A Response Surface Model of the Air Quality Impacts of Aviation

by

Tudor Maşek

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2008

May 2008

© Massachusetts Institute of Technology 2008. All rights reserved.

Author .......................................................... Department of Aeronautics and Astronautics

May 23, 2008

Certified by ..............................................................

Ian A. Waitz

Jerome C. Hunsaker Professor of Aeronautics and Astronautics
Department Head
Thesis Supervisor

Certified by ..............................................................

Karen E. Willcox

Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by ..............................................................

Prof. David L. Darmofal
Associate Department Head
Chair, Committee on Graduate Students
A Response Surface Model of
the Air Quality Impacts of Aviation

by

Tudor Mașek

Submitted to the Department of Aeronautics and Astronautics
on May 23, 2008, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

Aviation demand is expected to double in the coming decades, and there are growing concerns about its impacts on the environment. Governments seek to mitigate the impacts of aviation on climate, air quality, and noise by setting various emissions and noise regulations. However, there are complex interactions among these three impact pathways which must be carefully considered. The FAA is developing an integrated suite of software tools to allow policy makers to explore the tradeoffs among these environmental impacts for various regulatory options, and to weigh them against the costs to the aviation industry of those regulations. One component of this tools suite is the Aviation Environmental Portfolio Management Tool (APMT) which is designed to analyze industry economics and environmental impacts. Within APMT, there is a desire for faster models that can analyze multiple policy scenarios for decades into the future in order to inform policy decisions on a reasonable time scale. One particular need is that for a fast surrogate air quality model that relates changes in aviation activity to changes in ambient pollutant concentrations.

In this thesis, a response surface model (RSM) is developed for the high-fidelity, but time-consuming, Community Multiscale Air Quality (CMAQ) simulation system. The RSM relates changes in aviation emissions in the United States to changes in ambient concentrations of particulate matter, the main driver of the air quality impacts on public health. Specifically, the surrogate model takes in yearly inventories of landing-taxi-take-off cycle fuel burn, sulfur oxides, nitrogen oxides, and non-volatile primary particulate matter, and returns the resulting changes in ground level annual average ambient particulate matter concentrations. The RSM design space is set to capture likely emissions scenarios over the next 20 years. A low discrepancy sequence is used to generate the 27 CMAQ sample points in order to allow the flexibility of adding more CMAQ simulations as necessary without disrupting the coverage of the design space. Three formulations are then explored for the particulate matter RSM, two kriging models and one regression model. A leave-k-out cross-validation is performed to select the final RSM formulation and analyze its error behavior with the addition of successive CMAQ training points. Finally, the RSM is compared to a previous surrogate model based on the intake fraction method.

The ordinary least-squares regression model is found to perform better than the two kriging formulations, yielding a root-mean-square prediction error of around 1%. The error decays at a rate of just over 0.01% with the addition of each of the last 5 CMAQ runs. Running the RSM with a baseline emissions inventory then yields an estimate of the air quality and subsequent public health impacts of current aviation emissions. A Monte Carlo
simulation provides uncertainty distributions on the RSM outputs. The net increase in risk of adult premature mortality due to aviation—reported as the number of new incidences across the modeling domain—is estimated at 210, with a 95% confidence interval between 140 and 290. The total cost of the increased risk of adult premature mortality, as well as of several other health endpoints, is estimated at $1.21 billion with a 95% confidence interval between about $370 million and $2.15 billion (year 2000 US dollars). These estimates are roughly half of those given by the prior intake fraction model. Of these total impacts, 30% are found to stem from emissions of volatile organic compounds and volatile particulate matter from organics, another 30% from emissions of sulfur dioxide and volatile particulate matter from sulfur, 28% from nitrogen oxide emissions, and about 11% from non-volatile particulate matter emissions.

Thesis Supervisor: Ian A. Waitz
Title: Jerome C. Hunsaker Professor of Aeronautics and Astronautics
Department Head

Thesis Supervisor: Karen E. Willcox
Title: Associate Professor of Aeronautics and Astronautics
Acknowledgments

I owe many people a great deal of thanks for their support and guidance throughout my research and the for the completion of this thesis.

First I’d like to thank my advisors, without whom this work would not have been possible. Thank you Prof. Karen Willcox for pushing me when I needed to get the ball rolling, as well as your help and advice regarding surrogate modeling. Thank you Prof. Ian Waitz for you constant optimism that motivated me in my work, as well as your guidance on aircraft engines and emissions chemistry.

I am also obliged for the assistance of numerous other people in my work. Thank you Dr. Sarav Arunachalam for your technical help with CMAQ, and Eric Clark for your help with all those CMAQ simulations. Thank you Julien Rojo for sharing your knowledge when I first joined the project and answering all those emails. Thank you Anuja Mahashabde and Chris Sequeira for your help with emissions inventories and health impacts modeling. Thank you Garret Barter for your support with my endless Unix questions, and thanks to the ACDL for hosting me.

Finally, I am grateful for my friends and family who encouraged me throughout my years at MIT. Thank you Tri specifically for your interest in my project and our endless debates and technical discussions. Thank you Mom and Dad for your love and support; I know I don’t call home as often as you’d like, but I am forever indebted to you guys.
# Contents

1 Introduction ........................................... 17
   1.1 Context ........................................ 17
   1.2 Aviation Environmental Portfolio Management Tool ............... 18
   1.3 Motivation ...................................... 19
   1.4 Contributions .................................. 20
   1.5 Thesis Outline .................................. 21

2 Response Surface Model Approach ............................. 23
   2.1 Air Quality Modeling ............................... 24
      2.1.1 CMAQ Platform ................................ 24
      2.1.2 Aviation Emissions Inventories ................. 26
      2.1.3 Post-Processing of PM Concentrations .......... 27
   2.2 Emissions Trade-Space ............................. 29
      2.2.1 Input Variables ................................ 29
      2.2.2 First Order Approximation ..................... 31
      2.2.3 Input Variable Ranges ......................... 33
   2.3 Sampling Technique ................................ 36
      2.3.1 Full Factorial ................................ 36
      2.3.2 Latin Hypercube ................................ 37
      2.3.3 Centroidal Voronoi Tessellation .............. 37
      2.3.4 Low-Discrepancy Sequence .................... 38
   2.4 Model Formulation ................................ 40
      2.4.1 Regression .................................... 40
      2.4.2 Interpolation .................................. 43
2.4.3 Ordinary Kriging Equations

