

**Data Mining for State Space Orthogonalization in Adaptive Dynamic  
Programming**

Bancha Ariyajunya

Ying Chen\*

Victoria C. P. Chen

Seoung Bum Kim

## **Abstract**

Dynamic programming (DP) is a mathematical programming approach for optimizing a system that changes over time and is a common approach for developing intelligent systems. Expert systems that are intelligent must be able to adapt dynamically over time. An optimal DP policy identifies the optimal decision dependent on the current state of the system. Hence, the decisions controlling the system can intelligently adapt to changing system states. Although DP has existed since Bellman introduced it in 1957, exact DP policies are only possible for problems with low dimension or under very limiting restrictions. Fortunately, advances in computational power have given rise to approximate DP (ADP). However, most ADP algorithms are still computationally-intractable for high-dimensional problems. This paper specifically considers continuous-state DP problems in which the state variables are multicollinear. The issue of multicollinearity is currently ignored in the ADP literature, but in the statistics community it is well known that high multicollinearity leads to unstable (high variance) parameter estimates in statistical modeling. While not all real world DP applications involve high multicollinearity, it is not uncommon for real cases to involve observed state variables that are correlated, such as the air quality ozone pollution application studied in this research. Correlation is a common occurrence in observed data, including sources in meteorology, energy, finance, manufacturing, health care, etc.

ADP algorithms for continuous-state DP achieve an approximate solution through discretization of the state space and model approximations. Typical state space discretizations involve full-dimensional grids or random sampling. The former option requires exponential growth in the number of state points as the state space dimension grows, while the latter option is typically inefficient and requires an intractable number of state points. The exception is computationally-tractable ADP methods based on a design and analysis of computer experiments (DACE) approach. However, the DACE approach utilizes ideal experimental designs that are (nearly) orthogonal, and a multicollinear state space will not be appropriately represented by such ideal experimental designs. While one could directly build approximations over the multicollinear state space, the issue of unstable model approximations remains unaddressed. Our approach for handling multicollinearity employs data mining methods for two purposes:

(1) to reduce the dimensionality of a DP problem and (2) to orthogonalize a multicollinear DP state space and enable the use of a computationally-efficient DACE-based ADP approach. Our results demonstrate the risk of ignoring high multicollinearity, quantified by high variance inflation factors representing model instability. Our comparisons using an air quality ozone pollution case study provide guidance on combining feature selection and feature extraction to guarantee orthogonality while achieving over 95% dimension reduction and good model accuracy.

## **Keywords**

Data Mining, Design and Analysis of Computer Experiments, Approximate Dynamic Programming, Ozone Pollution

## **1. Introduction**

Dynamic programming (DP, Bellman 1957) is a common approach for developing intelligent systems that must dynamically adapt over time (e.g., White and Sofge 1992, Si et al. 2004). However, DP problems with high dimension are known to be difficult to solve. Exact solutions are only possible for small DP problems or under very limiting restrictions (linear dynamics, Gaussian random variables and quadratic cost). Approximate DP (ADP) methods now exist (Si et al. 2004), but still encounter issues in computational-tractability and approximation quality. In this paper, we focus on finite-horizon stochastic DP with *multicollinear* continuous-state variables. The ADP literature ignores the issues of multicollinearity. In some real world applications, multicollinearity can be quite high. We demonstrate our methods on an ozone pollution air quality case study from Atlanta, Georgia. The state variables consist of concentrations of ozone and its precursor gases that are monitored over multiple time stages defined from 4:00 AM through 7:00 PM. The decision variables are the reductions in precursor gas emissions at various locations in the defined time stages. The objective is to identify the minimal reduction in emissions in each time stage required to maintain the ozone level at or below the required standard set by the U.S. Environmental Protection Agency (EPA) throughout the day. For this case study,

high multicollinearity exists between the concentrations of ozone that are observed over time and at different locations around Atlanta.

Most continuous-state problems require an ADP algorithm that approximates the continuous optimal value function via sampling or discretization of the state space. Chen et al. (1999) introduced a statistical perspective of this approximation concept, by recognizing that state space discretization is analogous to design of experiments (DOE), and the approximation of the optimal value function is analogous to statistical modeling. This is now called a design and analysis of computer experiments (DACE) based ADP approach, where the computer experiment is the optimization that occurs in each (time) stage of the DP. In ideal DOE, orthogonality is a highly desired property (Chen et al. 2006); therefore, DOE is not appropriate for directly representing the state space when the state variables are multicollinear. Another issue due to the multicollinearity in a state space is that the collinear variables contain similar information, which means there is redundant information. A principal danger of this redundancy is it will result in overfitting of the statistical model. In statistics, even though the multicollinearity issue does not decrease the predictive power of the model within the fitted sample data set, small changes in the data or the model may change the coefficient estimates of the regression model erratically, which indicates the created regression model is not robust when applied to new data (Kutner et al. 2004). Therefore, considering these, this paper studies data mining (DM) approaches to address multicollinearity in a DP state space. DM methods are applied for two reasons: (1) to reduce the state space dimensionality of a DP problem, and (2) to orthogonalize a DP state space, so as to enable the use of ideal DOE.

## **2. Background and contribution**

A finite horizon DP has a finite number of discrete stages,  $t = 1, \dots, T$ . In each stage, there is a cost function  $c_t(\cdot)$ , and the objective of stochastic DP is to minimize expected cost  $E[c_t(\cdot)]$  over the  $T$  discrete stages and subject to certain constraints, where the expected value is taken over a random vector  $\varepsilon_t$  that represents uncertainty in the system. For a given current stage  $t$ , the state variables  $x_t$  specify the state of

the system at the beginning of stage  $t$ , and the decision (or control) variables  $u_t$  specify an action that takes place during stage  $t$ . The decision variables are chosen to minimize current plus future expected costs. The transition of the state variables from the current stage  $x_t$  to the next stage  $x_{t+1}$  is defined by the state transition function  $f_t(\cdot)$ , which depends on the state, decision, and random vectors.

