 23, 30, 37, 63, 64). In the Metamodeling Phase, we chose to maintain only these 15 point sources due to the tremendous reduction in dimension; however, all 25 grid squares were maintained.

3.2 Metamodeling Phase

The metamodeling process is illustrated in Figure 4. The experimental design specifies reductions in the emissions at various *regions* and *times*. Then the UAM is used to calculate the resulting ozone concentrations, from which maximum ozone for the four PAMS sites at various *times* is calculated. Based on the input NOx emissions and the output from the UAM, statistical methods are applied to build the metamodels that represent the ozone SDP transition functions. Through this process, further dimension reduction would be achieved.

First we specified the initial set of potential state and decision variables to be studied. For SDP stage t (t = 1, 2, 3, 4), the initial set of potential *decision* variables consisted of:

- Total NOx emissions over each grid square and at each point source during time period t: (25+15) = 40 variables.

A critical property of SDP is that the state variables specified at the beginning of stage t must contain all the required information to do all the calculations in stage t. For example, suppose the maximum ozone at Conyers occurring during time period 4 (4:00 PM – 7:00 PM) depends on the maximum ozone observed at Conyers in time periods 2 and 3, and the NOx emissions occurring in the center grid square (3,3) during time periods 1 and 3. Then the set of state variables at the beginning of SDP stage 4 must include information from time periods 1, 2, and 3. This is why we make the distinction between *time periods* and *SDP stages*. Thus, for SDP stage t (t = 1, 2, 3, 4), the initial set of potential *state* variables consisted of:

- Maximum hourly-averaged ozone at the four PAMS sites in time periods 0 though t–1. For SDP stage t = 4: 4*4 = 16 variables.
- Total NOx emissions over each grid square and at each point source during time periods 0 though t–1. For SDP stage t = 4: 4*(25+15)= 160 variables.

Thus, the initial number of state variables is as follows: 44 in SDP stage 1, 88 in SDP stage 2, 132 in SDP stage 3, and 176 in SDP stage 4. Since the largest continuous-state SDP problem in the literature involved 20 state variables (Tsai et al. 2004), dimension reduction is a must.

For the metamodeling process, the NOx emissions over the 25 grid squares, 15 point sources, and 5 time periods constituted a set of 200 design variables. A 500-point Latin hypercube experimental design was utilized to determine the NOx emissions for these 200 variables. The maximum ozone at the four PAMS sites in time periods 1, 2, 3, 4 constituted a set of 16 response variables. Separate regression metamodels were then constructed for each of the response variables. Specifically, the metamodel for maximum ozone in time period t could be a function of maximum ozone in earlier time periods and NOx emissions from time period t and earlier. A regression coefficient of determination (R^2) above 90% was achieved for all but two metamodels. These two, the Tucker and S. Dekalb sites in time period 2, will require further refinement in later work. The S. Dekalb metamodel is discussed further in Section 5.

4. Reduction of the State and Decision Variables

In each SDP stage we only need to maintain those state and decision variables required by the transition function metamodels for that stage and later stages. This greatly reduces the dimension of the SDP system. From the initial sets of potential state and decision variables listed in the previous section, Tables 1 and 2 list the required sets of variables. Define the following notation used in the tables:

p:	3-hour time period; p=0, 1, 2, 3, 4.
S:	PAMS site; S = C (Conyers), SD (South Dekalb), T (Tucker), Y (Yorkville)
O_p^S:	Maximum ozone at PAMS site S in time period p
(i, j):	square region in the 5×5 grid with coordinate (i, j), where i = 1, 2, 3, 4, 5 and j = 1, 2, 3, 4, 5, as defined in Section 3.
$E_p^{(i,j)}$:	Emission quantity within square (i, j) in time period p
pt(k):	point sources in the modeling domain; $k \in \{1, 2, 3, 4, 5, 6, 9, 12, 15, 21, 23, 30, 37, 63, 64\}$.
$E_p^{\mathrm{pt}(k)}$:	Emission quantity at point source pt(k) in time period p.

For each SDP stage, the variables from different time periods p are listed in different rows, and the different types of variables (e.g., grid squares vs. point sources) are listed in different columns. The column labeled "Carried Over" lists variables that were not required in the current SDP stage, but were needed in later stages. For example from Table 1, $E_0^{(3,2)}$ is not needed to do calculations in SDP stages 1 or 2, but is needed in stage 3. Thus, we must save $E_0^{(3,2)}$ until it is used in SDP stage 3.