3 RSM Implementation

3.1 Inventory Multipliers vs. Design Space Multipliers

3.2 Grid-Cell Models

3.3 Spatial Interpolation

3.3.1 Parametric Interpolation

3.4 Model Uncertainty

3.4.1 Emissions Inventories

3.4.2 Spatial Interpolation

3.4.3 Regression and Kriging Uncertainty

3.4.4 Health Impacts

3.4.5 Sequential Introduction of Uncertainty

3.5 Assumptions and Limitations

3.5.1 Emissions Processing

3.5.2 Air Quality Modeling

3.5.3 RSM Construction

3.5.4 RSM Uncertainties in the Context of Paired Analyses

4 Results

4.1 RSM Error Analysis

4.1.1 Cross-Validation

4.1.2 Final Model Selection

4.1.3 Error Decay with Additional CMAQ Runs

4.2 Health Impacts of Current Levels of Aviation

4.2.1 Monte Carlo Comparison of RSM to iF

4.2.2 RSM Point Estimate Comparison to CMAQ

4.2.3 Impact Apportionment to Input Emissions

5 Conclusions

5.1 Summary

5.2 Future Work

5.2.1 Current CMAQ Data
5.2.2 Future Air Quality Simulations

A Design of Experiments 83
B Spatial Interpolation 85
C Airport Mapping to Representative Profiles 87
D Grid Beta Plots 89
List of Figures

2-1 Model domain and airport locations ........................................ 27
2-2 Typical sampling techniques .................................................. 38
2-3 2D Halton sequence ............................................................... 39
2-4 Regression vs. interpolation .................................................... 43
2-5 Correlation functions ............................................................. 45
2-6 Variogram functions ............................................................... 45

3-1 Spatial interpolation: 1-dimension ........................................... 51
3-2 Spatial interpolation: α ........................................................... 52
3-3 Spatial interpolation: β ........................................................... 52
3-4 Population across the CMAQ domain ....................................... 56
3-5 Population across the CMAQ domain (log10) ............................. 56

4-1 RSM kriging model errors ...................................................... 67
4-2 RSM regression model errors ................................................... 67
4-3 CMAQ run histograms for error convergence cross-validation .... 70
4-4 RSM error convergence .......................................................... 70
4-5 RSM $R^2$ map ................................................................. 71
4-6 Change in PM concentrations due to aviation ....................... 76
4-7 Change in PM concentrations due to aviation (log$_{10}$) ............ 76
4-8 Increase in PM exposure due to aviation .................................. 77
4-9 Increase in PM exposure due to aviation (log$_{10}$) ................. 77

B-1 Spatial interpolation: fuel burn at airports in domain ............. 85
D-1 β$_1$ values across the domain ............................................... 90
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-2</td>
<td>$\beta_2$ values across the domain</td>
<td>91</td>
</tr>
<tr>
<td>D-3</td>
<td>$\beta_3$ values across the domain</td>
<td>92</td>
</tr>
<tr>
<td>D-4</td>
<td>$\beta_4$ values across the domain</td>
<td>93</td>
</tr>
<tr>
<td>D-5</td>
<td>Changes in PM due to postBPR nvPM emissions</td>
<td>94</td>
</tr>
</tbody>
</table>
List of Tables

1 List of Symbols (A-L) .................................................. 15
2 List of Symbols (M-V) .................................................. 16

2.1 CMAQ component versions. .......................................... 25
2.2 CMAQ grid projection parameters. .................................. 25
2.3 CMAQ output species. .................................................. 27
2.4 DoE variables and ranges .............................................. 35

3.1 Parametric interpolation .............................................. 51
3.2 Sample spatial interpolation ......................................... 52

4.1 Cross-validation combinations ........................................ 70
4.2 RSM to iF comparison inventories .................................. 72
4.3 iF vs RSM, pre-BPR correction, incidences ....................... 73
4.4 iF vs RSM, pre-BPR correction, valuations ....................... 73
4.5 iF vs RSM, post-BPR correction, incidences ...................... 74
4.6 iF vs RSM, post-BPR correction, valuations ...................... 74
4.7 Impact apportionment to input emissions ......................... 78

