A Decision-Making Framework for Ozone Pollution Control

COSMOS Technical Report 05-04

Zehua Yang

Abbott Laboratories, MS 2-30, 1921 Hurd Drive, Irving, TX 75038, U.S.A., zehua_yang@hotmail.com

Victoria C. P. Chen

Department of Industrial & Manufacturing Systems Engineering, The University of Texas at Arlington, Arlington, Texas 76019, U.S.A., vchen@uta.edu

Michael E. Chang

School of Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A., chang@eas.gatech.edu

Melanie L. Sattler

Department of Civil and Environmental Engineering, The University of Texas at Arlington, Arlington, Texas 76019, U.S.A., msattler@uta.edu

Aihong Wen

PROS Revenue Management, 3100 Main St. Suite 900, Houston, Texas 77002, U.S.A., awen@prosrm.com

Subject Classification: Environment. Dynamic Programming: Applications. Statistics: Data analysis.

Area of Review: Environment, Energy and Natural Resources

Abstract

In this paper, an intelligent Decision-Making Framework (DMF) is developed to help decision-makers identify cost-effective ozone control policies. High concentrations of ozone at the ground-level continue to be a serious problem in numerous U.S. cities. Our DMF searches for *dynamic* and *targeted* control policies that require a lower total reduction of emissions than current control strategies based on the "trial and error" approach typically employed by state government decision-makers. Our DMF utilizes a rigorous stochastic dynamic programming (SDP) formulation and incorporates an Atmospheric Chemistry Module to model how ozone concentrations change over time. Within the Atmospheric Chemistry Module, methods from design and analysis of computer experiments are employed to create SDP state transition equation metamodels, and critical dimensionality reduction is conducted to reduce the state space dimension in solving our SDP problem. Results are presented from a prototype DMF for the Atlanta metropolitan region.

1. Introduction

The increasing concentrations of ozone in the urban (and often rural) atmosphere are a major environmental issue. Ozone is considered to be a harmful pollutant because of its detrimental effects on humans and the natural ecosystem. The Clean Air Act empowers the U.S. Environmental Protection Agency (EPA) to set the National Ambient Air Quality Standards (NAAQS). Ozone is one of the EPA's six criteria pollutants regulated by NAAQS (http://www.epa.gov/air/criteria.html). The *one-hour EPA standard* defines an ozone exceedance as a daily peak hourly averaged ozone concentration that exceeds 0.12 parts per million (ppm). Given that modern ambient ozone monitoring devices are capable of precisely measuring ozone to within 0.001 ppm, the U.S. EPA adopted a standard rounding convention in the practical application of the NAAQS such that ozone concentrations of 0.124 ppm are considered to meet the NAAQS, and concentrations of 0.125 ppm are considered to exceed the NAAQS. An area is determined to be attaining the standard when the number of exceedance days per year, averaged over

three years, is less than or equal to one. For the stricter *8-hour EPA standard*, an area is attaining the standard when the three-year average of the annual fourth highest daily 8-hour average ozone concentration does not exceed 0.08 ppm (0.084 ppm with rounding). Additional non-attainment areas have been identified using the 8-hour standard. Non-attainment areas are required to improve air quality to the standard set by the EPA.

Ground-level (tropospheric) 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). Primary sources of NOx are power plants, automobiles and industry. VOCs have a wide variety of both anthropogenic sources (combustion, solvents) and natural sources (vegetation). In order to control ground-level ozone, it is typically necessary to control emissions of NOx and VOCs. Ozone exists naturally in the atmosphere. In the stratosphere, high concentrations of ozone protect the earth's surface from harmful ultraviolet radiation emitted by the sun. Much lower concentrations of ozone are found in the unpolluted troposphere resulting from a complex series of natural photochemical reactions and stratospheric intrusions. In the polluted troposphere, typically around urbanized areas where emissions of anthropogenic sources are greatest, photochemical reactions involving NOx and VOCs may increase ground-level ozone concentrations considerably. Such increases are highest in the summer when sunlight and heat increase the rate of photochemical formation; thus, ozone concentrations build during the course of a day and fall as the sun goes down. Ozone controls seek to reduce VOCs and/or NOx, depending on the characteristics of a non-attainment area, although nonlinearities in the air chemistry can occur when very high concentrations of NOx lead to lower ozone in local areas.

Due to legislative efforts and increasing concerns for air quality, many research studies have been conducted on ozone pollution control. Complex 3-D air quality photochemical models (e.g., Urban Airshed Model, U.S. EPA 1990; Comprehensive Air quality Model, http://www.camx.com/) have been developed to simulate air pollution emissions, chemical reactions, and atmospheric transport, in order to predict ozone concentrations and help government decision-makers evaluate control strategies. Current

approaches in the U.S. predominantly employ "trial and error" to identify "across-the-board" control strategies. Here "across-the-board" means that the emission reduction percentage of NOx or VOCs must remain the same for all emission sources across the entire urban area, 24 hours a day. Identifying promising control strategies is a challenge due to many issues, including the complex mechanisms of ozone formation, large numbers of emission sources in the target geographical domain, and conflicting social and political interests. Previous work in this area includes Seinfeld and Kyan (1971), Trijonis (1974), Kyan and Seinfeld (1974), and Loughlin et al. (2000).

Seinfeld and Kyan (1971) proposed a pioneering framework for determination of air pollution control strategies in 1970s. In this framework, a linear programming (**LP**) model was formulated and solved to optimally select control options which would reduce the total mass emissions to a projected level. Then a simple model based on the concept of well-mixed cells was employed to relate the total amount of emissions to the concentrations of a pollutant such as ozone. By solving these two problems for different values of total mass emissions, the optimal control strategies for different values of pollutant concentrations were obtained. Kyan and Seinfeld (1974) further extended the decoupled two-stage optimization model of Seinfeld and Kyan (1971) to an integrated non-linear optimization model. Trijonis (1974) used a simplified isopleth diagram approach to estimate the minimum required across-the-board VOCs and NOx reductions to achieve a target level of ozone, then solved an LP model to identify the best control strategies to achieve the minimum required emission reductions. Loughlin et al. (2000) applied genetic algorithms to search for cost-effective control strategies based on a simple computer air chemistry model, Ozone Isopleths Plotting Package – Research Version.

Most of the methods mentioned above aim to optimize control strategies in which the required emission reductions vary with emission sources, but none of them addressed the fact that the minimum required emission reductions need not be constant across different times of day to achieve the same results. By allowing emission reductions to change with both source location and time of a day, we may obtain more cost-effective control strategies than those that only account for the location of emissions. In

addition, simple analytical air chemistry models lack the capability of accounting for the complex terrain of target domain, complex weather conditions, and locality of emissions (Seinfeld 1988).

By contrast, our decision making framework (**DMF**) employs a comprehensive 3-D air chemistry photochemical model to search for cost-effective emission reduction strategies that *dynamically* vary over time, depending on the initial conditions of the day, and are *targeted* in critical locations and time periods. To accommodate continuous variables, time dependencies, and uncertainty, the DMF is based on a rigorous continuous-state stochastic dynamic programming (**SDP**) approach (e.g., Foufoula-Georgiou Kitanidis 1988, Chen et al. 1999). The advantage of dynamic programming (Bellman 1957) is its ability to model a system changing over time, and it has been applied to a number of environmental engineering problems (e.g., Shoemaker 1982, Yakowitz 1982, Bouzaher et al. 1990, Johnson et al. 1993, Mhaisalkar et al. 1993, Seppelt 2001, Tsai et al. 2004). Our DMF is illustrated in Figure 1. First, potential SDP state variables, decision variables, and time stages must be identified. These are then employed to construct the desired SDP cost objectives and constraints. More importantly, the Atmospheric Chemistry Module encompasses many steps using the comprehensive air chemistry photochemical model to identify the key state and decision variables, and ultimately estimate the SDP state transition equations. Finally, the SDP method brings all components together and solves for an optimal reduction policy. To enable a computationally-tractable solution, statistical methods from data mining and computer experiments (see reviews by Hastie et al. 2001; Tsui et al. 2006; Chen et al. 2003, 2006) are utilized

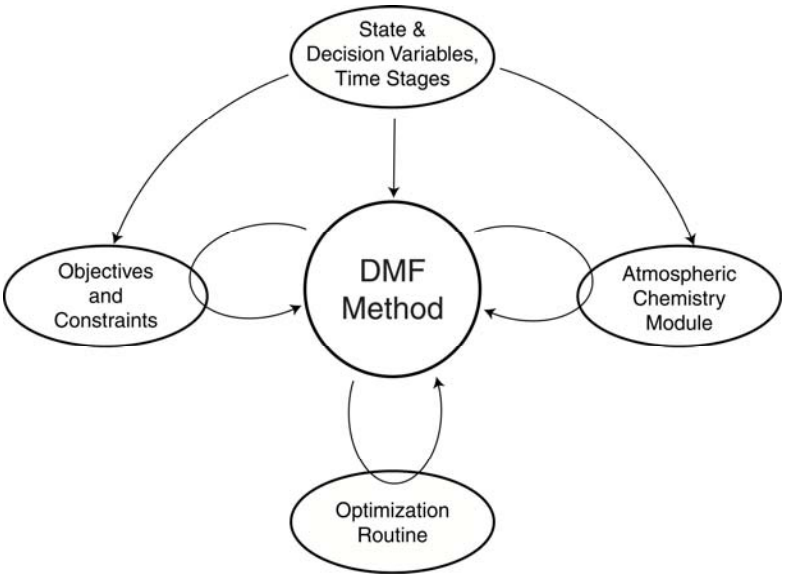

Figure 1: A Modular Decision Making Framework.

within the Atmospheric Chemistry Module and by the SDP solution method. We refer to a "control policy" as a solution from the SDP method, while a "control strategy" is an implementable action.