Looking at Table 1, instead of up to 176 state variables, the maximum number of state variables required is 25 in SDP stage 2. Similarly, looking at Table 2, instead of up to 40 decision variables, the maximum is 17 in SDP stage 1. Additional information can be garnered from the set of sufficient decision variables in Table 2. The included grid squares indicate the key regions whose NOx emissions influence ozone pollution. Finally, the fact that the largest set of decision variables is in SDP stage 1 emphasizes time period 7:00 AM – 10:00 AM as the most important time period for NOx emissions reductions. This corresponds to the morning rush hour and agrees with the consensus among air quality decision-makers.

5. Refinement of the Transition Function Metamodels

As mentioned at the end of Section 3, two transition function metamodels were unsatisfactory. The S. Dekalb metamodel for time period 2 achieved an R^2 of only 5.62%. To refine this metamodel, multivariate adaptive regression splines (**MARS**, Friedman 1991) were fit to the UAM data. In particular, we employed an implementation of MARS that utilizes automatic selection of the number of basis functions and gives priority to lower-order terms (Tsai and Chen 2005). This version of MARS achieved an R^2 of 98.58% with 9 univariate and 10 two-way interaction basis functions. Figure 5 illustrates the structure of the MARS metamodel as function of NOx emissions in grid square (3, 3) during time periods 1 and 2. Note from Figure 3, that the S. Dekalb PAMS site is located in this center grid square; thus, these NOx emissions would be most relevant.

Unlike those regression metamodels that achieved high R^2 values, the S. Dekalb MARS metamodel involves curvature that is potentially detrimental to the SDP optimization. The optimization within SDP seeks to reduce NOx emissions in order to reduce ozone concentrations. In Figure 5, if we follow the trend along the NOx emissions in time period 1, we see that at the high end of NOx emissions, ozone actually decreases. It is a known in atmospheric chemistry that if NOx concentrations are very high such that mixing with VOCs does not occur, then ozone cannot form and consequently decreases (Chameides et al. 1988). While it is valuable to model the atmospheric chemistry accurately, it is more important to calibrate the SDP optimization to find a solution that is reasonable. In this case, it would be *unreasonable* to suggest a control strategy that increases NOx to decrease ozone. Thus, in future work, we propose to remove this counterproductive structure from any metamodels. Procedures for doing this are currently in development.

6. Verification of the Metamodels

To verify the metamodels obtained in Section 3, two groups of comparisons were made. In the first comparison, the metamodels were used to predict maximum ozone concentrations during time periods 1 to 4 for eleven overall emission reduction strategies. In the second comparison, metamodels were used to predict maximum ozone for 50 randomly generated hypothetical scenarios. The daily maximum ozone concentrations from the UAM simulation were taken as the "true value." The error in each daily maximum ozone concentration predicted by the metamodels was measured by the absolute deviation percentage, which is defined as:

$$\frac{\left|\text{daily maximum by metamodels} - \text{true value}\right|}{\text{true value}} \times 100\% .$$

The results for the first comparison are given in Table 3. The absolute deviation percentages for the 50 hypothetical scenarios ranged from 0.2% to 4.6% with a mean of 2.5%.

Table 3 shows that when emission reduction is high, metamodel prediction tends to be lower than the UAM simulation; and vice-versa when emission reduction is low. Metamodel prediction error is higher in extreme scenarios. The lower absolute deviation percentages for the 50 hypothetical scenarios could be due to the fact the chances are small for one scenario to be close to an extreme case. Because the primary objective is identification of control strategies and not the accuracy of the metamodels, validation of the metamodels will be appropriately inferred by the effectiveness of control strategies in the UAM.

7. Concluding Remarks

A procedure for the Atmospheric Chemistry Module for an ozone pollution DMF has been established and implemented successfully on an ozone episode in metropolitan Atlanta. Important computational aspects, such as dimension reduction and efficient modeling were addressed. Currently, we are employing the constructed transition function metamodels from Section 3 with SDP to solve for potential control strategies. These control strategies will then be tested directly with the Atlanta UAM to verify their effectiveness. Future work must address refinement of the transition function metamodels to produce SDP optimization solutions that will lead to reasonable control strategies.