A.1 DoE multipliers input combinations .................................. 83

C.1 Mapping of 325 airports in domain to 3 representative profiles .... 87
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEDT</td>
<td>Aviation Environmental Design Tool</td>
</tr>
<tr>
<td>APMT</td>
<td>Aviation environmental Portfolio Management Tool</td>
</tr>
<tr>
<td>APU</td>
<td>Auxiliary Power Unit</td>
</tr>
<tr>
<td>BenMAP</td>
<td>environmental Benefits Mapping and Analysis Program</td>
</tr>
<tr>
<td>BPR</td>
<td>ByPass Ratio</td>
</tr>
<tr>
<td>BSU</td>
<td>Boise State University</td>
</tr>
<tr>
<td>BVB</td>
<td>Benefits Valuation Block</td>
</tr>
<tr>
<td>CAEP</td>
<td>Committee for Aviation and Environmental Protection</td>
</tr>
<tr>
<td>CMAQ</td>
<td>Community Multiscale Air Quality modeling system</td>
</tr>
<tr>
<td>CO</td>
<td>carbon monoxide</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>CRF</td>
<td>Concentration Response Function</td>
</tr>
<tr>
<td>CV</td>
<td>Cross-Validation</td>
</tr>
<tr>
<td>CVT</td>
<td>Centroidal Voronoi Tessellation</td>
</tr>
<tr>
<td>DACE</td>
<td>Design and Analysis of Computer Experiments</td>
</tr>
<tr>
<td>$dI_{ij}$</td>
<td>change in incidence of health endpoint $j$ for emissions scenario $i$</td>
</tr>
<tr>
<td>DoE</td>
<td>Design of Experiments</td>
</tr>
<tr>
<td>DRF</td>
<td>Dose Response Function</td>
</tr>
<tr>
<td>EDMS</td>
<td>Emissions and Dispersion Modeling System</td>
</tr>
<tr>
<td>EDS</td>
<td>Environmental Design Space</td>
</tr>
<tr>
<td>EI</td>
<td>Emissions Index</td>
</tr>
<tr>
<td>EPA</td>
<td>Environmental Protection Agency</td>
</tr>
<tr>
<td>EPAct</td>
<td>US Energy Policy Act of 2005 g</td>
</tr>
<tr>
<td>FAA</td>
<td>Federal Aviation Administration</td>
</tr>
<tr>
<td>FESG</td>
<td>Forecasting and Economics Support Group</td>
</tr>
<tr>
<td>FOA</td>
<td>First Order Approximation for aircraft particulate matter emissions</td>
</tr>
<tr>
<td>GRUMP</td>
<td>Global Rural-Urban Mapping Project</td>
</tr>
<tr>
<td>HC</td>
<td>hydrocarbon</td>
</tr>
<tr>
<td>ICAO</td>
<td>International Civil Aviation Organization</td>
</tr>
<tr>
<td>iF</td>
<td>intake Fraction</td>
</tr>
<tr>
<td>IPCC</td>
<td>Intergovernmental Panel on Climate Change</td>
</tr>
<tr>
<td>LAQ</td>
<td>Local Air Quality</td>
</tr>
<tr>
<td>LTO</td>
<td>Landing and TakeOff</td>
</tr>
</tbody>
</table>
Table 2: List of Symbols (M-V)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>MM5</td>
<td>Mesoscale (meteorology) Model version 5</td>
</tr>
<tr>
<td>NAAQS</td>
<td>National Ambient Air Quality Standards</td>
</tr>
<tr>
<td>NEI</td>
<td>National Emissions Inventory</td>
</tr>
<tr>
<td>NH&lt;sub&gt;4&lt;/sub&gt;</td>
<td>ammonia</td>
</tr>
<tr>
<td>NO&lt;sub&gt;x&lt;/sub&gt;</td>
<td>nitrogen oxides</td>
</tr>
<tr>
<td>NO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>nitrogen dioxide</td>
</tr>
<tr>
<td>NO&lt;sub&gt;3&lt;/sub&gt;</td>
<td>nitrate</td>
</tr>
<tr>
<td>nvPM</td>
<td>non-volatile primary Particulate Matter</td>
</tr>
<tr>
<td>OK</td>
<td>Ordinary Kriging</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least-Squares</td>
</tr>
<tr>
<td>PARTNER</td>
<td>Partnership for Air Transportation Noise and Emissions Reduction</td>
</tr>
<tr>
<td>PEB</td>
<td>Partial Equilibrium Block</td>
</tr>
<tr>
<td>PM</td>
<td>Particulate Matter</td>
</tr>
<tr>
<td>PM&lt;sub&gt;10&lt;/sub&gt;</td>
<td>coarse Particulate Matter (aerodynamic diameter ≤ 10µm)</td>
</tr>
<tr>
<td>PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>fine Particulate Matter (aerodynamic diameter ≤ 2.5µm)</td>
</tr>
<tr>
<td>PPM</td>
<td>Primary Particulate Matter</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>coefficient of determination</td>
</tr>
<tr>
<td>RE&lt;sub&gt;i&lt;/sub&gt;</td>
<td>Relative Error in change in Incidence for emissions scenario i</td>
</tr>
<tr>
<td>RMS</td>
<td>Root-Mean-Square</td>
</tr>
<tr>
<td>RSM</td>
<td>Response Surface Model</td>
</tr>
<tr>
<td>SAGE</td>
<td>System for Assessing Global Emissions</td>
</tr>
<tr>
<td>SANDWICH</td>
<td>Sulfate, Adjusted Nitrate, Derived Water, Inferred Carbonaceous</td>
</tr>
<tr>
<td>SEDAC</td>
<td>SocioEconomic Data and Applications Center</td>
</tr>
<tr>
<td>SMAT</td>
<td>Speciated Model Attainment Test</td>
</tr>
<tr>
<td>SMOKE</td>
<td>Sparse Matrix Operator Kernel Emissions</td>
</tr>
<tr>
<td>SN</td>
<td>Smoke Number</td>
</tr>
<tr>
<td>SO&lt;sub&gt;x&lt;/sub&gt;</td>
<td>sulfur oxides</td>
</tr>
<tr>
<td>SO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>sulfur dioxide</td>
</tr>
<tr>
<td>SO&lt;sub&gt;4&lt;/sub&gt;</td>
<td>sulfate</td>
</tr>
<tr>
<td>SSE</td>
<td>Sum of Squared Errors</td>
</tr>
<tr>
<td>TRB</td>
<td>Transportation Research Board</td>
</tr>
<tr>
<td>UNC</td>
<td>University of North Carolina</td>
</tr>
<tr>
<td>VOC</td>
<td>Volatile Organic Compounds</td>
</tr>
<tr>
<td>volPM-oil</td>
<td>volatile primary Particulate Matter from engine lubrication oil</td>
</tr>
<tr>
<td>volPM-org</td>
<td>volatile primary Particulate Matter from fuel organics</td>
</tr>
<tr>
<td>volPM-sulf</td>
<td>volatile primary Particulate Matter from fuel sulfur</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Context

Aviation is a growing business in the United States and the world. Global passenger traffic has grown at 2.4 times the rate of the average GDP growth rate from 1960 until 1999, at an average annual rate of 9%, with a doubling time of 8 years [22]. Although the aviation sector stagnated for a few years after the terrorist attacks of 2001, and although there are limits to the growth of aviation due to both capacity [24] and political constraints, aviation demand is forecast to double again over the next 20 years [20][61][62][63][19]. The aviation industry currently contributes $640 billion to the economy in the United States, representing 5.4% of the GDP and employing more than 9 million people [24].

At the same time, there are negative environmental externalities associated with aviation. Globally, aircraft emissions of greenhouse gases and their precursors contribute to climate change [22][20]. On the regional scale, aircraft emissions impact air quality, which in turn influences public health [20]. Premature mortality [56], cardiovascular and pulmonary hospital admissions, and asthma aggravation are all linked to poor air quality [51]. On the local scale, aircraft noise depreciates housing capital [35][45], and causes sleep and school disturbances [10].

A critical component to the growth of aviation and its future sustainability is the mitigation of its environmental impacts. International and government agencies attempt to address some of these negative impacts via certification standards and operational restrictions. Internationally, the Committee for Aviation and Environmental Protection (CAEP), a sub-group of the International Civil Aviation Organization (ICAO), which is itself an
agency of the United Nations (UN), holds formal meetings every three years to help guide ICAO environmental policies [20]. In the United States, the Environmental Protection Agency (EPA) sets aviation emissions regulations, and the Federal Aviation Administration (FAA) sets noise standards and operating procedures. These regulations seek to strike a balance between allowing the industry to grow and limiting its environmental damage.

1.2 Aviation Environmental Portfolio Management Tool

In the US, the FAA is developing an integrated suite of software tools to assess the various impacts of aviation. One component of this tool suite is the Aviation Environmental Portfolio Management Tool (APMT). APMT is designed to assess industry economics and environmental impacts under various technological, operational, policy, market, and environmental scenarios. Along with other components of the tool suite (EDS and AEDT), it contributes to enabling a comprehensive cost-benefit analysis for assessing different options for mitigating environmental impacts of aviation [66][67]. The tool suite is divided into four major functioning components: technology tradeoffs, economic flows, aircraft operations, and impacts valuation. The Environmental Design Space (EDS) computes various candidate aircraft for introduction into the fleet as a function of the assumed technology level over time. The APMT – Partial Equilibrium Block (PEB) calculates economic flows in the aviation market iteratively by accounting for aircraft acquisition and operating costs, passenger fares, and the resulting passenger demand. Next, the Aviation Environmental Design Tool (AEDT) simulates aircraft operations and accumulates aircraft noise and emissions inventories. Finally, the APMT – Benefits Valuation Block (BVB) computes the resulting environmental impacts of aviation, and monetizes the impacts to allow a direct comparison of the costs and benefits of aviation environmental policies. [40]

The BVB represents the current understanding of the effects of aviation on the environment. The global climate, regional air quality, and local noise impacts are all accounted for using the impact pathway approach described in the European ExternE methodology [11]. The air quality component (AQ) of the BVB considers the changes in health impacts due to changes in ambient ozone and particulate matter levels. In particular, the AQ focuses on the impacts of changes in particulate matter (PM) concentrations using the intake fraction (iF) concept [15]. The focus on PM is justified due to the dominance of PM on the health
effects of aviation [39][40].