For the purposes of developing a DMF prototype, we focus on an Atlanta case study for the ozone episode during the time period from July 29 – August 1, 1987. This remains to be one of the worst on record, and to date, no practical control strategy using typical approaches has been identified to handle this episode. Our objective was to see if our DMF could identify a control policy that would achieve the EPA standard more cost-effectively than the across-the-board approach. The details of the Atlanta Atmospheric Chemistry Module were presented by Yang et al. (2007). Hence, the primary contributions of the current paper are to develop the complete SDP model and to demonstrate our DMF prototype using the Atlanta case study. This is the first demonstration of numerically solving a continuous-state SDP problem with hundreds of potential state variables using a mining, metamodeling, and computer experiments based SDP solution method. The success of our DMF prototype shows the potential for identifying dynamic and targeted control strategies for air quality; such a DMF for air quality has never been imagined before, much less realized. More generally, our SDP approach could be applied to other highly complex stochastic dynamic systems, such as dynamic treatment strategies (Murphy 2003).

For ozone pollution control, our objective is to minimize the expected total cost of reducing emissions in order to maintain a maximum ozone level below the EPA standard. In the prototype DMF for Atlanta, we assume the same convex, monotonically increasing cost function for reduction of emissions at each source; however, real economic costs could be incorporated, if available. Although the practicality of the control policy is not considered, a solution from our DMF would provide insights and directions for new control strategies. A brief description of the Atmospheric Chemistry Module is given in the next section, our SDP approach is given in Section 3, and the Atlanta application is in Section 4.

2. Atmospheric Chemistry Module

Yang et al. (2007) developed the Atmospheric Chemistry Module, in which statistical methods from data mining and computer experiments are employed to model the relationships between emissions and ozone

represented by a comprehensive air chemistry photochemical model. We provide a brief description here to facilitate understanding of our DMF, and results for the Atlanta case study are summarized in Section 4.1. Unlike many other applications, it is not practical to change the levels of emissions in a metropolitan airshed and then observe the resulting air quality. Thus, collection of real experimental data is not an option. Advanced air chemistry photochemical models, such as the Urban Airshed Model, are reasonably accurate in simulating air quality (SIP 2001). Consequently, a state implementation plan (SIP) chooses a representative ozone episode, for which an advanced air chemistry photochemical model is painstakingly developed and then used by the state to demonstrate control strategies that lead to compliance with the EPA standard. The weather conditions in the air chemistry photochemical model mimic the weather conditions of the specified ozone episode.

Due to the computational requirements, direct optimization of ozone control strategies using a comprehensive air chemistry photochemical model has not been attempted to date. Methods from design and analysis of computer experiments enable construction of an approximation of the desired performance measure(s) from a computer experiment, such as a simulation model, via "metamodels."

The input data to a comprehensive air chemistry photochemical model include geographical domain boundary parameters, meteorological parameters, and emissions data. In constructing the model input for a specific metropolitan region, meteorological conditions, including wind direction, wind speed, temperature and sunlight, and emissions from various sources must be estimated. Emission sources are categorized as point sources (e.g., power plants, chemical plants, steel mills), mobile sources (e.g., on-road vehicles, ships, airplanes), and area sources (e.g., dry cleaners, paint shops, lawnmowers). In this paper, the mobile and area sources are combined into a single non-point source category. Typically, the model input is developed for a specific ozone episode that is representative for testing control strategies. The emissions that occurred during the modeled episode are called the *base case emissions*. Control strategies (generally) would specify actions to reduce these emissions.

In our DMF's Atmospheric Chemistry Module, we are interested in the relationships between the emissions (point and non-point sources) and the maximum ozone levels given the meteorological

conditions of the specified ozone episode for a modeled metropolitan region. Since we desire to vary emission reductions over time and space, the Atmospheric Chemistry Module's tasks are as follows:

1) Initialization: Identify spatial regions (non-point sources) and time periods in which emission reductions may be independently applied. Identify monitoring stations at which to estimate maximum ozone.
2) Mining: Construct an experimental design over emissions from point and non-point sources to collect "data" from the comprehensive air chemistry photochemical model. Mine this data to eliminate sources at which emission reduction has little effect on maximum ozone.
3) Metamodeling: Construct an experimental design that varies the emissions retained after data mining over the time periods. Use the resulting ozone data from the air chemistry photochemical simulation to fit metamodels for maximum ozone in each time period for each monitoring station. Potential explanatory variables include all emissions from current and previous time periods and maximum ozone in previous time periods.

The purpose of the Initialization Phase is to integrate the expertise of air quality decision-makers. Following this, the Mining and Metamodeling Phases are general and do not require expert input, although, it is always worthwhile to share the results of the mining and metamodeling steps with the domain experts, so as to ensure their reasonability. Given the number of emission sources, the Mining Phase requires construction of an appropriate experimental design, execution of the air chemistry photochemical model for each design point, and usage of an appropriate data mining method. The approach taken by Yang et al. (2007) was to construct a Latin hypercube design and employ a stepwise linear regression procedure with a significance level of 0.10. A potentially more efficient approach was proposed by Shih et al. (2006), which studied mining methods when the number of observations (design points) is *less than* the number of predictor variables (emission sources). Requiring few design points would reduce the number of expensive air chemistry photochemical model runs. Finally, the Metamodeling Phase typically requires a larger experimental design than the Mining task, so as to ensure

accurate modeling by the chosen approximation method. These metamodels will be employed as the SDP state transition equations described in Section 3.3. Yang et al. (2007) chose to employ linear regression metamodels for two reasons: (i) linear state transitions ensure convexity of the SDP optimization; (ii) nonlinearities in the air chemistry can occur when very high concentrations of NOx lead to lower ozone in local areas, and decision-makers know that increasing NOx to reduce ozone locally comes at a cost of increasing ozone concentrations elsewhere. Although only the linear regression metamodels are used in the current paper, Yang et al. (2007) also explored the use of multivariate adaptive regression splines (Friedman 1991, Tsai and Chen 2005).

Define

O_p^S Maximum ozone at monitoring station *S* in time period *p*,

$\boldsymbol{E}_p$ Vector of emissions for all sources in time period *p*,

$\boldsymbol{O}_p$ Vector of ozone levels at all monitoring stations in time period *p*.

An initialization time period $p = 0$ is included to initialize the SDP state variables. Then, a general statistical model for the maximum ozone level at monitoring station *S* in time period *p* can be written as:

$$O_p^S = f_p^S(\boldsymbol{E}_1, \boldsymbol{E}_2, ..., \boldsymbol{E}_p, \boldsymbol{O}_1, \boldsymbol{O}_2, ..., \boldsymbol{O}_{p-1}) + \varepsilon,$$

where $f(\cdot)$ is the function to be approximated and ε is the error term that follows a specified probability distribution. A good approximation does not require all the potential explanatory variables, and *model selection* is a key component of the modeling process. Of all the potential models for the response variables, we are interested in those with low complexity and good prediction accuracy, which are often obtained by minimizing prediction error on a validation data set or conducting cross-validation (Hastie et al. 2001). In the case of linear regression, we can use measures such as the coefficient of determination (R^2) or Akaike's information criterion (Kutner et al. 2004). Only those explanatory variables that appear in the metamodels will need to be included the SDP formulation; thus, critical *dimensionality reduction* occurs within the Atmospheric Chemistry Module. Further discussion on this is given in Section 3.1.