Acknowledgments

This research was supported by a Technology for Sustainable Environment (TSE) grant under the U. S. Environmental Protection Agency's Science to Achieve Results (STAR) program (Contract #R-82820701-0). We would also like to thank Paul Brooks and Hilda Valero for their statistical analysis work related to this research.

References

- Chameides, W. L. and Cowling, E. B. (1995) *The State of the Southern Oxidants Study (SOS): Research Accomplishments and Future Plans*, North Carolina State University, Raleigh, NC. http://www2.ncsu.edu:8010/ncsu/CIL/southern_oxidants/docs/intro.pdf.
- Chameides, W. L. and Kasibhatla, P. S. (1994) Growth of continental-scale metro-agro-plexes, regional ozone pollution, and world food production. *Science*, **264(5155)**, 74–77. http://www2.ncsu.edu:8010/ncsu/CIL/southern_oxidants/docs/intro.pdf.
- Chameides, W. L., Lindsay, R. W., Richardson , J., Kiang, C. S. (1988) The rise of biogenic hydrocarbons in urban photochemical smog – Atlanta as a case-study. *Scienc*e, **241(4872)**, 1473–1475.
- Chang, M. E. and Cardelino, C. (2000) Application of the Urban Airshed Model to forecasting next-day peak ozone concentrations in Atlanta, Georgia, *Journal of the Air & Waste Management Association*, **50**, 2010–2024.
- Chen, V. C. P. (1999) Application of orthogonal arrays and MARS to inventory forecasting stochastic dynamic programs. *Computational Statistics and Data Analysis*, **30**, 317–341.
- Chen, V. C. P., Ruppert, D., and Shoemaker, C. A. (1999) Applying experimental design and regression splines to high-dimensional continuous-state stochastic dynamic programming. *Operations Research*, **47**, 38–53.
- Chen, V. C. P., Tsui, K.-L., Barton, R. R., and Allen, J. K. (2003) A review of design and modeling in computer experiments. *Handbook in Statistics: Statistics in Industry*, **22**, 231–261, Khattree, R. and Rao, C. R. (eds.), Elsevier Science, Amsterdam.
- Friedman, J. H. (1991) Multivariate adaptive regression splines (with discussion), *Annals of Statistics*, **19**, 1–141.
- Sillman, S., Al-Wali, K., Marsik, F. J., Nowatski, P., Samson, P. J., Rodgers, M. O., Garland, L. J., Martinez, J. E., Stoneking, C., Imhoff, R. E., Lee, J-H., Weinstein-Lloyd, J. B., Newman, L. and Aneja, V. (1995) Photochemistry of ozone formation in Atlanta, GA: models and measurements. *Atmospheric Environment*, **29**, 3055–3066.
- SIP (2001) *Georgia's State Implementation Plan for the Atlanta Ozone Non-attainment Area*, Environmental Protection Division, Georgia Department of Natural Resources, Atlanta, Georgia, 2001, http://www.state.ga.us/dnr/environ/plans_files/plans/airsipsum.pdf
- Tsai, J. C. C. and Chen, V. C. P. (2005) Flexible and robust implementations of multivariate adaptive regression splines within a wastewater treatment stochastic dynamic program. *Quality and Reliability Engineering International*, **21**, 689–699.
- Tsai, J. C. C., Chen, V. C. P., Chen, J., and Beck, M. B. (2004) Stochastic dynamic programming formulation for a wastewater treatment decision-making framework. *Annals of Operations Research*, **132**, 207–221.
- U.S. EPA (1990) *User's Guides for the Urban Airshed Model*, EPA-450/4-90-007A-E.

Table 1: Monitored state variables.

SDP stage 1: 17 state variables				
p	**5×5 grid squares**	**Point Sources**	**Max O_3**	**Carried Over**
0	$E_0^{(1,3)}$ $E_0^{(2,3)}$ $E_0^{(2,5)}$ $E_0^{(4,3)}$ $E_0^{(5,3)}$	$E_0^{\mathrm{pt}(30)}$ $E_0^{\mathrm{pt}(64)}$	O_0^{C} O_0^{SD} O_0^{T}	$E_0^{(3,2)}$ $E_0^{(3,3)}$ $E_0^{(4,2)}$ $E_0^{(4,4)}$ $E_0^{\mathrm{pt}(5)}$ $E_0^{\mathrm{pt}(15)}$ $E_0^{\mathrm{pt}(63)}$