Particulate matter is categorized as either coarse (PM$_{10}$, diameter less than 10 microns) or fine (PM$_{2.5}$, diameter less than 2.5 microns). Fine particulate matter can deposit deeper in human lungs than coarse particulate matter, so its health impacts are more severe [11]. Incidentally, the focus of aviation health impact analyses is fine particulate matter (PM$_{2.5}$), since aircraft emissions of particulate matter fall into the PM$_{2.5}$ (hereafter referred to interchangeably as PM) category. Exposure to fine particulate matter increases the risk of various health endpoints [36][37]. Short-term exposure is linked to acute cardiac or respiratory events, while long-term exposure is associated with cardiovascular mortality, chronic respiratory disease, and lung cancer [56]. Thus the AQ uses the emissions outputs of the AEDT in conjunction with the iF approach and various concentration response functions (CRFs) to compute the increase in risk of these health endpoints that are attributable to aviation emissions.

1.3 Motivation

In a policy-making setting, the requirement for fast modeling tools is driven by the desire to analyze and compare the costs and benefits of multiple policy options for decades into the future. At the same time, model fidelity is also of prime importance in order to capture the effects of the policy scenarios as accurately as possible.

With respect to the AQ model, the Community Multiscale Air Quality modeling system (CMAQ)—the air quality modeling tool commonly used by the EPA for its regulatory impact analyses—is a high-fidelity model. However, its runtimes are prohibitively long for multiple policy analyses with 20 to 30 year time horizons. Indeed, the 108 month-long CMAQ simulations that were performed for this thesis required over 6 months to complete. This included data gathering, pre-processing of emissions, the actual model runs, post-processing and analysis of results, as well as two rounds of reruns to resolve processing issues. CMAQ is not well suited for multiple long-term policy analyses.

Therefore, the AQ model within the BVB uses a surrogate air quality model, namely the intake fraction model. The iF model greatly reduces run times, to the order of hours instead of months. Thus it represents a solid first step in the development of an air quality model appropriate for policy comparison settings. However, the source-receptor matrix
used in the intake fraction model was developed specifically for ground level mobile sources, and was at a county-level spatial resolution [15]. As such, the iF model was not optimized for aviation-specific emissions distributions.

Similarly, the EPA has developed a response surface model (RSM) for particulate matter for use in support of the National Ambient Air Quality Standards (NAAQS) analyses [58]. However, the scope of EPA PM RSM was far broader than that of the BVB iF model, including ground based point sources (factories, power plants, etc.), ground level mobile sources (automobiles, etc.), as well as ground level area sources (biogenic and agricultural emissions). Thus, although the EPA version of the PM RSM is promising, it is again not appropriate for the aviation policy setting.

Therefore, the goal of this thesis is to build a surrogate model that computes changes in ambient PM concentrations—and ultimately changes in health impacts—due to aviation emissions more accurately than the iF method and faster than CMAQ.

1.4 Contributions

The major contributions of this thesis are both a new surrogate air quality model for use in APMT, as well as the method used to develop the model. The PM Response Surface Model (RSM) represents a new method of computing changes in ambient PM concentrations due to various aviation emissions scenarios in the United States. The distribution of emissions (both horizontal and vertical), as well as the perturbations thereof used in the air quality model, are specific to aviation. The spatial resolution is finer than that of the iF model, using a 36km grid across the continental US as opposed to county level intake fractions. Finally, since the RSM is developed as a collection of individual grid-cell models, it has the ability to incorporate non-linear concentration response functions for the health impact analysis, in addition to the linear CRFs currently required for the iF method.

From a methodological viewpoint, the PM RSM is developed with flexibility in mind. Once the desired emissions trade-space is specified, the sampling method is chosen to allow for additional CMAQ runs if the surrogate model errors are too large, without un-balancing the RSM training set. Similarly, the RSM implementation is designed modularly, to allow for different model formulations. Finally, the error metric used to judge the performance of the RSM relative to CMAQ is defined considering the specific intended use of the RSM.
1.5 Thesis Outline

The thesis is organized in three major chapters: one for the conceptual approach of building a response surface model, another for the implementation details specific to the PM RSM, and a final chapter for the results and analysis of the final model. Chapter 2 presents the response surface methodology, Chapter 3 describes the implementation of the RSM, and Chapter 4 discusses the model validation, error analysis, and estimates for current aviation levels. Finally, Chapter 5 summarizes the thesis and gives recommendations for future work.
Chapter 2

Response Surface Model Approach

A response surface model (RSM) is a reduced-order approximation of a more complex and/or more time-consuming model. In typical mathematical or physical problems, the complexity of a model can be reduced by ignoring higher order terms or linearizing an equation. However, in computational applications with complex models or equations that are not known explicitly, the original full model to be approximated can often be treated as a black-box whose inner workings are not of interest. It is rather only a less complex approximation of the relationship between the inputs and outputs of the model that is desired. Thus a response surface model is simply a mapping of inputs to outputs that approximates the behavior of a more complex model or process.

The construction of a response surface model requires repeated sampling of the complex model in question, as well as a functional mapping of the inputs for those repeated trials to their outputs. A training set of model runs is used to generate the mapping of inputs to outputs, which then provides an approximation of the response of the complex model to new, un-sampled combinations of inputs. Specifically, the four components required to build a response surface model are:

1. the complex model to be approximated by the RSM,
2. a choice of input and output variables of interest,
3. a sampling technique that specifies what combinations of inputs to use in sampling the full model, and
4. a method of mapping any combination of the input variables to the output variables.
Once these four components are specified, a response surface model can be constructed.

This chapter presents the choices of the four components required to build a response surface model for the air quality impacts of aviation. Section 2.1 describes the air quality model used for the PM RSM, Section 2.2 covers the choice of input variables and output variables, as well as the ranges of the input variables that define the RSM emissions tradespace, Section 2.3 covers the combinations of the input variables used to sample the air quality model, and finally Section 2.4 presents the methods used to actually relate the inputs and outputs.

2.1 Air Quality Modeling

2.1.1 CMAQ Platform

The Community Multi-scale Air Quality model (CMAQ) is a three-dimensional atmospheric chemistry and transport modeling system. It is an Eulerian model, whose 3D grid is fixed relative to the ground after a proper map projection is applied. CMAQ computes the change in concentration of multiple chemical species in each grid cell for each time step by accounting for the emission, advection, diffusion, chemical formation, and removal of each species. Both gridded meteorology and gridded emissions input data are required in order to compute the advection and emission of each species in the model domain, as well as a set of initial and boundary conditions. The remaining diffusion, chemical formation, and removal of each species is computed within CMAQ at each time step in each grid cell as a function of those inputs [5]. CMAQ is capable of modeling at various geographic scales, from urban to regional to continental. It also treats multiple chemical species simultaneously in order to accurately model the interactions among them.