At the heart of a DP solution policy is the optimal value function  $V_t(x_t)$ . For the stochastic case of DP, it is defined for stage  $t$  as the minimum expected cost to operate the system from stage  $t$  forward to the end of the time horizon. Given the state  $x_t$  of the system at any stage  $t$ , the optimal value function can be obtained recursively using equation (1), and the optimal policy obtained from solving (1) will be employed to control the system at stage  $t$ .

$$\begin{aligned}
 V_t(x_t) &= \min_{u_t} E \{c_t(x_t, u_t, \varepsilon_t) + V_{t+1}(x_{t+1})\} \\
 \text{s.t. } \quad x_{t+1} &= f_t(x_t, u_t, \varepsilon_t), \text{ for } t = 1, \dots, T \\
 u_t &\in \Gamma_t, \text{ for } t = 1, \dots, T \\
 \text{where } V_T(x_T) &= \min_{u_T} E \{c_T(x_T, u_T, \varepsilon_T)\}
 \end{aligned} \tag{1}$$

All possible values for  $V_t(x_t)$  can be acquired if there are a finite number of states. However, when the state space is continuous, as it is for ozone pollution, it is impossible to get all possible values. Typically, an approach for approximating the solution in continuous-state ADP is to first form a finite grid of discretization points and then approximate the optimal value function. However, in  $n$ -dimensional problems, a grid of points with  $p$  levels in each dimension requires  $p^n$  points, yielding exponential growth in the number of state points, as the number of state variables increases. This is one form of the ‘‘curse of dimensionality’’ which makes DP computationally intractable. More generally, discretization is a form of sampling the state space. In a reinforcement learning based ADP approach, sampling is inefficient, requiring millions of state points (Busoniu et al. 2010, Geramifard et al. 2013).

With the statistical perspective of Chen et al. (1999), the sampling of the state space is controlled via DOE methods and flexible statistical modeling methods can be employed for value function approximation to create a computationally-tractable ADP algorithm. The DACE-based ADP approach is

presented in Algorithm 1. In each DP stage  $t$ , an experimental design is applied to specify values of the state variables. The computer experiment is executed for these designed state values to yield the optimized objective, which is the response of the experiment. Then a statistical model is used to fit these data in order to approximate the optimal value function.

1. For each stage  $t$ : use DOE to sample  $N$  points from the state space  $\{x_{jt}\}_{t=1}^N$ .
2. In each stage  $t = T-1, \dots, 1$ :
  - (a) For each sampled state point  $x_{jt}, j = 1, \dots, N$ , solve the minimization problem (1), where  $t < T-1$ , the optimal value function  $V_{t+1}(\cdot)$  is estimated by  $\hat{V}_{t+1}(\cdot)$ .
  - (b) Construct the estimated  $\hat{V}_t(\cdot)$  via a statistical model using the data from step 2(a).

Algorithm 1: DACE-based ADP algorithm for optimizing a multi-dimensional,  $T$ -stage, continuous-state problems (Chen et al. 1999)

The focus of this paper is to address the case when the sampled state points yield multicollinearity among the state variables. This occurs naturally in applications in which state variables are correlated. Issues due to sampling are ignored in the reinforcement learning based ADP approach, including the popular Q-learning method, because this approach assumes it is feasible to keep sampling indefinitely until “learning” has been sufficiently achieved. By contrast, the DACE-based ADP approach directly seeks more efficient sampling via DOE in order to reduce computational effort. However, ideal experimental designs are orthogonal or “uniformly space-filling” (Chen et al. 2006) and are inappropriate for representing state spaces with multicollinearity. It is possible to transform a space-filling design to follow specific probability distributions by using the concept of  $F$ -discrepancy (Fang and Wang 1994). Cervellera and Macciò (2016) utilized  $F$ -discrepancy within a DACE-based ADP algorithm for solving an inventory forecasting DP problem. While  $F$ -discrepancy could potentially be utilized to generate sampling in multicollinear spaces, it will not address the presence of unimportant variables, and the multicollinearity will still affect the quality of the statistical model needed to approximate the optimal value function.

The contribution of this paper is two-fold: (1) employ DM feature selection to identify unimportant or redundant features and (2) employ DM feature extraction to directly address the multicollinearity via orthogonalization. In real world DP applications, the state transition function  $f_i(\cdot)$  in equation (1) is also unknown, but may be estimated from data. DM feature selection can be employed when building state transition models to eliminate unnecessary state variables, so as to enable a lower-dimensional DP state space, a smaller experimental design, and reduced computational effort when solving for the optimal value function. DM feature selection for real world DP applications has been previously studied by Yang et al. (2007), Yang et al. (2009), Lin et al. (2014), and Shih et al. (2014). DM feature extraction will provide the orthogonality needed to apply ideal DOE, so as to generate sampling that will enable better statistical models for approximating the optimal value function. This is the first paper to study DM feature extraction for ADP. The next section provides a review of the relevant data mining techniques and describes how they can be used to reduce the dimension of a DP problem and orthogonalize the state space. In Section 4, state transition modeling for the Atlanta ozone pollution case study is conducted with a comparison of different DM scenarios. Included are scenarios that employ DM feature selection without DM feature extraction for comparison purposes. Finally, concluding remarks are given in the last section.

### **3. Data mining techniques**

In extremely high-dimensional DP problems with, say, over 100 state variables, directly conducting DACE-based ADP would require a large experimental design under the assumption that all these state variables are important. However, not all of them are vital in practice, and it is not known in advance which ones should be excluded and which ones should be maintained. Hence, DM feature selection methods are able to provide important dimension reduction to decrease computation. Figure 1 is an example from the Atlanta ozone pollution case study. It is clearly seen that the state variable “ykm3p1” has high correlation with the state variable “ykm3p2”, and the state variable “ykm3p2” is very highly correlated with “ykm3p3”. These are ozone concentration variables observed at different time points. In Figure 2, these four state variables have been transformed from the original multicollinear state space to

an orthogonalized space, eliminating the multicollinearity. An ideal experimental design is not appropriate for Figure 1, but is appropriate for Figure 2.

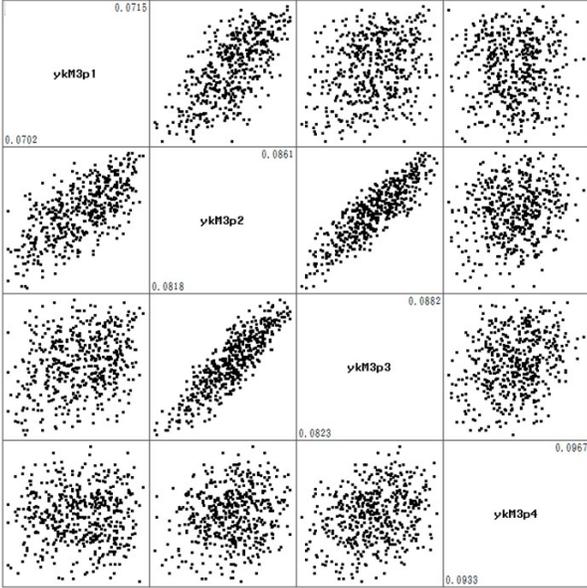


Figure 1. The relationship between state variables in the original multicollinear state space