3. Stochastic Dynamic Programming Approach

A continuous-state SDP model over *T* discrete time stages can be written as (Chen et al. 1999):

$$\min_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_T} E\left\{\sum_{t=1}^{T} c_t(\boldsymbol{x}_t,\boldsymbol{u}_t,\boldsymbol{\varepsilon}_t)\right\}$$
$$\text{s.t. } \boldsymbol{x}_{t+1} = f_t(\boldsymbol{x}_t,\boldsymbol{u}_t,\boldsymbol{\varepsilon}_t), \text{ for } t = 1,\ldots,T-1 \text{ and } (\boldsymbol{x}_t,\boldsymbol{u}_t) \in \Gamma_t, \text{ for } t = 1,\ldots,T, \quad (1)$$

where $\boldsymbol{x}_t \in R^n$ is the state vector; $\boldsymbol{u}_t \in R^m$ is the decision vector; $c_t : R^{n+m+l} \to R^1$ is a known cost function for period *t*; the known function $f_t(\cdot)$ defines the state transition from stage t to stage $t+1$; $\Gamma_t \subset R^{n+m}$ is the set of constraints; and the expectation is taken over the random vector $\boldsymbol{\varepsilon}_t \in R^l$. For the ozone pollution problem, *T* time periods defined in the Atmospheric Chemistry Module determine SDP stages 1 through *T*. Details on the state and decision variables, objectives and constraints, and state transition equations are described in Sections 3.1, 3.2 and 3.3, respectively. The SDP solution method is given in Section 3.4.

3.1 State and Decision Variables

In SDP stage *t* (*t* = 1, …, *T*), the set of decision variables corresponds to the emission sources (point and non-point) that will be controlled in time period *t*. It should be noted that we can only control those emission sources that are explicitly represented within the air chemistry photochemical model of the specified ozone episode; hence, emissions beyond the boundaries of the modeling domain cannot be controlled. The set of state variables in SDP stage *t* consists of the complete history of maximum ozone and emissions across the modeled domain, i.e., the maximum ozone levels and emissions occurring in time periods $p < t$. In the Atmospheric Chemistry Module, ozone metamodels were developed to predict maximum ozone in time period *p* given the complete history prior to *p* and the set of emission variables in time period *p*. However, not all the historical variables and emission variables are utilized within these ozone metamodels. Consequently, the decision vector in SDP stage *t* need only maintain those emission variables that are required by the ozone metamodels for time period *t*. Similarly, we can reduce the set of maintained state variables. Clearly, historical variables required by the ozone metamodels for time period

t must be maintained in the state vector for SDP stage *t*. In addition, a historical variable occurring prior to time period *t* that is required by an ozone metamodel in a time period $p > t$ must be maintained in the state vector for SDP stage *t*, regardless of whether it was needed in time period *t*, in order to pass that information on to the later time period. More precisely, we define the following sets of variables and set operations to obtain the *minimal set* of state variables for each SDP stage *t* (stored below in X_t).

V_t = Set of historical variables needed to predict maximum ozone levels in time period *t*.

X_t = Set of variables to be maintained as state variables for SDP stage *t*.

U_t = Set of emission variables to be maintained as decision variables for SDP stage *t*.

Then, the relationships among these variable sets can be derived backwards as follows:

$$\begin{cases} U_T = \{v \mid v \in V_T, v \text{ is an emission variable for time period } T\} \\ X_T = V_T - U_T \end{cases}$$

and

$$\begin{cases} U_t = \{v \mid v \in V_t \cup X_{t+1}, v \text{ is an emission variable for time period } t\} \ , \ t = T-1,\ldots,2,1 \\ X_t = V_t \cup X_{t+1} - U_t, \ t = T-1,\ldots,2,1 \end{cases}.$$

Consequently, unlike the typical SDP formulation in equation (1), the state and decision variable spaces can be considerably different from stage to stage. The dimension of the SDP problem is then taken to be the maximum state dimension among all stages.

3.2 Objectives and Constraints

The primary objective of an air quality control strategy is to maintain the air quality such that it satisfies the EPA standard, e.g., the one-hour EPA ozone standard of 0.12 ppm. This could be enforced via a strict constraint, but to enable a smoother SDP value function, we chose a penalty approach. In our DMF prototype the purpose of the cost function is to identify the emission reductions that provide the most reduction in ozone, so that minimal emission reductions are employed to achieve the EPA standard. As long as the cost function is increasing in form and the emission reductions across sources are weighted equally, this purpose is served. It is not presumed that this type of cost function represents reality since in

actuality control strategies will have different costs. Further discussion on costs may be found in Section 4.4.

For our SDP model, our objectives are:

1) If the EPA standard cannot be attained due to factors not controlled by the decision variables, such as meteorological conditions, initial conditions, etc., then find the SDP control policy to minimize the ozone level.
2) If the EPA standard can be attained, then find a SDP control policy that attains the ozone goals with the least expected cost.

To take these objectives into account, the cost function is formulated for each SDP stage t ($t = 1, 2, ..., T$) and can be stated as follows:

$$c_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{\varepsilon}_t) = \alpha \sum_{u_t^i \in \boldsymbol{u}_t} W_t^i c_e(u_t^i) + \beta \sum_S c_{\max}(O_t^S),$$

where O_t^S denotes the maximum ozone level at monitoring station S in time period t, which is predicted by a statistical metamodel developed in the Atmospheric Chemistry Module as a function of $\boldsymbol{x}_t$, $\boldsymbol{u}_t$, and $\boldsymbol{\varepsilon}_t$; u_t^i is a decision variable in $\boldsymbol{u}_t$ and corresponds to the fraction of emissions reduced at the source indexed by i; W_t^i is a scaling factor for source i; $c_e(\cdot)$ is an increasing emission reduction cost function associated with emission reduction; and $c_{\max}(\cdot)$ is an increasing penalty function on maximum ozone. Constants α and β must be chosen such that the penalty cost dominates the total emission reduction cost when maximum ozone exceeds the EPA standard. Both the emission reduction cost function and penalty function follow quintic forms similar to those used in Chen et al. (1999). The choice of smooth convex functions enabled the use of a fast sequential quadratic programming optimization routine from the NAG library (NAG 1991, 2002).

In our DMF, the emission reduction cost function has the form

$$c_e(u) = \begin{cases} 0 & u \le 0 \\ 4u^3 - 4u^4 & 0 < u < 0.5 \\ u - 0.25 & u \ge 0.5 \end{cases},$$

where u is the fraction of the reduced emission, i.e., $u_t^i = \frac{M_t^i - E_t^i}{M_t^i}$, where M_t^i is the base case (maximum) level of emissions at source i in time period t, and E_t^i is the corresponding reduced emission (hence, $M_t^i - E_t^i$ is the amount of emissions that is removed). When there is no reduction taken on an emission source, then the cost is zero. A constraint of $u_t^i \le 1$ is needed to prevent removal of more emissions than were emitted (and this constraint can be adjusted to further limit emission reductions). Since the emission reduction cost function $c_e(\cdot)$ is based on the fraction of emission reduction from the base case, the unscaled cost of reducing the same fraction for a source with high base case emissions is the same as a source with low base case emissions. In reality, it should cost more to reduce this fraction for the source with high base case emissions because more reduction would be required. Hence, the scaling factor $W_t^i = \frac{M_t^i}{M}$, where M is the total base case emissions summed over the modeling domain, is used to scale the emission reduction cost for source i in time period t based on the proportion of base case emissions generated by this source, where a source with higher base case emissions would have a higher W_t^i.

The penalty function was chosen to have the form

$$c_{\max}(x) = \begin{cases} 0, & x \le 0.118 \\ 2.5 \times 10^{11}(x - 0.118)^3 - 6.25 \times 10^{13}(x - 0.118)^4, & 0.118 < x < 0.12 \\ 10^6(x - 0.119), & x \ge 0.12 \end{cases}.$$

When the maximum ozone level (x) is above 0.12 ppm, the penalty cost will exceed 1000. Thus, the total emission cost should be scaled (using α) such that the maximum total emission reduction cost should not exceed 1000. When the maximum ozone level is less than 0.118 ppm, the penalty cost is zero (0.118 ppm is used instead of 0.125 ppm, the standard to 3 decimal places, to allow for a margin of error, since

photochemical modeling contains numerous sources of uncertainty). Finally, it should be noted that there are modeling uncertainties in the system and the DMF, so in practice, the 0.118 and 0.12 values in the penalty function could be lowered to provide a greater margin for error. However, this would primarily affect the amount of reduction, as opposed to the times and locations of reductions.