For the PM RSM work, CMAQ was used to simulate the response of PM concentrations to aviation emissions across the continental US. Table 2.1 lists the versions of the various CMAQ components used for the RSM simulations. CMAQ was used to perform 27 simulations for the PM RSM development. Due to the long CMAQ run times, the modeling work was split up amongst MIT, UNC, and BSU. Each was responsible for roughly one third of the simulations, while UNC provided technical guidance as well as processing the input emissions.
Table 2.1: CMAQ component versions.

<table>
<thead>
<tr>
<th>Compiler, library, or module</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>pgf90 (Fortran compiler)</td>
<td>7.0.2</td>
</tr>
<tr>
<td>gcc (C compiler)</td>
<td>4.1.1</td>
</tr>
<tr>
<td>NetCDF</td>
<td>3.6.0</td>
</tr>
<tr>
<td>I/O API</td>
<td>3.0</td>
</tr>
<tr>
<td>CMAQ</td>
<td>4.5</td>
</tr>
<tr>
<td>MCIP (met data pre-processor)</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Model Geographic Domain

The physical domain for the CMAQ runs includes the contiguous US, parts of Canada and Mexico, as well as Cuba. It extends from 126°W to 66°W longitude and from 24°N to 52°N latitude. The parameters for the 2-dimensional projection of the domain of interest are listed in Table 2.2. The model domain is then divided into 148 north-south columns, 112 east-west rows, and 14 vertical layers, with the horizontal edge of each grid cell measuring 36 km. Figure 2-1 (a) depicts the model domain, with the CMAQ grid overlaid on the map of the continental US.

Table 2.2: CMAQ grid projection parameters.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ₁</td>
<td>Projection type</td>
<td>Lambert conformal conic</td>
</tr>
<tr>
<td>Φ₂</td>
<td>Standard latitude 1 (southern)</td>
<td>33°N</td>
</tr>
<tr>
<td>γ</td>
<td>Standard latitude 2 (northern)</td>
<td>45°N</td>
</tr>
<tr>
<td>λₓ=0</td>
<td>Central meridian</td>
<td>97°W</td>
</tr>
<tr>
<td>λᵧ=0</td>
<td>Longitude of projection center</td>
<td>97°W</td>
</tr>
<tr>
<td>Φᵧ=0</td>
<td>Latitude of projection center</td>
<td>40°N</td>
</tr>
</tbody>
</table>

Representative Months

Each CMAQ run simulates one year of emission and dispersion of precursor species, and the subsequent formation and advection of air pollutants. However, due to the long CMAQ run times, the year-long simulation is broken up into four separate runs, with one representative month-long simulation for each season. These four months are February, April, July, and October, chosen to match the months simulated for the previous EPA RSM [58]. Further, each month-long simulation includes a 5-day run-up period in order to reduce the
dependence of the model outputs during the month of interest on the initial conditions of the simulation.

Weather and Background Emissions Data

The CMAQ simulations for the development of the PM RSM use the weather and background emissions data from the US Energy Policy Act of 2005 (EPAct) study [39]. The weather data is simulated data from the Mesoscale Model v5 (MM5) for 2001 [34], and the background emissions data is from the 2001 National Emissions Inventory (NEI) with aviation sources removed [60].

2.1.2 Aviation Emissions Inventories

In addition to the background emissions data, the CMAQ simulations also require aviation emissions in order to compute the changes in PM concentration due to aviation. Aircraft emissions at 325 airports were computed using the Emissions and Dispersion Modeling System (EDMS) and processed through the Sparse Matrix Operator Kernel Emissions (SMOKE) tool before use with CMAQ. The airport locations are shown in Figure 2-1(b). These inventories are for landing and take-off (LTO) operations only (i.e. only up to 3,000 ft). The baseline emissions at these airports—also borrowed from the EPAct study—represent 95% of commercial jet aircraft operations for which flight plans were filed between June 2005 and May 2006 [39]. Although the aviation emissions inventories were aggregated to the ground level, they need to be distributed in CMAQ’s vertical layers. Emissions at each of the 325 airports are therefore mapped to the vertical emission profiles at one of three representative airports. The three representative airports for which specific vertical emissions profiles have been developed are Atlanta (ATL), Chicago (ORD), and Providence (PVD), and the mapping is provided in Appendix C. The aviation emissions at each airport were then distributed within the first seven layers (corresponding to the first 3,000 ft) of the CMAQ modeling domain according to these three representative profiles.

These aviation emissions inputs were then multiplied by a constant to simulate an increase (or decrease) of aviation levels across the United States. Each emissions species was multiplied by a different constant to change the ratios of the different aviation emissions types in order to simulate a change in fleet composition. The multiplicative constants were the same at all airports, although in principle these can be varied in order to simulate
Airports in non-attainment areas
* Additional airports modeled

Figure 2-1: CMAQ simulation domain and modeled airport locations. Note, the CMAQ runs only include emissions at airports in the continental US. Source: EPAct study report draft [39].

regional differences in aviation. Changing the values of these emissions multiplicative factors provides the way of obtaining the different CMAQ samples required to build the PM RSM.

2.1.3 Post-Processing of PM Concentrations

Output Species

The final output of interest for the RSM is the annual average ground level ambient concentration of particulate matter. However, since particulate matter is comprised of many chemical species, these individual species must all be added together from the CMAQ outputs. The six PM components to be included in the total PM concentration are: ammonia, sulfates, nitrates, elemental carbon, organic carbon, and crustal material. The specific CMAQ outputs that are used for the PM RSM are listed in Table 2.3 [58].

<table>
<thead>
<tr>
<th>PM component</th>
<th>CMAQ name</th>
<th>Component CMAQ species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia (NH₄)</td>
<td>PM_AMM</td>
<td>ANH4I + ANH4J</td>
</tr>
<tr>
<td>Sulfates (SO₄)</td>
<td>PM_SULF</td>
<td>ASO4I + ASO4J</td>
</tr>
<tr>
<td>Nitrates (NO₃)</td>
<td>PM_NITR</td>
<td>ANO3I + ANO3J</td>
</tr>
<tr>
<td>Elemental Carbon</td>
<td>PM_EC</td>
<td>AECI + AECJ</td>
</tr>
<tr>
<td>Organic Carbon</td>
<td>PM_ORG_TOT</td>
<td>AORGAI + AORGAJ + AORGBI + AORGBJ + 1.167*(AORGPAI + AORGPAJ)</td>
</tr>
<tr>
<td>Crustal Material</td>
<td>PM_OTH</td>
<td>A25I + A25J</td>
</tr>
</tbody>
</table>
SMATing

The Speciated Model Attainment Test (SMAT) is a process used by the EPA to reconcile differences between simulated ambient pollutant concentrations and actual monitor values throughout the United States. EPA practice is to SMAT modeled air quality results before using them for health impact analyses [39][46][55][53]. SMATing involves two major steps: the first is interpolating the current (baseline) values of each of the six PM components across the domain of interest from the EPA monitor data, and the second is applying the ratios of policy modeled concentrations to baseline modeled concentrations to each of the six component species. The SMATed total PM concentrations are then simply the sum of the six SMATed component species, for each grid cell in the model domain.