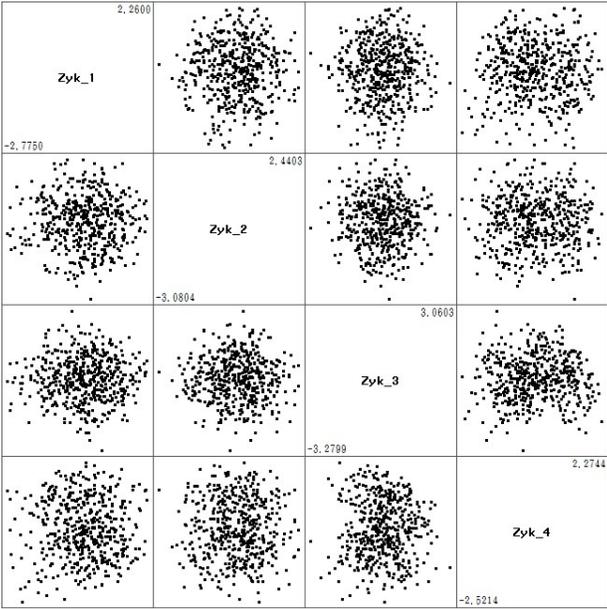


Figure 2. The relationship of state variables in orthogonalized space

To eliminate multicollinearity and reduce dimensionality, the state space can be orthogonalized by using DM tools before implementing a DACE-based ADP algorithm. The DM techniques used in this study can be divided into two types, namely feature selection for dimension reduction and feature extraction for orthogonalization. Together, these tools can generate an efficient and orthogonal DP state space.

### **3.1. Feature selection**

Feature (variable) selection DM techniques are used to reduce the dimensionality of a DP problem by identifying a key subset of features. The feature selection techniques used in this study include stepwise regression, classification and regression trees (Breiman et al. 1984), and a multiple testing procedure based on the false discovery rate (FDR, Benjamini and Hochberg 1995). These techniques were previously studied by Shih et al. (2014), who identified that FDR performed well for the Atlanta ozone pollution case study from Yang et al. (2007) and Yang et al. (2009).

#### **3.1.1. Stepwise regression**

Stepwise regression is an automatic variable selection procedure that uses forward selection and backward elimination processes. In the forward selection process, variables are added one by one to the model if they are statistically significant. Of the variables already included in the model, the backward process deletes insignificant variables. Forward selection and backward elimination processes are repeated until no more variables can be added or deleted. In this study, the significance level threshold for a variable to enter or to stay in the model was specified at 0.05.

#### **3.1.2. Classification and regression trees**

Classification and regression trees (CART) developed by Breiman et al. (1984) have become a very popular data mining tool for supervised learning. We employ only the regression tree version since our variables are continuous. The CART forward algorithm uses binary recursive partitioning to separate the variable space into rectangular regions based on the similarity of the response values. In this research, regression trees are conducted using CART software from Salford Systems ([www.salfordsystems.com](http://www.salfordsystems.com)).

For variable selection, this software provides “variable importance scores.” The variable that receives a 100 score is the most influential variable for prediction, followed by other variables based on their relative importance to the most important variable. However, there are some different options for calculating the scores, and selecting the threshold of the scores to identify important variables may be subjective.

### 3.1.3. Multiple testing procedure based on the false discovery rate (FDR)

Variable selection using FDR typically divides a dataset into  $c$  groups based on a categorical response variable. For each predictor variable ( $x_i$ ), we test the differences in the  $c$  samples, using a  $t$ -test or  $F$ -test. For an  $n$ -dimensional problem, a collection of hypothesis tests and the corresponding  $p$ -values  $\{p_i\}_{i=1}^n$ , where  $p_i$  is the  $p$ -value for testing the null hypothesis for variable  $x_i$  (where a rejected null hypothesis corresponds to a significant variable). In the literature, it is standard to choose a  $p$ -value threshold ( $\alpha$ ) and declare the variable  $x_i$  is significant if and only if the corresponding  $p$ -value has  $p_i \leq \alpha$ . The FDR is defined as the “expected proportion of false positives among all the hypotheses rejected” (Benjamini and Hochberg 1995). The general FDR-procedure to identify significant variables is shown as follows:

1. Choose a fixed  $\alpha$ , where  $0 \leq \alpha \leq 1$ .
  2. Find  $\hat{t} = \max [i: p_i \leq \frac{i}{m} \cdot \frac{\alpha}{\pi_0}]$ , where  $\pi_0 (= \frac{m_0}{m})$  denotes the proportion of true  $H_i$ .
  3. If  $\hat{t} \geq 1$ ,  $\Omega = \{\text{All rejected } H_i \text{ with } p_i < p_{(\hat{t})}\}$  with  $\text{FDR}(\Omega) \leq \alpha$ .
- If  $\hat{t} = 0$ , do not reject any hypothesis since  $\Omega = \emptyset$ .

In this study,  $\alpha = 0.05$  and  $\pi_0 = 1$  are pre-specified.

## 3.2. Feature extraction

Feature extraction DM techniques create new orthogonal features based on transformations of the original features that can supply useful information for modeling (Kim 2009). The new orthogonal features are linear combinations of the original features. Feature extraction can be used for both orthogonalization and dimension reduction. Principal component analysis (PCA) and partial least squares

(PLS) are the feature extraction tools used in this study. Brief descriptions of PCA and PLS are presented in the following subsections.

### 3.2.1. Principal component analysis

PCA can be regarded as a method to compute a new coordinate system formed by principal components (PCs), a.k.a, latent variables or scores, which are orthogonal. Only a small number of the most informative PCs are used. In PCA, correlated original variables ( $X$ ) with  $p$  columns (variables) and  $n$  rows (samples or observations) are transformed to uncorrelated PCs ( $Z$ ) which are linear combinations of  $X$  and are defined in equation (2). Each consecutive PC is orthogonally selected in descending order of the proportion of explained variation in  $X$ .

$$Z = XE \quad (2)$$

where  $E = [E_1, E_2, \dots, E_p]$ ,  $Z = [Z_1, Z_2, \dots, Z_p]$ .

The eigenvectors of the covariance matrix of  $X$  are  $E = [E_1, E_2, \dots, E_p]$ , with corresponding ordered eigenvalues ( $\lambda_1 > \lambda_2 > \dots > \lambda_p$ ), where  $\lambda_i$  indicates the variance of  $Z_i$ . Thus, the first PC ( $Z_1$ ) identifies the direction of highest variation in the original data  $X$ . The second PC ( $Z_2$ ) is orthogonal to the first one, and explains the direction of the next highest variation in the data, and so forth. PCA will produce  $p$  PCs, if the original data  $X$  has  $p$  dimensions. The PCs illustrate the latent structure of  $X$  and can be employed as regressors to predict a response in the regression model.