3.3 State Transition Equations

The state variables of SDP stage $t + 1$ are determined by state transition equations, which are functions of the state, decision, and random variables of SDP stage t. The state and decision variables have been identified in Section 3.1, such that in each SDP stage t, the necessary variables are available to calculate the predicted maximum ozone levels using the statistical metamodels; thus, these metamodels specify the state transitions for maximum ozone to SDP stage $t + 1$. Using $[O_t^S]_{t+1}$ to denote the maximum ozone occurring in time period t that is saved as a state variable in SDP stage $t + 1$, the state transitions can be written in the form:

$$[O_t^S]_{t+1} = f_t^S(\boldsymbol{x}_t, \boldsymbol{u}_t) + \varepsilon_t^S, \qquad t = 1, \ldots, T \ ,$$

where $\boldsymbol{x}_t$ is the state vector of SDP stage t; $\boldsymbol{u}_t$ is the decision vector of SDP stage t; $f_t^S(\cdot)$ is the metamodel approximation for $[O_t^S]_{t+1}$; and ε_t^S is a random variable that corresponds to the random error in the transition. Note: $[O_T^S]_{T+1}$ is the final maximum ozone level that is controlled by a constraint in SDP stage T. Other state variables in SDP stage $t + 1$, namely emissions and maximum ozone levels occurring in time periods $p < t + 1$, simply take their values from a state variable or emission reduction variable E_t^i of SDP stage t, and the state transitions for these variables are called *identity transitions*.

The stochastic variable ε_t^S is based on the random error in a multiple linear regression model, and we assume these follow a normal distribution with mean zero and variance σ_t^S. The error variances σ_t^S are estimated using the mean square for error (MSE) from each regression model. Residual analyses of the regression models were conducted to validate that normality was reasonably satisfied. However,

for computational reasons, the continuous normal distribution was not incorporated directly in the DMF. Instead, the simple two-point discretization for the standard normal random variable (Z) from Chen et al. (1999) was employed, as shown below.

Z	$P[Z = z]$
−1	0.5
+1	0.5

Furthermore, it is assumed that all the stochastic components in the system are independent. These discretized distributions for ε_t^S are used to approximate the expected value in equation (1). Finer discretizations would yield better approximations of the expected value in the same manner as numerical integration, and future research seeks to represent these uncertainties more realistically using the variability in actual data.

3.4 Solving the SDP model

Chen et al. (1999) introduced the OA/MARS SDP method, which employed orthogonal array (OA, see Chen 2001) experimental designs to discretize a high-dimensional continuous state space and multivariate adaptive regression splines (MARS, Friedman 1991) to construct continuous approximations of the SDP future value functions. A variant of OA/MARS using an automatic stopping rule for MARS (Tsai and Chen 2005) was applied to a 20-dimensional wastewater treatment problem (Tsai et al. 2004), and a version using neural networks in place of MARS was applied to a 30-dimensional water reservoir problem (Cervellera et al. 2006).

Our ozone pollution DMF employed another version similar to the OA/MARS approach, using an experimental design based on a low-discrepancy sequence (a.k.a., number-theoretic method). Much literature has been devoted to the construction of low-discrepancy sequences (see Chen et al. 2003 for background in the context of computer experiments), and application to SDP has only been studied recently (Wen 2005). Well-known constructions include good lattice points, good points, Halton sequences, Hammersley sequences, Sobol´ sequences, Faure sequences, and Niederreiter-Xing sequences (e.g., Halton 1960, Hammersley 1960, Sobol´ 1967, Xing and Niederreiter 1995, Faure 2001). The

advantage of low-discrepancy sequences over other experimental designs is that the discretization points are distributed in the space evenly according to a certain definition of discrepancy (Niederreiter 1992).

4. Prototype Decision-Making Framework for Atlanta

For the Atlanta application, the DMF was used to study the ozone episode during July 29 – August 1, 1987, which is one of the worst on record for Atlanta. To date, no practical control strategy using typical approaches has been identified to handle this episode, and our ultimate objective is use our DMF to see if a dynamic and targeted control strategy could be more cost-effective. A comprehensive air chemistry photochemical model was constructed using the EPA's Urban Airshed Model, which is recommended by EPA for State Implementation Plan (SIP) demonstration, and has been used extensively to study ozone pollution in urban Atlanta (Chameides et al. 1992; Chameides and Cowling 1995; Saylor et al. 1999; Chang and Cardelino 2000; SIP 1998, 2001). The base Urban Airshed Model was developed by Systems Applications International (SAI 1989) for Atlanta.

The modeling domain for Atlanta encompasses a 160×160 kilometer square region containing the metropolitan area. Our Atlanta DMF prototype only considers those emissions represented in the Atlanta Urban Airshed Model for the July 29 – August 1, 1987 ozone episode; emissions transported from outside the modeling domain are not studied. Further discussion on this is given in the Appendix. This Urban Airshed Model computes hourly-averaged ozone levels on a 40×40 grid over the modeling domain. Emissions in Atlanta are represented by 102 *point sources* (e.g., power plant smoke stacks) and cell regions in the 40×40 grid. Base case emissions (without any controls) were estimated by the Georgia Environmental Protection Division (SIP 1998). Meteorological conditions, such as temperatures, wind directions and wind speeds recorded for the 1987 episode, were duplicated in the model input. The Atlanta Urban Airshed Model was run on a 360 MHz Sun Ultra workstation. The first two days (July 29 – 30) are used to initialize the system, and the ozone episode begins on the third day. We studied only the third day to demonstrate our DMF approach. The Atlanta Atmospheric Chemistry Module is briefly described in Section 4.1, and the SDP control policy for July 31 is discussed in Sections 4.2 and 4.3.

4.1 Atlanta Atmospheric Chemistry Module

To reduce ozone concentrations, emissions of NOx and VOCs are candidates for reduction. Atlanta, in particular, is "NOx-limited," which means that targeting VOCs is not effective (Chameides et al. 1988). For our Atlanta DMF, we chose to control only 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 in the SDP model would still be exceedingly large. Thus, in order to achieve a computationally-tractable DMF, a major step in the Atmospheric Chemistry Module is dimension reduction, conducted by the 3 phases described in Section 2. We summarize here the results of Yang et al. (2007) to facilitate understanding of the Atlanta case study. In the Initialization Phase, they devised the following setup:

a) Aggregate the Atlanta 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. The time periods include only daylight hours, since sunlight is required for ozone formation.

c) Control maximum ozone concentrations only at the four Atlanta 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.

In the Mining Phase, a *small* 149-point Latin hypercube experimental design (constructed such that all variance inflation factors were less than 7.0) determined NOx emissions from zero up to the base case level, and stepwise regression with a significance level of 0.10 was used to eliminate unimportant variables. Different time periods were not considered. Of the original 102 point sources, only 15 were found statistically significant, while others were clearly unimportant. Nine out of the 25 grid squares were *not* statistically significant; however, all were retained for analyses in the next phase.

In the Metamodeling Phase, the NOx emissions for the 25 grid squares and 15 significant point sources were varied separately over the 5 time periods using a 500-point Latin hypercube experimental design (with all variance inflation factors less than 1.6). The maximum hourly-averaged 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 state transition equation 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.

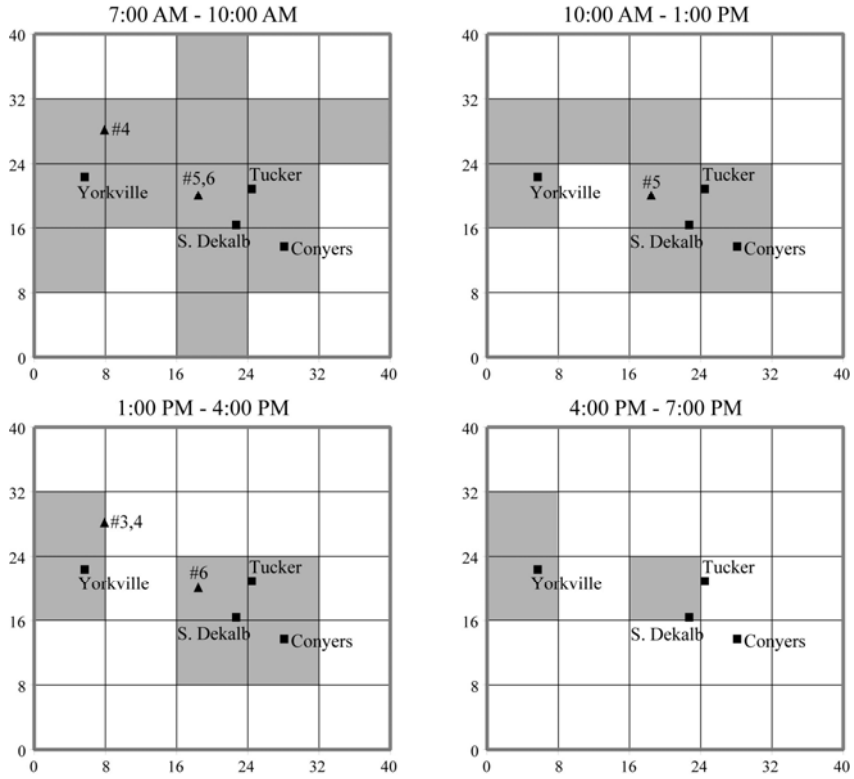

Figure 2: Minimum set of decision variables by time period on the aggregated 5×5 grid. Maintained point emission sources are marked by triangles and maintained non-point emission sources are indicated by shaded grid squares. Maximum ozone is observed at the Atlanta PAMS sites (Conyers, S. Dekalb, Tucker, Yorkville), which are marked by squares.