SDP stage 2: 25 variables				
p	**5×5 grid squares**	**Point Sources**	**Max O_3**	**Carried Over**
0	$E_0^{(1,3)}$ $E_0^{(2,3)}$ $E_0^{(4,2)}$ $E_0^{(4,3)}$ $E_0^{(4,4)}$	$E_0^{\mathrm{pt}(5)}$ $E_0^{\mathrm{pt}(15)}$ $E_0^{\mathrm{pt}(63)}$		$E_0^{(3,2)}$ $E_0^{(3,3)}$
1	$E_1^{(1,3)}$ $E_1^{(1,4)}$ $E_1^{(3,2)}$ $E_1^{(3,5)}$ $E_1^{(4,2)}$ $E_1^{(4,3)}$ $E_1^{(5,4)}$		O_1^{C} O_1^{SD} O_1^{T} O_1^{Y}	$E_1^{(2,4)}$ $E_1^{(3,3)}$ $E_1^{(3,4)}$ $E_1^{\mathrm{pt}(6)}$

SDP stage 3: 23 variables				
p	**5×5 grid squares**	**Point Sources**	**Max O_3**	**Carried Over**
0	$E_0^{(3,2)}$ $E_0^{(3,3)}$			
1	$E_1^{(1,3)}$ $E_1^{(1,4)}$ $E_1^{(3,2)}$ $E_1^{(3,3)}$	$E_1^{\mathrm{pt}(6)}$	O_1^{SD} O_1^{Y}	$E_1^{(2,4)}$ $E_1^{(3,4)}$
2	$E_2^{(1,3)}$ $E_2^{(1,4)}$ $E_2^{(3,2)}$ $E_2^{(3,3)}$ $E_2^{(4,3)}$		O_2^{C} O_2^{SD} O_2^{T} O_2^{Y}	$E_2^{(2,4)}$ $E_2^{(3,4)}$ $E_2^{\mathrm{pt}(5)}$

SDP stage 4: 19 variables				
p	**5×5 grid squares**	**Point Sources**	**Max O_3**	**Carried Over**
1	$E_1^{(1,4)}$ $E_1^{(2,4)}$ $E_1^{(3,3)}$ $E_1^{(3,4)}$			
2	$E_2^{(1,4)}$ $E_2^{(2,4)}$ $E_2^{(3,2)}$ $E_2^{(3,4)}$	$E_2^{\mathrm{pt}(5)}$		
3	$E_3^{(1,4)}$ $E_3^{(3,2)}$ $E_3^{(3,3)}$ $E_3^{(4,2)}$	$E_3^{\mathrm{pt}(3)}$ $E_3^{\mathrm{pt}(4)}$ $E_3^{\mathrm{pt}(6)}$	O_3^{C} O_3^{SD} O_3^{T}	

Table 2: Decision variables.

SDP stage 1: 17 Decision variables		
5×5 grid squares	**Point Sources**	**Carried Over**
$E_1^{(1,2)}$ $E_1^{(1,3)}$ $E_1^{(1,4)}$ $E_1^{(2,3)}$ $E_1^{(3,1)}$ $E_1^{(3,3)}$ $E_1^{(4,2)}$ $E_1^{(4,3)}$ $E_1^{(4,4)}$	$E_1^{\mathrm{pt}(4)}$ $E_1^{\mathrm{pt}(5)}$	$E_1^{(2,4)}$ $E_1^{(3,4)}$ $E_1^{(3,2)}$ $E_1^{(3,5)}$ $E_1^{(5,4)}$ $E_1^{\mathrm{pt}(6)}$

SDP stage 2: 9 decision variables		
5×5 grid squares	**Point Sources**	**Carried Over**
$E_2^{(1,3)}$ $E_2^{(1,4)}$ $E_2^{(3,2)}$ $E_2^{(3,3)}$ $E_2^{(4,2)}$ $E_2^{(4,3)}$		$E_2^{(2,4)}$ $E_2^{(3,4)}$ $E_2^{\mathrm{pt}(5)}$

SDP stage 3: 9 decision variables		
5×5 grid squares	**Point Sources**	**Carried Over**
$E_3^{(1,3)}$ $E_3^{(3,2)}$ $E_3^{(3,3)}$ $E_3^{(4,2)}$ $E_3^{(4,3)}$		$E_3^{(1,4)}$ $E_3^{\mathrm{pt}(3)}$ $E_3^{\mathrm{pt}(4)}$ $E_3^{\mathrm{pt}(6)}$

SDP stage 4: 3 decision variables		
5×5 grid squares	**Point Sources**	**Carried Over**
$E_4^{(1,3)}$ $E_4^{(1,4)}$ $E_4^{(3,3)}$		

Table 3: Percentage absolute deviations for eleven overall emission reduction strategies.