### 3.2.2. Partial least squares

The model structures of PLS and PCA are very similar. A critical difference between them is that with PLS utilizes the response. The new orthogonal variables (PLS components,  $Z$ ) are chosen to maximize the covariance between  $X$  (predictors) and  $Y$  (responses). PLS can be regarded as a merging of PCA and ordinary least squares. The covariance of  $X$  and  $Y$  merges high variance of  $X$  and high correlation with  $Y$ . The PLS components  $Z$  are achieved by exploring a weight vector  $w$  which maximizes the covariance between the scores of  $X$  and  $Y$  as shown in equation (3), then regresses  $Z$  on  $X$  and  $Y$  through equations (4)-(5), and finally, the prediction model for  $Y$  from the original  $X$  can be obtained by equation (6).  $E$  and  $F$  are residual matrices, and  $P$  and  $Q$  are loading matrices. PLS components  $Z$  can be acquired from many

algorithms, but in this study, Wold's PLS (Wold et al. 2001) is applied, where each PLS component  $Z$  and weight  $w$  are orthogonal ( $Z_i^T Z_j = 0$ ,  $w_i^T w_j = 0$ ;  $i \neq j$ ).

$$Z = Xw \quad (3)$$

$$X = ZP^T + E \quad (4)$$

$$Y = ZQ^T + F \quad (5)$$

$$\hat{Y} = ZQ^T = XwQ^T = X\hat{B} \quad (6)$$

$$\text{where } \hat{B} = wQ^T$$

In general, PLS performs better for prediction than PCA since the new orthogonal predictors  $Z$  are chosen by incorporating information in response  $Y$ . In the following section, combinations of DM scenarios are employed for the Atlanta ozone pollution case study from Yang et al. (2007).

#### 4. Atlanta ozone pollution case study

One of the main reasons for this research is to enable the use of ideal experimental designs for a DACE-based DP solution method when the state variables are highly correlated. The Atlanta ground-level ozone pollution problem from Yang et al. (2007) is selected as our case study because the ozone state variables at different monitoring stations and at different time periods are highly correlated. In addition, the air quality computer model used in the Atlanta ozone problem, called the Atlanta Urban Airshed Model (UAM), is computationally impractical to implement directly in an ADP algorithm. Therefore, more efficient approaches are needed.

Natural ozone that stays in the upper atmosphere is good for our earth. This stratospheric ozone protects the earth from harmful ultraviolet (UV) rays. However, ozone is a harmful pollutant when it is generated in the troposphere because ground-level ozone irritates human respiratory systems and damages vegetation. Ground-level ozone is not emitted directly, but is formed by the chemical reactions of nitrogen oxides ( $\text{NO}_x$ ) and volatile organic compounds (VOCs) in sunlight. Hence, ozone concentrations fall at night, but rise during the day. Therefore, in order to control ground-level ozone, it is required to

control emissions of NO<sub>x</sub> and VOC. However, Atlanta is “NO<sub>x</sub>-limited”, which indicates that targeting VOC emissions is not effective. Thus, in this case study, only NO<sub>x</sub> emissions are controlled. To control NO<sub>x</sub>, we must reduce the sources of NO<sub>x</sub>, which are categorized as point sources and non-point sources. Power plants and other heavy industry are considered point sources of NO<sub>x</sub> emissions, while other sources, such as automobiles and small industry, are treated as non-point sources.

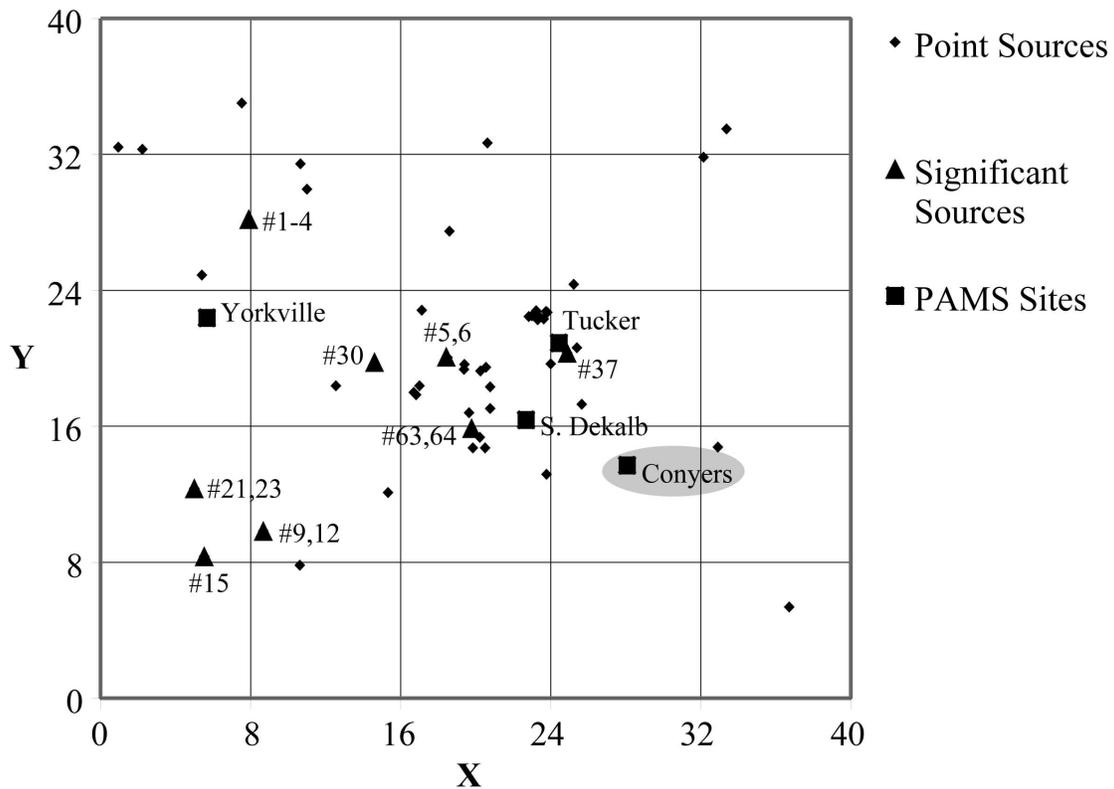


Figure 3. Illustration of the emission sources and the sites for the four Photochemical Assessment Monitoring Stations (PAMS) of the Atlanta ozone pollution case study (Yang et al. 2007)

Figure 3 shows a spatial representation of the Atlanta area in the UAM 40 x 40 grid covering a 160 x 160 kilometer square region of the metropolitan area. Yang et al. (2007) aggregated the 40 x 40 grid into a 5 x 5 grid to spatially represent the non-point source emissions for the Atlanta metropolitan area. A total of 102 point sources are represented as dots or triangles in Figure 3, where the triangles are those that were identified as significant by Yang et al. (2007). Ozone concentrations are monitored by four PAMS

sites, located at Conyers, S. Dekalb, Tucker, and Yorkville. Conyers, in the shaded oval, is commonly the location of the highest ozone concentrations, and our computational study focuses on modeling ozone concentrations for Conyers.