In each SDP stage we only need to maintain those state and decision variables required by the state transition equation metamodels for that stage and later stages. This greatly reduces the dimension of the SDP system. For Atlanta, the minimum set of decision variables is shown in Figure 2. The minimal

numbers of required state variables for SDP stages 1 through 4, shown in Figure 3, were 17, 25, 23, and 19, respectively. Hence, the effective dimensionality of this SDP problem is 25.

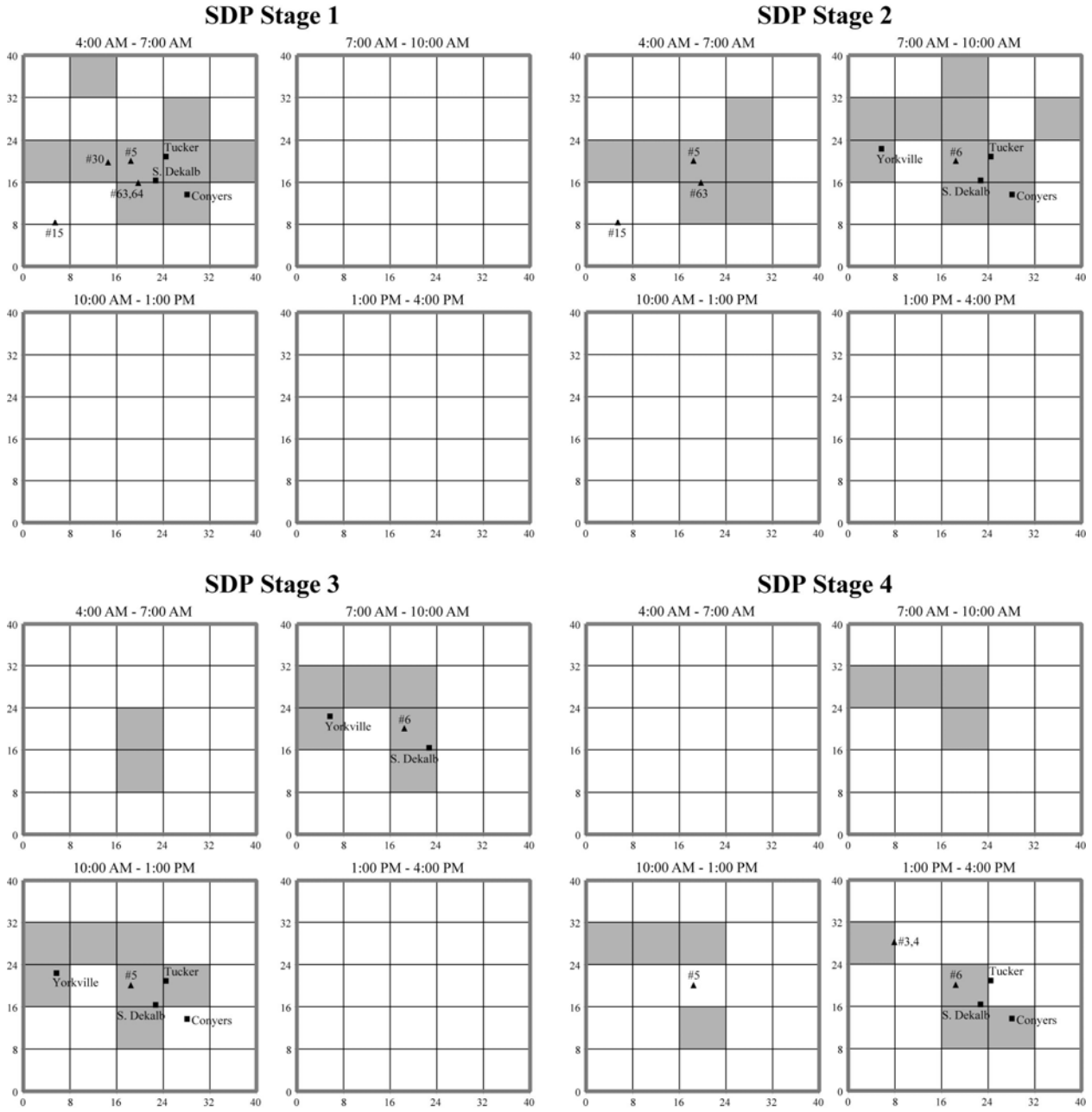

Figure 3: Minimum set of state variables for SDP stages 1–4 on the aggregated 5×5 grid. Maximum ozone state variables are marked by squares, point source emission state variables are marked by triangles, and non-point source emission state variables are indicated by shaded grid squares.

4.2 SDP Control Policies for Atlanta

In the Atlanta SDP model, time periods 1 through 4 were modeled as SDP stages, and time period 0 was considered only as an initialization time period. The emissions from time periods 1 through 4 corresponded to the decision variables in the four SDP stages. As stated in Section 3.4, our SDP solution method employed an experimental design based on a low-discrepancy sequence to discretize the continuous SDP state spaces and continuous MARS approximations of the continuous SDP future value functions. For state space discretization, point sets of 2000 points from Sobol´ sequences were used. These sequences were generalized by Niederreiter (1988) and Xing and Niederreiter (1995). A public-domain Sobol´ sequence generator obtained from http://ldsequences.sourceforge.net/ was used. Table 1 shows the resulting numbers of basis functions for the MARS future value function approximations in each stage and the corresponding run times.

Table 1: MARS results for approximating the SDP future value functions and corresponding run times.

Stage	Number of basis functions selected by MARS	Running time (hh:mm:ss)
1	1754	41:08:28
2	1089	13:44:33
3	120	0:18:55
4	290	0:31:03

Although the optimal solutions that were obtained by the backward SDP solution algorithm could be used to interpolate a control policy, a forward "re-optimization" technique (Tejada-Guibert et al. 1993) has been shown to provide better control policies. For a given initial state vector entering the first SDP stage, $\boldsymbol{x}_1 \in R^n$, the state transition equation metamodels from the Atmospheric Chemistry Module are used to evolve the system, and the optimal decisions are solved forward using equation (1). The "re-optimized" SDP decision vector for initial state $\boldsymbol{x}_1$ can be denoted as a T-tuple, ($\boldsymbol{u}_1$, …, $\boldsymbol{u}_T$), which specifies where, when, and by how much emissions should be reduced. The re-optimized SDP decision vector is then simulated in the Urban Airshed Model to obtain the resulting ozone levels.

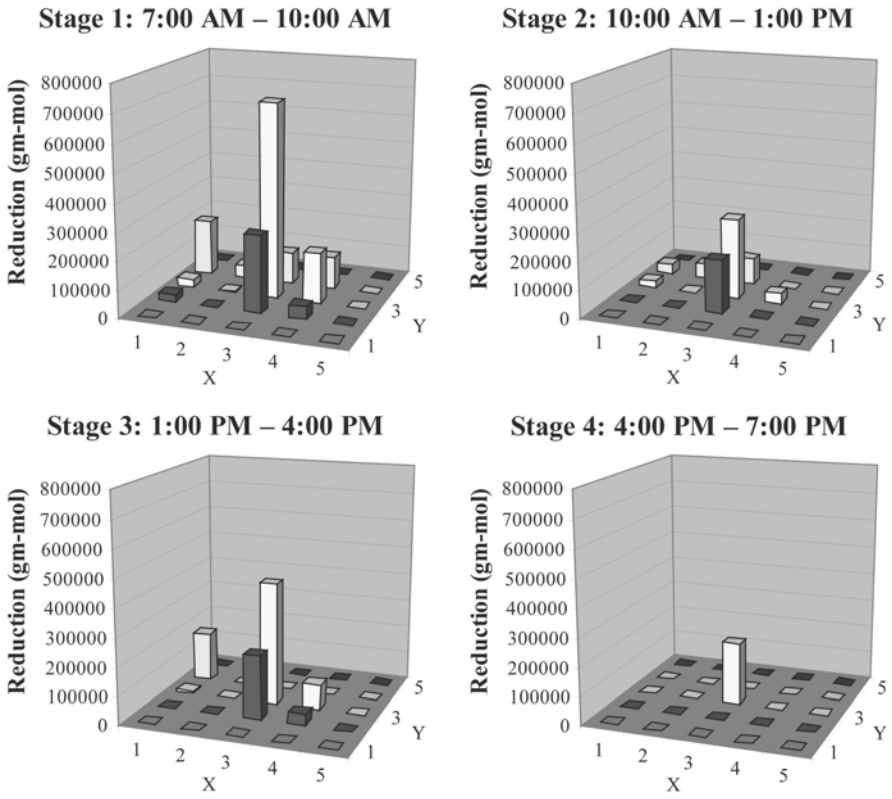

Figure 4: SDP Control Policy for July 31. Reductions in each grid square are combined for point and non-point sources. In Stage 1, point source 4 contributes 77.20% of the reduction shown in grid square (1,4), and point sources 5 and 6 contribute 6.23% and 7.05% of the reduction shown in grid square (3,3). In Stage 2, point source 5 contributes 10.94% of the reduction shown in grid square (3,3). In Stage 3, point sources 3 and 4 contribute 53.39% and 31.53% of the reduction shown in grid square (1,4), and point source 6 contributes 12.41% of the reduction shown in grid square (3,3).