Emission Reduction %	UAM	Metamodels	Absolute Deviation %
100%	1.05E-01	9.44E-02	10.00%
90%	1.12E-01	1.02E-01	9.08%
80%	1.19E-01	1.10E-01	8.13%
70%	1.26E-01	1.17E-01	6.77%
60%	1.32E-01	1.25E-01	5.10%
50%	1.37E-01	1.33E-01	3.16%
40%	1.43E-01	1.40E-01	1.78%
30%	1.48E-01	1.48E-01	0.07%
20%	1.52E-01	1.56E-01	2.36%
10%	1.55E-01	1.63E-01	5.07%
0%	1.58E-01	1.71E-01	8.11%
Mean Absolute Deviation %			5.42%

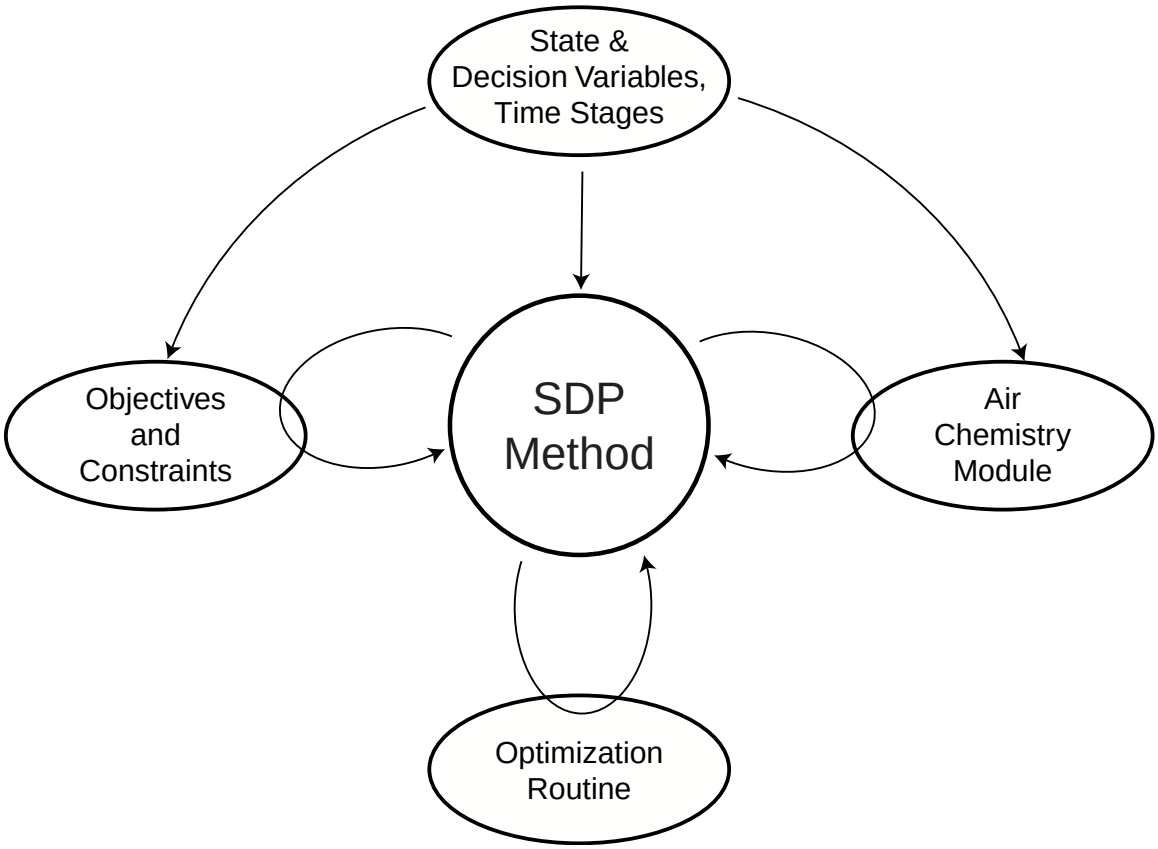

Figure 1: A modular decision-making framework for ozone pollution.