The current version of the PM RSM does not use SMATed PM data since the PM SMATing software is not yet available. As a result, the RSM output PM concentrations are not directly applicable to health impact analyses. However, preliminary health impact results are presented later in this thesis, with the caveat that the analysis does not strictly follow EPA practices.

Particle-Bound Water

EPA health impact analyses also require using “wet” PM concentrations, meaning that the estimated PM masses must contain all particle-bound water [46][13][59][52]. This requirement stems from the fact that the concentration response functions (CRFs) used in health impact analyses were developed from EPA monitor values of PM concentrations, which include particle-bound water mass.

The current version of the RSM only uses “dry” PM concentrations, since only dry ion concentrations are averaged during each hour of CMAQ outputs. Although averaged water outputs are also available, the proper way to compute wet PM mass is to compute the water apportionment at each model time step, and then only average the wet PM masses. However, this was not done for the CMAQ simulation used in the RSM development, and a satisfactory method of apportioning water to the average dry concentrations of the different PM species has not been identified. Thus, once again, the health impact analysis in this thesis does not strictly follow EPA practices, and a proper water apportionment step would be required to reconcile this difference.
2.2 Emissions Trade-Space

Once the complex model that serves as the basis of the RSM has been selected, the next step is to choose the input and output variables to relate using the RSM. This section presents the choice of RSM input and output variables, as well as the ranges of each of those variables that define the design space of the RSM.

Choice of Response Surface Model Variables

The purpose of the PM RSM is to provide reasonably accurate estimates of changes in air quality due to changes in commercial aviation technology, operations, and fuels in the United States. More specifically, the RSM needs to relate changes in aircraft emissions to changes in ambient concentrations of particulate matter. Thus, the output of the RSM is chosen to be the annual average PM concentration (for each CMAQ grid cell), with the input variables chosen to represent the different types of emissions at different airports across the US. While the RSM output variable is well-defined, the input variables require closer consideration.

2.2.1 Input Variables

Two factors influence the choice of the specific input variables for the RSM. The first consideration is the number of input variables available. Since the process of mapping CMAQ inputs to outputs requires estimating RSM parameters, the more independent input variables there are, the more CMAQ runs are needed to get statistically significant estimates for those RSM parameters. Second, the ideal input variables should be uncorrelated so that the entire realm of possible future aviation emissions scenarios can be described using the least number of independent input variables.

Since CMAQ simulation times (including data acquisition, pre-processing of inputs, and post-processing of results) are considerably long, the most severe constraint on the choice of input variables comes from the extremely limited budget of training runs available. As only 27 CMAQ runs are available, no more than a few independent variables can be chosen without considerably sacrificing the accuracy of the model. This limit of only a few variables immediately restricts the choice of specific input variables. For instance, although the ideal RSM should respond accurately to changes in emissions at individual airports in the US,
since the aircraft emissions inventories used for the CMAQ runs include 325 airports, this level of independence is impossible with only a few input variables. Moreover, these 325 input variables can not be truly independent, since a large fraction of flights taking off from any airport in the US also land at another airport in the US; there are many domestic flights which link pairs of grid cells in the CMAQ domain. A reasonable compromise with regards to this spatial issue could thus involve grouping those 325 airports geographically into several (3-5) regions, or several airport types (e.g. large international vs. regional), for each of which the level of aviation emissions could be varied independently.

However, assuming that the typical policy for which the PM RSM is being developed has a larger impact on the aircraft fleet composition than on the spatial distribution of emissions—i.e. assuming that the regional effects of a policy are fairly similar, but that there may be a general trend in the US fleet—the focus of the choice of input variables shifts from geographic resolution to fleet or emissions inventory parameters. For example, a national NOX stringency policy that limits the mass of NOX that aircraft engines are allowed to produce per unit thrust [19][54] will probably have a greater impact on the national ratio of NOX emissions relative to non-volatile PM, SOX, and VOC emissions, and a lesser impact on the regional differences of those relative ratios. This shift in focus for the RSM input variables considerably reduces the set of variables to choose from, since there are far fewer types of aircraft emissions than there are airports in the United States.

The next steps are to consider which types of aircraft emissions have an impact on PM formation, describe the relationship between the various emissions and a set of independent variables, and choose a small subset of those independent variables as inputs to the RSM. Due to the long CMAQ run times previously alluded to, an upper limit of four independent variables was imposed on the RSM input.

**Impacts of Various Emissions on PM Formation**

Particulate matter is not composed of a single chemical species, but rather comprises many different compounds. Although primary PM species (non-volatile PM, as well as volatile PM from fuel sulfur content, fuel organics, and lubrication oil) are formed in the exhaust of aircraft engines, secondary PM formation occurs over minutes to days in the atmosphere through a series of chemical and microphysical processes involving sulfates, nitrates, and organic compounds [40]. Therefore, aircraft emissions of both primary PM as well as precursor
species to secondary PM formation need to be accounted for. Particularly, aircraft emissions of non-volatile PM (nvPM), volatile PM (volPM-sulfur, volPM-organics, and volPM-oil), nitrogen oxides (NO\textsubscript{x}), sulfur oxides (SO\textsubscript{x}), and volatile organic compounds (VOC) are all needed in order to obtain an estimate of the effect of aircraft operations on ambient PM concentrations.

Furthermore, a relationship between a small set of (in this case limited to four) independent variables and the different aircraft emissions types is also still required. For example, emissions of nvPM and NO\textsubscript{x} are correlated, since if aircraft operations at a given airport were to increase, one would expect both emissions of nvPM and NO\textsubscript{x} (as well as other species) to increase. Therefore, a method of decoupling the effects of policies on the different emissions types is needed in order to ensure that the final choice of RSM inputs involves only independent variables.

### 2.2.2 First Order Approximation

Presently, there is no standardized and validated measurement technique for PM, so the International Civil Aviation Organization (ICAO) has developed a method of approximating aircraft engine PM emissions until such a measurement technique becomes available. The latest version of this approximation, named the First Order Approximation version 3 (FOA3), provides a set of equations that relates aircraft fuel burn, fuel sulfur content, engine smoke number, and various emissions indices to primary PM emissions [69]. This set of equations is presented below (Equations 2.1 - 2.3) in simplified form. Similar equations for the emissions of secondary PM precursors are also included (Equations 2.4 - 2.6). Although the various emissions indices (EIs) are measured for individual aircraft engine types, an inventory (or fleet-wide) emissions index can also be defined for the various emissions types. The EI for a given emissions species is simply the mass of that species emitted per unit of fuel burned. These indices vary depending on aircraft type and specific operating condition. For the purposes of the PM RSM, the focus is on the fleet average emissions indices: the total mass of a species emitted divided by the total mass of fuel burned. Also of note is the fact that no method currently exists for estimating the impact of lubrication oil on primary PM emissions. FOA3 rather assumes that PM from lubrication oil is included.
in the organic fraction of volatile PM [69].