The objective of the Atlanta ozone pollution DP problem is to minimize the reduction of NO<sub>x</sub> emissions needed to prevent ozone concentrations from exceeding the EPA standard, which was 0.12 parts per million in this research (and more recently has been decreased, see <http://www.epa.gov/air/criteria.html>). To reduce ozone concentrations, emission controls are applied to specific point and non-point sources in different time periods covering daytime. Since the chemical reaction that generates ozone occurs in the presence of the sun, only time periods from 4:00 AM to 7:00 PM are considered for reducing emissions. Specifically, five 3-hour time periods are defined: time period 0 is from 4:00 AM to just before 7:00 AM, time period 1 is from 7:00 AM to just before 10:00 AM, time period 2 is from 10:00 AM to just before 1:00 PM, time period 3 is from 1:00 PM to just before 4:00 PM and time period 4 is from 4:00 PM to just before 7:00 PM. Time period 0 is an initialization period, and DP decisions (emission reductions) are controlled in time periods 1 through 4.

Using the DP formulation in equation (1), state and decision variables of the Atlanta ozone pollution case study can be defined as follows. State variables ( $x_t$ ) at the beginning of time period  $t$  include past ozone concentrations at the monitoring sites and past NO<sub>x</sub> emissions at the point sources and non-point sources across the 5 x 5 grid over metropolitan Atlanta area (see Figure 3). Decision variables ( $u_t$ ) are the reductions in emissions at point and non-point sources to be chosen in DP stage  $t$ . The stagewise cost function in time period  $t$  is denoted by  $c_t(\cdot)$  and depends on the state  $x_t$ , the decision  $u_t$ , and a stochastic vector  $\varepsilon_t$  that represents uncertainty in the evolution of the state. Please refer to Yang et al. (2009) for more details.

In Yang et al. (2007), the Atlanta UAM model was used as a computer model for generating data on the relevant air chemistry. A 500-point Latin Hypercube experimental design was used to select a representative sample of point and non-point emission reductions in the different time periods. These 500 emission control settings were executed in the UAM, and the resulting ozone concentrations were

obtained. Given these data, a statistical metamodel is constructed to represent the state transition function  $f(\cdot)$  in equation (1). Figure 4 shows their metamodeling process that uses the emission control settings from experimental design and the ozone concentrations from the UAM to construct metamodel surrogates of the UAM. In this paper, we again utilize these data and this general state transition metamodeling process, focusing on Conyers ozone concentrations. The primary difference is that we now handle the naturally occurring multicollinearity that exists between the ozone concentrations in the different time periods and at the different monitoring sites.

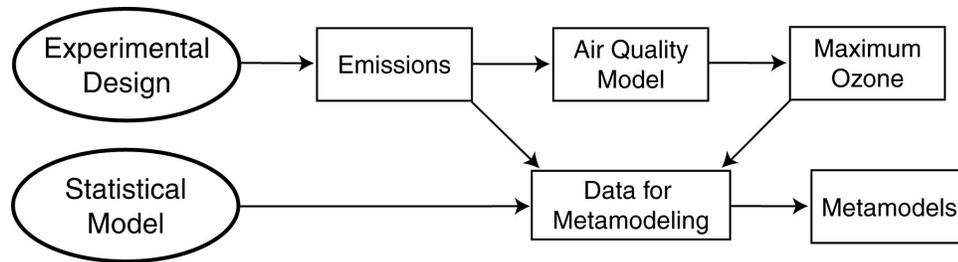


Figure 4. Process of developing metamodel (Yang et al. 2007)

#### 4.1. Data mining scenarios

In our evaluation study, 19 DM scenarios were tested, which are shown in Table 1. Most of these start with a feature selection method to reduce the dimension of the original problem, such as using stepwise regression, FDR, or regression trees. Then orthogonalization is performed using PCA or PLS. For example, scenario A-3 uses stepwise regression on the original dataset to select a subset of state variables, then utilizes PCA on the selected subset to orthogonalize them, and finally conducts stepwise regression again with the respect to the original response on the orthogonal predictors (PCs) to select a final subset of PCs that can efficiently represent the state space. However, in Table 1, it is noted that there are two versions using FDR to conduct the pre-feature selection: one is FDR with a two-category (binary) response and the other is FDR with a continuous response. For FDR with a binary response, we first employ the median of the response to split the response values into two groups and then apply the original

FDR procedure from Benjamini and Hochberg (1995). For FDR with a continuous response, we first apply ordinary least squares to obtain the estimates for all variables; second,  $t$ -statistics for the estimates are calculated; third, the  $p$ -values of all the estimates from  $t$ -statistics can be obtained; finally, the original FDR procedure from Benjamini and Hochberg (1995) is utilized to select important variables by taking advantage of a set of  $p$ -values. Finally, note that scenarios A-1, B-1, C-1, and D-1 do not utilize feature extraction for orthogonalization; hence, these scenarios represent prior existing approaches for comparison purposes. Scenario E-1 utilizes feature extraction only on the ozone concentrations (observed at the PAMS sites) since these variables are the primary contributors for high multicollinearity in the dataset.

Table 1. Proposed DM scenarios

Scenario	Pre-Feature Selection	Feature Extraction	Post-Feature Selection
A-1	Stepwise Regression	None	
A-2	Stepwise Regression	PCA	
A-3	Stepwise Regression	PCA	Stepwise Regression
A-4	Stepwise Regression	PLS	
B-1	FDR w / (2)Categorized Response	None	
B-2	FDR w / (2)Categorized Response	PCA	
B-3	FDR w / (2)Categorized Response	PCA	Stepwise Regression
B-4	FDR w / (2)Categorized Response	PLS	
C-1	FDR w / Continuous Response	None	
C-2	FDR w / Continuous Response	PCA	
C-3	FDR w / Continuous Response	PCA	Stepwise Regression
C-4	FDR w / Continuous Response	PLS	
D-1	Regression Tree	None	
D-2	Regression Tree	PCA	
D-3	Regression Tree	PCA	Stepwise Regression
D-4	Regression Tree	PLS	
E-1	-	PAMS sites – PCA	Stepwise Regression
E-2	-	PCA	Stepwise Regression
F	-	PLS	