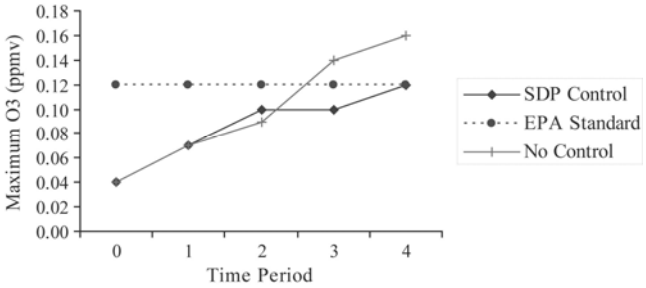

Figure 5: The trajectory of maximum O_3 levels on July 31: No control vs. SDP control policy.

Figures 4 and 5 illustrate the SDP results for the base case emissions on July 31. As can been seen in Figure 4, the SDP decisions are quite different for the different time periods, and the emission

reductions are clearly targeted in certain regions. More importantly, Figure 5 shows that the SDP control policy enabled attainment of the EPA standard in the Atlanta Urban Airshed Model. Since the chemical reactions that lead to ozone are driven by sunlight and heat, precursors (NOx and VOCs) by mid-day are more likely to be involved in these reactions than emissions occurring late in the day. Even though air quality decision-makers are aware of this knowledge, it is not used in their current development of control strategies. A progressive decision-maker would seek control strategies that achieve these optimal emission reductions; however, implementable actions that achieve these reductions may not exist. For example, an action that seeks to reduce mobile emissions in a specific square region in the morning might limit traffic during the morning rush hour in that region. This action would be implementable only if the desired traffic patterns can be enforced. In addition, although this action would reduce the targeted emissions, it might simultaneously increase emissions during other time periods or locations. If these transferred emissions occur during times or locations that have lower impact on maximum ozone, then we would still see an overall reduction in maximum ozone. In seeking implementable actions, additional testing may be conducted via the DMF by adding physical/societal constraints to the set of constraints $(\boldsymbol{x}_t, \boldsymbol{u}_t) \in \Gamma_t$ in SDP equation (1) or altering the cost function if more detailed cost information is available. For example, a constraint on a power plant's emissions in a specified time period, such as $u \leq 0.20$, sets a maximum reduction limit of 20% for that point source in that time period. Practical issues are discussed further in Section 4.4.

The SDP solution, given by the MARS future value function approximations, is valid for a range of initial conditions. For the purposes of illustrating the (re-optimized) SDP control policy under different conditions, 50 hypothetical scenarios were created by randomly generating initial state vectors. Further results for these scenarios are given in the Appendix. The Atlanta Urban Airshed Model was calibrated for each scenario and then used to calculate the corresponding maximum ozone levels. Variation in the decision variables is grouped into three categories in Table 2. About one-third of the sources exhibited no variation; these were reduced by either 0% or 100% in all scenarios. The remaining

sources exhibited some variation, demonstrating the *dynamic* nature of the SDP solution. Overall, nearly two-thirds of the decision variables are similar across all scenarios, but the remaining optimized decisions depend on the initial conditions. With regard to identifying control strategies, sources that exhibited little to no variation could be leveraged to create static controls (that are still targeted in time and location); however, the ultimate efficiency does require some dynamic controls that depend on the initial conditions of the day. Implementation of dynamic controls is certainly not practical at this time.

Table 2: Categories of decision variables based on variation in the control policies for the 50 scenarios. Emission reduction variable in square (*i, j*) in time period *p* is denoted by $E_p^{(i,j)}$. Emission reduction variable at point source pt(*k*) in time period *p* is denoted by $E_p^{\mathrm{pt}(k)}$. In the category of slight variation," each emission reduction variable has the same value for more then 90% of the scenarios or has very small variance (variance less than 0.02 when reductions converted to percentages of base case emissions).

Stage	No variation	Slight variation	Significant variation
1	$E_1^{(1,4)}$ $E_1^{(4,3)}$ $E_1^{\mathrm{pt}(4)}$ $E_1^{(3,2)}$ $E_1^{\mathrm{pt}(6)}$ $E_1^{(3,1)}$	$E_1^{(1,3)}$ $E_1^{(2,3)}$ $E_1^{(3,3)}$ $E_1^{\mathrm{pt}(5)}$ $E_1^{(3,4)}$ $E_1^{(3,5)}$	$E_1^{(1,2)}$ $E_1^{(4,2)}$ $E_1^{(4,4)}$ $E_1^{(2,4)}$ $E_1^{(5,4)}$
2	$E_2^{(1,3)}$ $E_2^{(2,4)}$ $E_2^{(3,4)}$ $E_2^{\mathrm{pt}(5)}$	$E_2^{(1,4)}$ $E_2^{(3,2)}$	$E_2^{(3,3)}$ $E_2^{(4,2)}$ $E_2^{(4,3)}$
3	$E_3^{(4,2)}$	$E_3^{(3,2)}$ $E_3^{\mathrm{pt}(6)}$	$E_3^{(1,3)}$ $E_3^{(3,3)}$ $E_3^{(4,3)}$ $E_3^{(1,4)}$ $E_3^{\mathrm{pt}(3)}$ $E_3^{\mathrm{pt}(4)}$
4	$E_4^{(1,3)}$ $E_4^{(1,4)}$		$E_4^{(3,3)}$

4.3 Cost-effectiveness of SDP Control Policies

To demonstrate cost-effectiveness, across-the-board approaches with varying percent reductions were simulated in the Atlanta Urban Airshed Model for all 50 scenarios, and the resulting maximum ozone levels are shown in Table 3. To achieve equivalent maximum ozone levels as the SDP control policies, an across-the-board reduction of 50% to 60% is needed. As seen in Table 4 of the Appendix, the SDP control policies require a similar total emission reduction in the first time period (7:00 AM – 10:00 AM) and much lower reductions in the later time periods. Since the across-the-board approach requires maintaining the stated percent reduction over the entire 24 hours, it is clear that our DMF approach yields

more cost-effective control policies. Even within the 7:00 AM – 7:00 PM time period controlled by our DMF, the total percent reduction is lower using the SDP control policy.

Table 3: The maximum hourly-averaged ozone level over the 50 scenarios for different reductions using the across-the-board approach.

Overall Reduction Percentage	Maximum O_3 (ppm)
100%	0.092
90%	0.092
80%	0.100
70%	0.111
60%	0.121
50%	0.131
40%	0.139
30%	0.145
20%	0.150
10%	0.155
0%	0.158

4.4 Beyond a Prototype

The knowledge gained from the results of our DMF suggest that (1) it is much more cost-effective to target reductions by time and location, and (2) it would be worthwhile for decision-makers to start thinking about dynamic controls. This is demonstrative of the power of the described DMF for finding heretofore undiscovered opportunities for effectively and efficiently managing air quality. The results shown, however, should not yet be considered a practical solution for improving air quality in the metropolitan Atlanta area, nor have they been considered as such by any authority. While the cost functions used here are reasonable for this demonstration, they are not based on any real experience or survey of expected costs. Further, they do not reflect the many other potential continuous and discontinuous constraints that decision makers must consider, e.g. considerations of economic and environmental justice, political will, social acceptance, or the ability to enforce controls. It is possible to include these issues in a much more complicated model, and to use real cost data that reflects local anomalies and regional trends, but that is well beyond the scope of this study. It is expected that if this DMF were to be applied to any region for the purpose of identifying control strategies beyond the

prototypical demonstration shown here, a significant fraction of the project would consist of developing reasonable, realistic and comprehensive cost functions and constraints. Once in hand, the approach described here can be used to find an optimum solution.