\[ nvPM = K_1 * (fuel burn) * (nvPM EI) \]  
\[ volPM_{sulfur} = K_2 * (fuel burn) * (fuel sulfur content) * (sulfur conversion efficiency) \]  
\[ volPM_{organics} = K_3 * (fuel burn) * (HC EI) * (HC conversion efficiency) \]  
\[ NO_x = K_4 * (fuel burn) * (NO_x EI) \]  
\[ SO_2 = K_5 * (fuel burn) * (fuel sulfur content) \]  
\[ VOC = K_6 * (fuel burn) * (HC EI) \]

Equation 2.1 is simplified by rewriting the original FOA3 equation in terms of EI as opposed to engine smoke number (SN). All of the equations are further simplified by lumping the various numerical constants, such as molecular weights and scaling factors based on trends, into a single constant for each equation (\(K_1\) through \(K_6\)). If the changes in the remaining variables (fuel burn, fuel sulfur content, etc.) are expressed as multiplicative factors or alternatively percent changes applied to some baseline emissions levels, the numerical values of these constants become inconsequential. Therefore, in order to simplify the equations, the independent variables that are to serve as inputs to the RSM are expressed as multiplicative factors.

Finally, Equations 2.1 - 2.6 explicitly show that the various emissions types are not independent. All of the emissions types scale with fuel burn, while each individual species is further determined by an emissions index (EI) or a conversion efficiency. It is only these component variables—fuel burn, nvPM EI, fuel sulfur content, sulfur conversion efficiency, HC EI, HC conversion efficiency, and NO\(_x\) EI—that are independent. Thus, it is apparent that the four RSM inputs need to be chosen from this set of independent variables that determine the overall emissions inventories.

Thus, since the sulfur conversion efficiency and the HC conversion efficiency can be regarded as constants, and since the HC EI is also held constant in order to reduce the dimensionality of the problem (due to its lower impact on public health [40]), the four remaining independent variables are fuel burn, fuel sulfur content, nvPM EI, and NO\(_x\) EI.
Final Choice of Input and Output Variables

In summary, this section described the rationale behind the selection of the RSM input and output variables. The final choices reflect the restriction to only four input variables imposed by the low number of CMAQ simulations available and the desire for independent inputs in order to more fully describe the emissions trade-space with fewer inputs. The input variables chosen are multiplicative factors with respect to the current fleet and aircraft operations for:

1. fuel burn,
2. fleet average fuel sulfur content,
3. inventory nvPM EI, and
4. inventory NOx EI.

The multiplicative factors for these variables are each fixed to be constant over the geographic domain. This restriction limits what types of policies can be analyzed using the PM RSM to only national level policies with small regional differences in effects, but it is necessary in order to reduce the dimensionality of the problem.

Finally, the output variable of interest is the annual average ambient concentration of PM at the Earth’s surface, since it is this ultimate PM concentration that impacts public health. The PM RSM thus needs to map multiplicative factors for fuel burn, average fuel sulfur content, inventory nvPM EI, and inventory NOx EI to changes in ambient PM concentrations.

2.2.3 Input Variable Ranges

Having chosen the RSM input variables, the next step is to set the range of each variable that define the RSM design space. The design space needs to be large enough to cover the range of possible aviation emissions scenarios for up to 20 years in the future, the typical policy decision horizon. However, the larger the design space, the farther apart the 27 sample points are within that space, so the lower the accuracy of the RSM due to possible non-linearities in the response of PM concentrations to emissions. The rationale for the ranges of the input variables is as follows:
Fuel Burn Multiplier

Aircraft fuel efficiencies have historically improved at a rate of about 1%-2% per year [31][68][19]. However, since the RSM is meant to model aviation emissions scenarios for up to 20 years into the future, it is the number of aircraft flying, not their individual fuel efficiencies, that is most important in determining the range of fuel burn multipliers to use for the CMAQ runs. Since aviation forecasts project demand to grow by as much as 150% over the next 20 years at certain particular airports [61][62][63][19], the annual 1%-2% improvements in engine fuel efficiencies will be overtaken by the increases in traffic volume. Thus, in order to model the full range of aviation scenarios from no aviation to a 150% increase, the range of the multiplicative factors applied to the baseline total fuel burn is set at 0 to 2.5.

US-Average Fuel Sulfur Content Multiplier

Reducing the fuel sulfur content has a direct impact on reducing PM emissions. The current fuel sulfur standard for aviation is 3000 parts-per-million (ppm) [21]. While most Jet A aviation fuel contains 400-800 ppm sulfur [40][69], it is conceivable that a policy reducing the aviation fuel sulfur content to around 15 ppm (the new highway diesel sulfur content standard [57]) could be introduced in the next 20 years. There is also a possibility that aviation fuel sulfur level will increase—for example, if the aviation sulfur standard is not lowered and the lowest sulfur fuel goes to other users. Therefore, the range for the fuel sulfur level must include the current fuel standard (3000 ppm) as the upper bound, as well as the new highway diesel standard (15 ppm) as the lower bound. Assuming the current fleet average is 600 ppm, the range of multiplicative factors applied to the baseline fuel sulfur content is set at 0.025 to 5.

Inventory NO\textsubscript{X} Emissions Index Multiplier

NO\textsubscript{X} EIs have increased historically as engine pressure ratios have increased [19][54]. The current fleet exhibits a wide range of NO\textsubscript{X} EIs due to a large distribution of engine pressure ratios and rated thrust, suggesting perhaps a large range on the NO\textsubscript{X} EI percent change. However, previous emissions stringency scenario analyses show that the percent changes in NO\textsubscript{X} inventories are much smaller than the percent stringencies in emissions. For instance, CAEP IP13 shows that a 30% NO\textsubscript{X} stringency (reducing the NO\textsubscript{X} emissions certification
standard by 30%) only yields up to a 3.8% reduction in fleet NO\textsubscript{x} emissions inventories over an 18-year period [19]. This is due to the fact that inertia in the fleet prevents the full impact of a given stringency from being observed until much later in the future, when all the aircraft which are currently above the new stringency level are retired from service. With this idea of fleet inertia in mind, the desired range for the NO\textsubscript{x} EI multiplicative factors is rather small, but considering that CMAQ may not be able to capture such small differences between simulations, the range is fixed at 0.7 to 1.1.

**Inventory nvPM Emissions Index Multiplier**

Finally, since FOA3 is an approximation method to be used until more reliable PM measurement methods become available, and since no similar PM policy analyses exist to date, the future trends in fleet nvPM EI are far less certain. Further, there is a large discrepancy between the PM estimates given by the FOA3.0 method developed by ICAO and the FOA3a methodology which is a more conservative approach currently used by the EPA. The difference in fleet primary PM emissions is a factor of 4.4 between the two methods [46]. The inventories used for the CMAQ runs used FOA3.0 to calculate aviation PM emissions, so the range on the inventory nvPM EI multiplier must capture both the large discrepancy built into FOA3 and the uncertainty in the future fleet nvPM EIs. Therefore, the range of multiplicative factors applied to the baseline inventory nvPM EI is set at 0.25 to 3.6.