#### 4.2. Evaluating the metamodel

Each DM scenario in Table 1 was evaluated on the Atlanta ozone pollution data described for Figure 4 (500-Latin Hypercube design points with ozone concentrations from the UAM) to predict ozone concentrations at the Conyers monitoring site (shaded in Figure 3) for time stages 1 through 4. For each

time stage, the number of predictors is specified in Table 2. This set of predictors from Yang et al. (2007) following a mining phase in which they identified only 15 of the original 102 point sources as being significant (identified as triangles in Figure 3). Consequently, the 5 x 5 grid of 25 non-point sources plus the 15 point sources constitute 40 decision variables for emission control. Since the state space maintains past information on ozone concentrations at 4 monitoring sites and NO<sub>x</sub> emissions at 40 sources, each time period generates 44 state variables. For example, the state variables entering time stage 4 consist of ozone concentrations and NO<sub>x</sub> emissions from time periods 0, 1, 2, and 3 to yield 176 state variables.

Table 2. Number of predictors for the Atlanta ozone problem

	State Space ( $x_t$ ) (Past Ozone & NO <sub>x</sub> )	Decision Space ( $u_t$ ) (NO <sub>x</sub> emission)	Total # Predictors
<b>Stage-1</b>	44 (40 + 4)	40 (5 x 5 + 15)	84
<b>Stage-2</b>	88 (44 x 2)	40	128
<b>Stage-3</b>	132 (44 x 3)	40	172
<b>Stage-4</b>	176 (44 x 4)	40	216

Each DM scenario was evaluated using following performance measures:

- (1) Model  $R^2$ , which measures how well the model fits to the data.
- (2) Number of variables left in the model, which represents the ability to reduce dimension.
- (3) Variance Inflation Factor (VIF), which indicates the level of instability in the parameter estimates due to multicollinearity. Note: Feature extraction yields the ideal VIF of 1.0.
- (4) Prediction error percentage (%Error) calculated using 10-fold cross-validation, which measures model prediction accuracy.

Evaluation results are presented in the following sections by time stage.

#### 4.2.1. Results for the Conyers metamodel in stage 1

The results of the 19 DM scenarios for the Conyers metamodel in stage 1 are shown in Table 3. In this table and Tables 4-6, we highlight the best results in boldface for each measure. From Table 3, for all scenarios, the metamodel of Conyers at stage 1 has low  $R^2$ , which indicates a high level of uncertainty

(unexplained variability) in this time stage of the ozone pollution system. It is interesting to note that while scenarios A-1, B-1, C-1, and D-1 do not use feature extraction, the VIFs are close to the ideal value of 1.0. This is achievable in presence of multicollinearity when the highly collinear variable combinations are excluded as a consequence of a DM feature selection method. For dimension reduction, from 84 variables, modeling with PLS or continuous-response FDR requires only one (latent) variable. For model accuracy, it can be seen that stepwise regression yields the smallest prediction error percentages and higher  $R^2$ , regardless of DM feature extraction. It is interesting to note that scenarios E-1 and E-2, while similar in concept, yield different results. Comparing these two scenarios, it appears that applying PCA only on the PAMS ozone concentrations performs better than applying PCA on the all the variables. Overall, the combination of Stepwise-PLS (scenario A-4) achieves the best of both dimension reduction and model accuracy, while also guaranteeing an orthogonal state space.

Table 3. Results of the DM scenarios for the Conyers metamodel in stage 1

Scenario		$R^2$	Vars. left in model	VIF	%Error
A-1	Stepwise	<b>0.2646</b>	7	(1.0006–1.0137)	<b>1.08925</b>
A-2	Stepwise-PCA	<b>0.2646</b>	7	1	<b>1.08925</b>
A-3	Stepwise-PCA-Stepwise	<b>0.2597</b>	4	1	<b>1.09229</b>
<b>A-4</b>	<b>Stepwise-PLS</b>	<b>0.2636</b>	<b>1</b>	1	<b>1.08741</b>
B-1	FDR	0.2219	3	(1.003–1.008)	1.10580
B-2	FDR-PCA	0.2219	3	1	1.10580
B-3	FDR-PCA-Stepwise	0.2208	2	1	1.10352
B-4	FDR-PLS	0.2205	<b>1</b>	1	1.10384
C-1	conFDR	0.1894	<b>1</b>	1	1.13779
C-2	conFDR-PCA	0.1894	<b>1</b>	1	1.13779
C-3	conFDR-PCA-Stepwise	0.1894	<b>1</b>	1	1.13779
C-4	conFDR-PLS	0.1894	<b>1</b>	1	1.13779
D-1	Tree	0.1937	2	1.0006	1.13840
D-2	Tree-PCA	0.1937	2	1	1.13840
D-3	Tree-PCA-Stepwise	0.1937	2	1	1.13840
D-4	Tree-PLS	0.1936	<b>1</b>	1	1.13890
E-1	PAMSsites-PCA-Stepwise	0.2476	7	(1.002–1.014)	<b>1.08925</b>
E-2	PCA-Stepwise	<b>0.2646</b>	20	1	1.15553
F	PLS	<b>0.3048</b>	<b>1</b>	1	1.18912

#### 4.2.2. Results for the Conyers metamodel in stage 2

To construct the Conyers metamodel in stage 2, the variables from stage 1 are also included in the state space, for a total of 128 predictor variables. The results are shown in Table 4. The addition of these variables enables a much higher  $R^2$  that is close to 1.0 for all scenarios, and the prediction error percentage also drops compared to stage 1. For the scenarios without feature extraction (A-1, B-1, C-1, and D-1), their VIFs are fairly close to the ideal value of 1.0, but higher than in stage 1. With more variables, we can now see more differences in the dimension reduction results, where the worst scenarios are E-1, E-2, F, and stepwise regression scenarios A-1, A-2, and A-3. This result indicates that conducting feature selection with FDR or trees prior to feature extraction is helpful for dimension reduction. In particular, PLS alone (scenario F) does not eliminate as many variables as PLS combined with any of the feature selection methods. Overall, as in stage 1, PLS is again very successful for dimension reduction, requiring as few as one or 3 latent variables. FDR is also successful again, but for stage 2, the binary-response FDR (scenarios B-1 to B-4) is better while the continuous-response FDR (scenarios C-1 to C-4) was better for stage 1. For model accuracy, similar to stage 1, stepwise regression scenarios A-1 through A-4 and E-1 have the smallest prediction accuracy percentages and highest  $R^2$  values. Finally, as in stage 1, Stepwise-PLS (scenario A-4) is the best overall choice.