5. Concluding Remarks

Although tremendous scientific advancements have been made in air quality modeling, cost-effective ozone management is still a challenging issue. In this research, statistical computer experiments, stochastic dynamic programming, and a comprehensive 3-D air chemistry photochemical model, were utilized to develop a computationally-tractable DMF that searches for *dynamic* and *targeted* control policies. The DMF can help decision-makers identify where, when, and how much emissions should be reduced to achieve the EPA ozone standard more cost-effectively than an across-the-board approach. To reduce the computational requirements of our DMF, parallel computing can be utilized in the Atmospheric Chemistry Module and in solving SDP (see Eschenbach et al. 1995). If cost and emission reduction information is available, implementable actions could also be explored via our DMF.

Appendix

The modeling domain for Atlanta is drawn over the state of Georgia in Figure 6. While this encompasses a fairly large region beyond the city of Atlanta, since regulators were not able to meet their attainment objective in Atlanta by controlling sources only in Georgia for this same ozone episode described in our study, the State of Georgia joined the Ozone Transport Assessment Group (OTAG, Environmental Council of the States, http://www.epa.gov/ttn/naaqs/ozone/rto/otag/finalrpt/, last updated on 3/9/2006, accessed on 7/20/2007) and sought emissions controls from upwind sources. The OTAG study was convened by the U.S. EPA in 1995 to quantify the amount of upwind emissions that must be controlled in order for downwind areas in the eastern U.S. to meet heretofore impracticable clean air goals. In 1998, the U.S. EPA issued a ruling called the “NOx SIP Call” plan that required states that were upwind of nonattainment areas to reduce emissions to a threshold budgeted amount as determined by the OTAG study (U.S. EPA 1998). Controls in the upwind areas, however, did not fully mitigate the ozone problem

Figure 6: Location of Urban Airshed Model domain for Atlanta.

in the nonattainment areas, including in Atlanta. In turn, the State of Georgia again was required to consider controls on sources for which it has the authority to regulate -- namely those sources in Georgia. In recognition of the jurisdiction over which state regulators have authority, and their pertinent marginal decision-making needs, it is appropriate that options for control are constrained to only those within the defined system.

Table 4 summarizes the decisions from the re-optimization process for each of the 50 hypothetical scenarios of Section 4.2, and shows the resulting maximum ozone level calculated by the Atlanta Urban Airshed Model. In each time period, the total percent reduction was calculated as:

$$100 \times \frac{\text{sum of emission reduction requirements in time period } p}{\text{total emissions in time period } p}.$$

No emission reductions are taken before 7:00 AM or after 7:00 PM. Of the 50 scenarios, 35 attained the EPA standard of 0.125, with the maximum ozone level ranging from 0.1210 to 0.1296 ppm. As mentioned in Section 3.2, parameters of the penalty cost function can be adjusted to allow a greater margin for error.

Table 4: Maximum hourly-averaged ozone and reductions in emissions for the 50 scenarios as percentages of the total base case emissions within each time period. Maximum and minimum values across the scenarios are shown in boldface, and the average is calculated at the bottom.

Scenario No.	Maximum O_3 (ppm)	Reduction requirement (% of total base case emission)			
		7AM–10AM	10AM–1PM	1PM–4PM	4PM–7PM
1	0.1231	53.68%	29.60%	32.56%	6.69%
2	0.1226	53.63%	29.40%	33.68%	6.74%
3	0.1241	**45.49%**	20.72%	33.40%	**13.63%**
4	0.1214	53.81%	35.79%	26.54%	7.99%
5	0.1294	55.04%	**8.95%**	44.60%	3.39%
6	0.1233	49.49%	32.00%	28.80%	6.67%
7	0.1227	53.87%	28.44%	37.30%	6.56%
8	0.1211	53.98%	36.39%	30.78%	9.57%
9	0.1235	48.56%	34.17%	33.30%	5.97%
10	0.1286	53.36%	16.18%	38.26%	3.50%
11	0.1293	53.61%	16.19%	38.26%	3.59%
12	0.1282	53.78%	16.18%	38.26%	3.39%
13	0.1233	53.69%	31.63%	29.92%	6.45%
14	0.1210	53.84%	33.25%	25.48%	9.71%
15	0.1283	55.09%	16.18%	38.23%	3.39%
16	0.1221	53.87%	26.98%	34.40%	6.36%
17	0.1221	54.98%	20.71%	35.80%	5.16%
18	0.1211	53.67%	33.25%	25.00%	9.44%
19	0.1291	54.10%	16.18%	33.40%	3.44%
20	0.1223	53.84%	28.87%	35.10%	6.47%
21	**0.1296**	53.82%	16.18%	38.26%	3.51%
22	0.1224	53.84%	20.77%	38.25%	3.49%
23	0.1287	53.87%	12.06%	37.90%	8.09%

24	0.1234	48.19%	29.27%	32.84%	6.71%
25	0.1223	52.84%	28.96%	34.37%	6.50%
26	0.1231	53.72%	30.15%	32.26%	4.88%
27	0.1287	52.79%	9.53%	**44.63%**	3.53%
28	0.1231	49.54%	29.14%	32.24%	6.77%
29	0.1229	53.64%	28.70%	34.71%	6.59%
30	0.1210	52.78%	33.25%	25.57%	9.62%
31	0.1292	52.79%	16.18%	38.26%	3.47%
32	0.1238	53.82%	34.27%	29.76%	6.68%
33	0.1219	52.75%	20.73%	38.25%	3.51%
34	0.1211	53.95%	33.25%	24.94%	9.41%
35	0.1233	48.27%	29.58%	36.31%	6.48%
36	0.1233	48.19%	32.64%	32.53%	6.52%
37	0.1296	53.90%	16.18%	35.82%	3.41%
38	0.1285	**55.21%**	16.18%	33.40%	**3.38%**
39	**0.1210**	53.79%	33.25%	25.54%	9.66%
40	0.1226	54.01%	27.83%	34.13%	6.53%
41	0.1286	53.83%	16.18%	38.26%	3.43%
42	0.1211	53.84%	33.25%	24.94%	9.41%
43	0.1219	52.67%	37.95%	24.92%	10.95%
44	0.1226	48.22%	30.45%	33.58%	3.42%
45	0.1263	53.64%	13.46%	43.64%	5.80%
46	0.1225	53.86%	20.71%	36.24%	8.86%
47	0.1229	48.22%	29.32%	36.85%	5.63%
48	0.1231	53.68%	30.05%	31.42%	6.65%
49	0.1211	53.80%	**38.22%**	**24.77%**	9.62%
50	0.1293	55.16%	16.19%	33.46%	13.06%
Average	**0.1195**	**85.10%**	**46.57%**	**42.77%**	**27.92%**

Acknowledgements

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

Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press.

Bouzaher, A., J. B. Braden, and G. V. Johnson (1990). "Dynamic programming approach to a class of nonpoint source pollution control problems." *Management Science*, **36(1)**, pp. 1–15.

Cervellera, C., V. C. P. Chen, and A. Wen (2006). "Optimization of a Large-Scale Water Reservoir Network by Stochastic Dynamic Programming with Efficient State Space Discretization." *European Journal of Operational Research*, **171**, pp 1139–1151.

Chameides, W. L., R. W. Lindsay, J. Richardson, C. S. Kiang (1988). "The Role of Biogenic Hydrocarbons in Urban Photochemical Smog – Atlanta as a Case Study." *Science*, **241(4872)**, pp. 1473–1475.

Chameides, W. L., F. Fehsenfeld, M. O. Rodgers, C. Cardelino, J. Martinez, D. Parrish, W. Lonneman, D. R. Lawson, R. A. Rasmussen, P. Zimmerman, J. Greenberg, P. Middleton, and T. Wang (1992). "Ozone Precursor Relationships in the Ambient Atmosphere." *Journal of Geophysical Research*, **97**, pp. 6037–6055.

Chameides, W.L. and E. B. Cowling (1995). *The State of the Southern Oxidants Study (SOS): Research Accomplishments and Future Plans.* http://www2.ncsu.edu:8010/ncsu/CIL/southern_oxidants/docs/intro.pdf.

Chang, M. E. and C. Cardelino (2000). "Application of the Urban Airshed Model to Forecasting Next-Day Peak Ozone Concentrations in Atlanta, Georgia." *Journal of the Air & Waste Management Associate*, **50**, pp. 2010–2024.