Figure 2: Location of UAM domain for Atlanta, Georgia (SIP 2001).

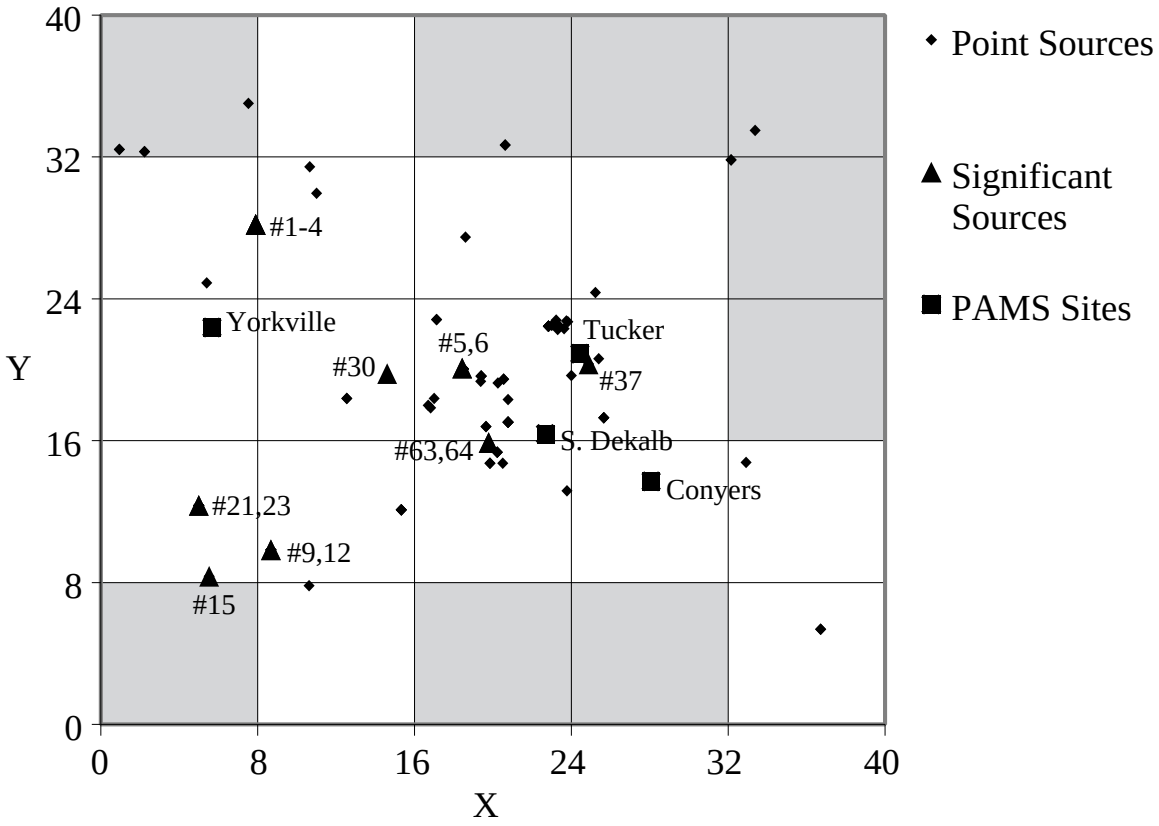

Figure 3: From the Mining Phase: Aggregated 5×5 grid and all point sources over the 40×40 UAM grid.

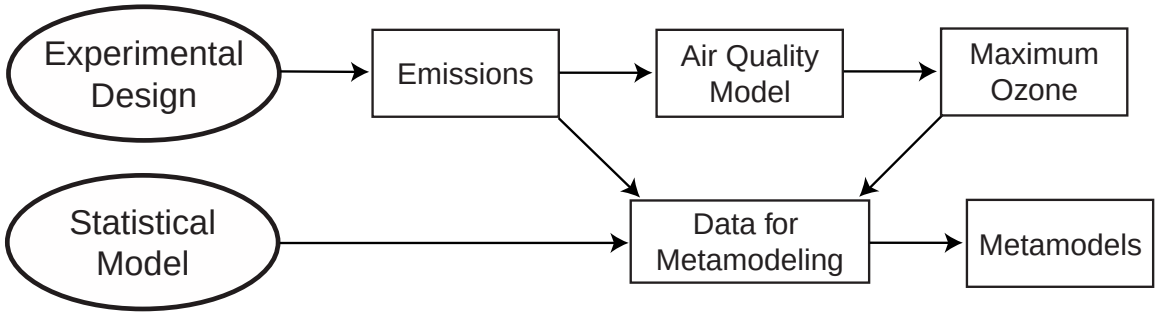

Figure 4: Flow chart describing the Metamodeling Phase.