Table 4. Results of the DM scenarios for the Conyers metamodel in stage 2

Scenario		$R^2$	Vars. left in model	VIF	%Error
A-1	Stepwise	<b>0.9937</b>	10	(1.009–1.260)	<b>0.33001</b>
A-2	Stepwise-PCA	<b>0.9937</b>	10	1	<b>0.33001</b>
A-3	Stepwise-PCA-Stepwise	<b>0.9937</b>	10	1	<b>0.33001</b>
A-4	<b>Stepwise-PLS</b>	<b>0.9934</b>	<b>3</b>	1	<b>0.34289</b>
B-1	FDR	0.9894	<b>3</b>	(1.01–1.23)	0.45582
B-2	FDR-PCA	0.9894	<b>3</b>	1	0.45582
B-3	FDR-PCA-Stepwise	0.9894	<b>3</b>	1	0.45582
B-4	FDR-PLS	0.9894	<b>1</b>	1	0.45582
C-1	conFDR	0.9900	6	(1.01302–1.24359)	0.45066
C-2	conFDR-PCA	0.9900	6	1	0.45066

C-3	conFDR-PCA-Stepwise	0.9900	6	1	0.45066
C-4	conFDR-PLS	0.9897	3	1	0.45497
D-1	Tree	0.9894	4	(1.00064–1.23411)	0.45593
D-2	Tree-PCA	0.9894	4	1	0.45593
D-3	Tree-PCA-Stepwise	0.9894	4	1	0.45593
D-4	Tree-PLS	0.9894	3	1	0.45652
E-1	PAMsites-PCA-Stepwise	<b>0.9946</b>	14	(1.02–2.11)	<b>0.34485</b>
E-2	PCA-Stepwise	<b>0.9935</b>	110	1	0.42443
F	PLS	<b>0.9947</b>	10	1	0.43429

### 4.2.3. Results for the Conyers metamodel in stage 3

To construct the Conyers metamodel in stage 3, the variables from stages 1 and 2 are also included in the state space, for a total of 172 predictor variables. The results are shown in Table 5. The  $R^2$  values are still close to 1.0 for all scenarios, but slightly lower than in stage 2, and the prediction error percentage rises compared to stage 2. This indicates slightly higher uncertainty in stage 3 modeling than stage 2. For the scenarios without feature extraction (A-1, B-1, C-1, and D-1), their VIFs are fairly close to the ideal value of 1.0. For dimension reduction, similar to stage 2, the worst scenarios are E-1, E-2, and stepwise regression scenarios A-1, A-2, and A-3. More importantly, PLS and binary-response FDR are again very successful, requiring as few as 1–3 latent variables. For model accuracy, as in stages 1 and 2, stepwise regression scenarios A-1 through A-4 and E-1 have the smallest prediction accuracy percentages and highest  $R^2$  values. Overall, Stepwise-PLS (scenario A-4) is again the best choice.

Table 5. Results of the DM scenarios for the Conyers metamodel in stage 3

Scenario		$R^2$	Vars. left in model	VIF	%Error
A-1	Stepwise	<b>0.9847</b>	21	(1.02–1.10)	<b>0.51086</b>
A-2	Stepwise-PCA	<b>0.9847</b>	21	1	<b>0.51086</b>
A-3	Stepwise-PCA-Stepwise	<b>0.9846</b>	20	1	<b>0.51669</b>
A-4	<b>Stepwise-PLS</b>	<b>0.9846</b>	<b>3</b>	1	<b>0.52170</b>
B-1	FDR	0.9659	3	(1.004–1.011)	0.75258
B-2	FDR-PCA	0.9659	3	1	0.75258
B-3	FDR-PCA-Stepwise	0.9659	3	1	0.75258
B-4	FDR-PLS	0.9652	1	1	0.75727
C-1	conFDR	0.9727	7	(1.00281–1.01647)	0.72409
C-2	conFDR-PCA	0.9727	7	1	0.72409
C-3	conFDR-PCA-Stepwise	0.9727	7	1	0.72409

C-4	conFDR-PLS	0.9727	2	1	0.72316
D-1	Tree	0.9659	4	(1.00605–1.01342)	0.75327
D-2	Tree-PCA	0.9659	4	1	0.75327
D-3	Tree-PCA-Stepwise	0.9659	3	1	0.74759
D-4	Tree-PLS	0.9659	2	1	0.75677
E-1	PAMSSites-PCA-Stepwise	<b>0.9848</b>	21	(1.01–1.32)	<b>0.50636</b>
E-2	PCA-Stepwise	<b>0.9871</b>	135	1	0.71002
F	PLS	<b>0.9879</b>	11	1	0.71337

#### 4.2.4. Results for the Conyers metamodel in stage 4

To construct the Conyers metamodel in stage 4, the variables from stages 1–3 are also included in the state space, for a total of 216 predictor variables. The results are shown in Table 6. The  $R^2$  values are still close to 1.0 for all scenarios and only very slightly lower than stage 3, but the prediction error percentage clearly rises compared to stage 3. For the scenarios without feature extraction (A-1, B-1, C-1, and D-1), there are VIFs that are too high, exceeding an inflation factor of 5. Only scenario C-1 using continuous-response FDR yields acceptable VIFs that are close to 1.0. Hence, for stage 4, feature extraction becomes necessary to address the multicollinearity. For dimension reduction, as in stages 2 and 3, the worst scenarios are E-1, E-2, and stepwise regression scenarios A-1, A-2, and A-3. More importantly, PLS and FDR are again successful for dimension reduction, requiring 4–9 variables. While this is more than the previous stages, it is still over 95% reduction in dimension. For model accuracy, as in the previous stages, stepwise regression scenarios A-1 through A-4 and E-1 have the smallest prediction accuracy percentages and highest  $R^2$  values. Once again, Stepwise-PLS (scenario A-4) is the best overall.