Chen, V. C. P. (2001). "Measuring the Goodness of Orthogonal Array Discretizations for Stochastic Programming and Stochastic Dynamic Programming." *SIAM Journal of Optimization*, **12**, pp. 322–344.

Chen, V. C. P., D. Ruppert, and C. A. Shoemaker (1999). "Applying Experimental Design and Regression Splines to High Dimensional Continuous-State Stochastic Dynamic Programming." *Operations Research*, **47**, pp. 38–53.

Chen, V. C. P., K.-L. Tsui, R. R. Barton, and J. K. Allen (2003). "A review of design and modeling in computer experiments." *Handbook in Statistics: Statistics in Industry*, **22**, pp. 231–261, Khattree, R. and Rao, C. R. (eds.), Elsevier Science, Amsterdam.

Chen, V. C. P., K.-L. Tsui, R. R. Barton, and M. Meckesheimer (2006). "Design, Modeling, and Applications of Computer Experiments." *IIE Transactions*, **38**, pp. 273–291.

Eschenbach, E. A., C. A. Shoemaker, and H. Caffey (1995). "Parallel Processing of Stochastic Dynamic Programming for Continuous State Systems with Linear Interpolation." *ORSA Journal on Computing*, **7**, pp. 386–401.

Faure, H. (2001). "Monte-Carlo and Quasi-Monte-Carlo Methods for Numerical Integration." *Combinatorial & Computational Mathematics (Pohang, 2000)*, pp. 1–12, River Edge, NJ: World Science Publishing.

Foufoula-Georgiou, E. and P. K. Kitanidis (1988). "Gradient Dynamic Programming for Stochastic Optimal Control of Multidimensional Water Resources Systems." *Water Resources Research*, **24**, pp. 1345–1359.

Friedman, J. H. (1991). "Multivariate Adaptive Regression Splines (with discussions)." *Annals of Statistics*, **19**, pp. 1–141.

Halton, J. H. (1960). "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals." *Numerische Mathematik*, **2**, pp. 84–90.

Hammersley, J. M. (1960). "Monte Carlo methods for solving multivariable problems." *Annals of the New York Academy of Sciences*, **86**, pp. 844–874.

Hastie, T., J. H. Friedman, and R. Tibshirani (2001). *Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag.

Johnson, S. A., J. R. Stedinger, C. A. Shoemaker, Y. Li, and J. A. Tejada-Guibert (1993). "Numerical Solution of Continuous-State Dynamic Programming Using Linear and Spline Interpolation." *Operations Research*, **41**, pp. 484–500.

Kyan, C. P. and J. H. Seinfeld (1974). "On Meeting the Provisions of the Clean Air Act." *AIChE*, **20(1)**, pp 118–127.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li (2004). *Applied Linear Statistical Models*, 5th edition, New York: McGraw-Hill Irwin.

Loughlin, D. H., S. Ranjithan, J. W. Baugh, and E. D. Brill (2000). "Application of Genetic Algorithms for the Design of Ozone Control Strategies." *Journal of the Air and Waste Management Association*, **50**, pp.1050–1063.

Mhaisalkar, V. A., J. K. Bassin, R. Paramasivam, and P. Khanna (1993). "Dynamic programming optimization of water-treatment-plant design." *Journal of Environmental Engineering*, **119(6)**, pp. 1158–1175.

Murphy, S. A. (2003). "Optimal dynamic treatment regimes." *Journal of the Royal Statistical Society, Series B*, **65(Part 2)**, pp. 331–366.

NAG (1991). *The NAG Fortran Library Manual (Mark 15)*. NAG Inc.

NAG (2002). *The NAG C Library Manual (Mark 7)*. NAG Inc.

Niederreiter, H. (1988). "Low-Discrepancy and Low-Dispersion Sequences." *Journal of Number Theory*, **30**, pp. 51–70.

Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia: SIAM.

SAI (1989). *Urban Airshed Model Study of Five Cities: Low-Cost Application of the Model to Atlanta and Evaluation of the Effects of Biogenic Emissions on Emission Control Strategies*, SYSAPP-89/117, Systems Applications International, 1989

Saylor, R. D., W. L. Chameides, M. E. Chang (1999). "Demonstrating attainment in Atlanta using urban airshed model simulations: Impact of boundary conditions and alternative forms of the NAAQS." *Atmospheric Environment*, **33(7)**, pp. 1057–1064.

Seinfeld, J. H. (1988). "Ozone air quality models: A critical review." *Journal of Air Pollution Control and Waste Management*, **38(5)**, pp. 616–647

Seinfeld, J. H. and C. P. Kyan (1971). "Determination of Optimal Air Pollution Control Strategies." *Socio-Economic Planning Science*, **5**, pp. 173–190.

Seppelt, R. (2001). "Hierarchical dynamic programming and applications in ecosystem management." *Environmental Modelling and Software*, **16(4)**, pp. 377–386.

Seppelt, R. (2001). "Hierarchical dynamic programming and applications in ecosystem management." *Environmental Modelling and Software*, **16(4)**, pp. 377–386.

Shih, D. T., V. L. Pilla, S. B. Kim, J. M. Rosenberger, and V. C. P. Chen (2006). "Efficient Computer Experiment Based Optimization through Variable Selection." In *Proceedings of the Informs Artificial Intelligence & Data Mining Workshop*, Pittsburgh, PA.

Shoemaker, C. A. (1982). "Optimal Integrated Control of Univoltine Pest Populations with Age Structure." *Operations Research*, **30**, pp. 40 – 61.

Sillman, S., K. Al-Wali, F. J. Marsik, P. Nowatski, P. J. Samson, M. O. Rodgers, L. J. Garland, J. E. Martinez, C. Stoneking, R. E. Imhoff, J-H. Lee, J. B. Weinstein-Lloyd, L. Newman and V. Aneja (1995). "Photochemistry of ozone formation in Atlanta, GA: models and measurements." *Atmospheric Environment*, **29**, pp. 3055–3066.

Sobol´, I. M. (1967). "The distribution of points in a cube and the approximate evaluation of integrals." *USSR Comput. Math. Math. Phys.*, **7**, pp. 86–112.

SIP (1998). *State Implementation Plan for the Atlanta Ozone Non-attainment Area*, Georgia Department of Natural Resources, Environmental Protection Division, Air Protection Branch; April 28, 1998

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

Tejada-Guibert, J. A., S. A. Johnson, J. R. Stedinger (1993) "Comparison of Two Approaches for Implementing Multi-Reservoir Operating Policies Derived Using Dynamic Programming." *Water Resources Research*, **29**, pp. 3969–3980.

Trijonis, J. C. (1974). "Economic Air Pollution Control Model for Los Angeles County in 1975." *Environmental Science and Technology*, **8(9)**, pp. 811–826

Tsai, J. C. C. and V. C. P. Chen (2005). "Flexible and Robust Implementations of Multivariate Adaptive Regression Splines within a Wastewater Treatment Stochastic Dynamic Program." *Quality and Reliability Engineering International*, **21**, pp. 689–699.

Tsai, J. C. C., V. C. P. Chen, J. Chen, and M. B. Beck (2004). "Stochastic Dynamic Programming Formulation for a Wastewater Treatment Decision-Making Framework." *Annals of Operations Research*, Special Issue on Applied Optimization under Uncertainty, **132**, pp. 207–221.

Tsui, K.-L., V. C. P. Chen, W. Jiang, and Y. A. Aslandogan (2006). "Data Mining Methods and Applications." In *Springer Handbook of Engineering Statistics*, pp. 651-669, Pham, H. (ed.), Springer-Verlag, London.

U.S. EPA (1990). *User's Guides for the Urban Airshed Model*, EPA-450/4-90-007A-E.

U.S. EPA (1998). "Finding of Significant Contribution and Rulemaking for Certain States in the Ozone Transport Assessment Group Region for Purposes of Reducing Regional Transport of Ozone," Federal Register: October 27, 1998, Volume 63, Number 207, pp. 57355–57404.

Wen, A. (2005). *Statistics-Based Approaches for Stochastic Optimal Control*. Ph.D. Dissertation, The University of Texas at Arlington.

Xing, C. and H. Niederreiter (1995). "A Construction of Low-Discrepancy Sequences Using Global Function Fields." *Acta Arith.*, **73**, pp. 87–102.

Yakowitz, S. (1982). "Dynamic Programming Applications in Water Resources." *Water Resources Research*, **18**, pp. 673–696.

Yang, Z., V. C. P. Chen, M. E. Chang, T. E. Murphy, and J. C. C. Tsai (2007). "Mining and Modeling for a Metropolitan Atlanta Ozone Pollution Decision-Making Framework." *IIE Transactions Special Issue on Data Mining*, **39**, pp. 607–615.