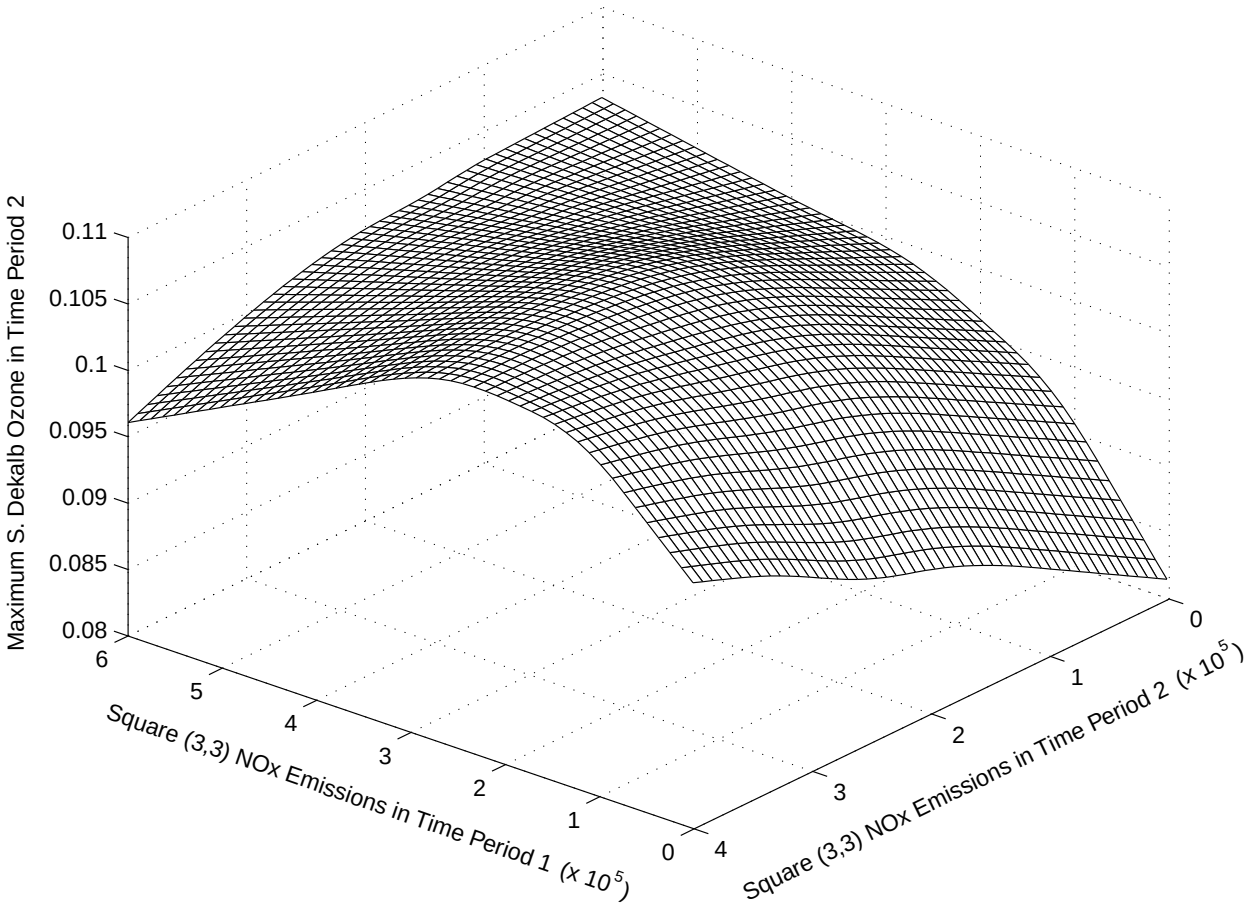

Figure 5: Transition function metamodel for maximum ozone concentrations at the S. Dekalb monitoring station during the 10:00 AM 1:00 PM time period, shown as a function of NOx emissions during the 7:00 AM 10:00 AM and 10:00 AM 1:00 PM time periods in the square region that contains the S. Dekalb monitoring station.