Table 6. Results of the DM scenarios for the Conyers metamodel in stage 4

Scenario		$R^2$	Vars. left in model	VIF	%Error
A-1	Stepwise	<b>0.9841</b>	26	(1.04–44.9)	<b>0.76287</b>
A-2	Stepwise-PCA	<b>0.9841</b>	26	1	<b>0.76287</b>
A-3	Stepwise-PCA-Stepwise	<b>0.9841</b>	25	1	<b>0.76405</b>
A-4	<b>Stepwise-PLS</b>	<b>0.9841</b>	<b>9</b>	1	<b>0.76289</b>
B-1	FDR	0.9628	<b>9</b>	(1.05–56.15)	1.09164
B-2	FDR-PCA	0.9628	<b>9</b>	1	1.09164
B-3	FDR-PCA-Stepwise	0.9627	<b>8</b>	1	1.08940
B-4	FDR-PLS	0.9627	<b>7</b>	1	1.09064
C-1	conFDR	0.9548	<b>9</b>	(1.003–1.185)	1.25593
C-2	conFDR-PCA	0.9548	<b>9</b>	1	1.25593

C-3	conFDR-PCA-Stepwise	0.9548	<b>9</b>	1	1.25593
C-4	conFDR-PLS	0.9548	<b>4</b>	1	1.25641
D-1	Tree	0.9676	12	(1.01–12.10)	1.03436
D-2	Tree-PCA	0.9676	12	1	1.03436
D-3	Tree-PCA-Stepwise	0.9675	11	1	1.03789
D-4	Tree-PLS	0.9676	<b>9</b>	1	1.03437
E-1	PAMSSites-PCA-Stepwise	<b>0.9836</b>	26	(1.03–6.31)	<b>0.78045</b>
E-2	PCA-Stepwise	<b>0.9864</b>	167	1	1.03480
F	PLS	<b>0.9877</b>	<b>7</b>	1	1.09891

## 5. Concluding remarks

Using the ozone pollution case study, we have compared the performance of a variety of DM scenarios that employ feature selection and feature extraction in various combinations. Across the four DP stages, it is clearly seen that scenario A-4 is the one that best achieves significant dimension reduction while maintaining good model accuracy and guaranteeing orthogonality. This scenario utilizes stepwise regression first then conducts PLS. With regard to feature extraction, both PCA and PLS guarantee orthogonality, but PCA was clearly inferior to PLS for dimension reduction, likely because PLS uses response information to select the latent variables, while PCA does not. With regard to feature selection, stepwise regression can achieve good model accuracy, but not necessarily good dimension reduction. On the other hand, FDR can achieve good dimension reduction, but is inferior with regard to model accuracy. Tree modeling falls in between, with both dimension reduction and model accuracy that is moderate in performance, neither worst nor best.

The primary impact gained from our results is with regard to the multicollinearity. In stages 1 through 3, acceptable VIFs are achieved without feature extraction, however, in stage 4, unacceptably high VIFs is exhibited. While model accuracy is not affected, it must be understood that this result assumes future data will lie within the same multicollinear space. High VIFs directly measure the instability (high variance) in a model's parameter estimates, which means that small changes in the data or even in the modeling approach can lead to very different parameter estimates. In statistics, this instability indicates a non-robust model that may not generalize well (Kutner et al. 2004). These high VIFs also have

implications for DACE-based ADP. When the VIFs are close to 1.0, that indicates the multicollinearity among the selected set of variables is low and an ideal experimental design may still be employed for DACE-based ADP. However, for stage 4, low VIFs could not be achieved for all models and feature extraction becomes necessary in order to appropriately implement DACE-based ADP.

In summary, it is recommended that stepwise regression followed by PLS be employed when high multicollinearity is present among the state variables. However, a disadvantage created by feature extraction is the creation of the latent variables in place of the original variables. The DP formulation in equation (1) and corresponding solution algorithm will need to be restructured in terms of the latent variables ( $z$ ), and all state transition functions will need to be recreated. Future work will formulate candidate processes for incorporating the latent variables in DACE-based ADP. Computational effort will be evaluated, where decreased computation is expected due to dimension reduction, but some increase in computation may result from handling latent variables in the ADP algorithm. We plan to test orthogonalized DACE-based ADP on additional applications, including one related to electric vehicle charging stations. Finally, the ultimate comparison will demonstrate the impact of high multicollinearity on the performance of the ADP solution policy, so as to promote the use of the DM methods presented in this paper.

## **Acknowledgements**

This research was partially supported by the National Science Foundation (ECCS-0801802) and the Dallas-Fort Worth International Airport.

## References

- Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press, Princeton, NJ.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of The Royal Statistical Society Series B-Methodological*, 57(1), pp. 289-300.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth, Inc., Monterey, CA.
- Busoniu, L., Babuska, R., De Schutter, B., and Ernst, D. (2010). *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press, Boca Raton, FL.
- 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(1), pp. 38–53.
- Chen, V. C. P., Tsui, K.-L., Barton, R. R., and Meckesheimer, M. (2006). A Review of Design and Modeling in Computer Experiments. *IIE Transactions*, 38(4), pp. 273–291.
- Cervellera, C. and Macciò, D (2016). F-Discrepancy for Efficient Sampling in Approximate Dynamic Programming. *IEEE Transactions on Cybernetics*, 46(7), pp. 1628-1639.
- Fang, K. T. and Wang, Y. (1994). *Number-Theoretic Methods in Statistics*. Chapman & Hall, New York, NY.
- Geramifard, A., Walsh, T., Tellex, S., Chowdhary, G., Roy, N., How, J. (2013). A Tutorial on Linear Function Approximators for Dynamic Programming and Reinforcement Learning, *Foundation and Trends in Machine Learning*, 6(4), pp. 375–454.
- Kim, S. B. (2009). Feature Extraction/Selection in High-Dimensional Spectral Data. In *Encyclopedia of Data Warehousing and Mining*, Wang, J. (eds.), Information Science Reference, Hershey PA, pp. 863-869.

- Kutner, M., Nachtsheim, C., Neter, J., and Li, W. (2004). *Applied Linear Regression Models*-5th Edition. McGraw-Hill, Homewood, IL.
- Lin, C. F., LeBoulluec, A. K., Chen, V. C. P., Gatchel, R. J. (2014). An Adaptive Pain Management Framework. *Health Care Management Science*, 17(3), pp. 270-283.
- Shih, D. T., Kim, S. B., Chen, V. C. P., Rosenberger, J. M., Pilla, V. L. (2014). Efficient Computer Experiment-Based Optimization through Variable Selection. *Annals of Operations Research*, Special Issue on Data Mining & Informatics, 216, pp. 287–305.
- Si, J., Barto, A. G., Powell, W. B., and Wunsch, D. (2004). *Handbook of Learning and Approximate Dynamic Programming*. Wiley, New York, NY.
- White, D. A., Sofge, D. A. (1992). *Handbook of Intelligent Control*. Van Nostrand, New York, NY.
- Yang, Z., Chen, V. C. P., Chang, M. E., Murphy, T. E., Tsai, J. C. C. (2007). Mining and Modeling for a Metropolitan Atlanta Ozone Pollution Decision-Making Framework. *IIE Transactions*, Special Issue on Data Mining, 39, pp. 607–615.
- Yang, Z., Chen, V. C. P., Chang, M. E., Sattler, M. L., Wen, A. (2009). A Decision-Making Framework for Ozone Pollution Control. *Operations Research*, 57(2), pp. 484-498.