**Response Surface Model Emissions Trade Space**

The emissions trade space for the PM RSM is thus defined by the four independent input variables and their ranges. These ranges are summarized in Table 2.4 below. The design points that make up the training set of CMAQ runs for use in the RSM development must fall within this emissions trade-space. Furthermore, any new input combinations for which

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel burn multiplier</td>
<td>0</td>
<td>2.5</td>
</tr>
<tr>
<td>Fuel sulfur content multiplier</td>
<td>0.025</td>
<td>5</td>
</tr>
<tr>
<td>Inventory NO\textsubscript{x} EI multiplier</td>
<td>0.7</td>
<td>1.1</td>
</tr>
<tr>
<td>Inventory nvPM EI multiplier</td>
<td>0.25</td>
<td>3.6</td>
</tr>
</tbody>
</table>
the CMAQ behavior is to be estimated using the PM RSM should also fall within this
design space. In essence, the design space defines the range of inputs for which the RSM is
optimized, and running the RSM with inputs outside of this space may yield much higher
errors.

2.3 Sampling Technique

Having defined the RSM trade-space through the specification of the input variables and
their ranges, the final step in defining the design of experiments (i.e. determining the
specific inputs to be used in the CMAQ simulations) is the choice of the sampling technique
employed to explore the design space. This section presents the choice of the sampling
method used to pick points within that design space to serve as the training set for the
RSM.

The goal of the sampling technique is to cover the design space uniformly so as to avoid
areas with too many samples, or worse, areas that are not sampled at all. In covering the de-
sign space, the sampling technique should explore various combinations of the independent
inputs so that their relative importance and effects on the outputs can be estimated.

Finally, since the degree of non-linearity in the response of PM concentrations to the
various emissions species is currently unknown, and since the initial set of 27 CMAQ sim-
ulations may be insufficient for the RSM (i.e. if the RSM yields errors that are deemed
unacceptably high), the flexibility of adding more points to design of experiments later
without un-balancing or re-computing the training set is highly desirable.

Here, several different approaches are possible, each with various advantages and disad-
vantages discussed below.

2.3.1 Full Factorial

The simplest sampling method is to define a given number of “levels” (numerical values)
for each “factor” (independent variable), and then sample each possible combination of the
resulting factor levels. Each variable can have the same number of levels, or there could be
different numbers of levels for each variable. The resulting set of points is guaranteed to
cover the design space uniformly in each dimension (and in all dimensions if each factor has
the same number of levels), and the combinations of inputs required are easy to determine.
However, there are two major drawbacks of the full factorial method. The first is the exponentially increasing number of samples required as the number of levels is increased. For example, to create a full factorial sample with only two levels for each of the four input variables requires $2^4 = 16$ CMAQ runs. However, if the number is increased to three levels per factor, the number of simulations required grows to $3^4 = 81$. Thus, the large number of points required for a full factorial sample make this approach rather unattractive. The second drawback is the impossibility of adding new points to the design of experiments without unbalancing the training set, so this also makes the full factorial approach unsuitable for the PM RSM.

2.3.2 Latin Hypercube

The Latin hypercube is another common sampling technique that addresses the problem of exponentially increasing samples with added factor levels. A Latin hypercube experimental design involves choosing as many sample points as there are levels for each factor, and then distributing those points throughout the design space such that the projection of those points onto each dimension is distributed evenly. In other words, each level of each factor must be sampled exactly once (instead of as many times as there are levels for the remaining variables, as is the case for a full factorial design). Consequently, Latin hypercubes require that the number of levels for each factor be the same.

Although Latin hypercubes reduce the number of sample points required in a high-dimensional design space, they can potentially leave large portions of that space “uncovered”. While “space-filling” methods can address this by moving the Latin hypercube points to cover the design space more uniformly, the problem of choosing where to place new points while keeping the training set balanced persists.

2.3.3 Centroidal Voronoi Tessellation

A more sophisticated approach that ensures even coverage of a design space, regardless of its dimensionality or the initial number of samples desired, is a Centroidal Voronoi Tessellation (CVT) [9]. CVT distributes the requested number of points throughout a specified high-dimensional space through iterative calculations since there is no direct formula for determining the final locations of the points. However, the biggest disadvantage of the CVT technique is that the set of sample points becomes unbalanced when adding a new sample
point. Thus all of the previous sample locations must be recomputed when adding a new point in order to ensure that the entire new set is still evenly distributed. Although the iterative calculations can be performed very quickly on a computer, the drawback of CVT as a sampling technique for the CMAQ runs is that changing the previous set of input combinations requires that all previous CMAQ simulations be re-run.

In summary of the first three options discussed, Figure 2-2 depicts a full factorial design, a Latin hypercube, and a Centroidal Voronoi Tessellation in the unit square.

![Typical sampling techniques](image)

Figure 2-2: Typical sampling techniques: a full factorial design with 16 points (a), a Latin hypercube with 15 points (b), and a centroidal voronoi tessalation with 14 points (c). Notice that for the full factorial approach (a), the number of sample points must equal the product of the number of levels for each factor, while for the Latin hypercube (b), the number of levels must be the same for each factor. Neither of these restrictions apply to the CVT design (c), which seeks to cover the design space uniformly without knowledge of specific factor levels.

### 2.3.4 Low-Discrepancy Sequence

It becomes apparent that the ideal sampling technique should require few points, should ensure that those points are distributed evenly, and should provide a way to place new points in the design space without un-balancing or recomputing the locations of the previous points. Low-discrepancy sequences do exactly that; the discrepancy of the sequence is simply a measure of the departure of its distribution of points from a uniform distribution. A low discrepancy sequence is thus a deterministic sequence of points, in any number of dimensions, whose distribution approaches a uniform distribution as more and more points are added.

Such a sequence was developed by van der Corput in one dimension [64]. Hammersley
extended the sequence from one dimension into any \( k \)-dimensional space by combining \( k \) van der Corput sequences with different prime bases [17]. Finally, Halton modified the Hammersley sequence by removing the dependence of the sequence on the number of sample points. While Hammersley required that the base for the sequence in the first dimension be the number of points desired \( N \), Halton dropped this first dimension to create a \((k - 1)\)-dimensional sequence from a \( k \)-dimensional Hammersley sequence [16]. Although the Halton sequence is not as efficient as the Hammersley sequence, in the sense that its discrepancy decays slower as the number of points increases, the Halton sequence has the desirable property of being extensible from \( N \) to \( N + 1 \) points without recomputing the first \( N \) points.

For the PM RSM development, a Halton low-discrepancy sequence is chosen to provide the combinations of input variables for the CMAQ simulations. As desired, the Halton sequence can be used to compute additional combinations of inputs without affecting the previous combinations of inputs. Thus the training set of CMAQ runs can be expanded easily as required for the PM RSM. Figure 2-3 shows the first 10, 30, and 120 points of the Halton sequence in the unit square. As can be seen in the figure, one caveat is that low-discrepancy sequences only approach the ideal uniform distribution as the number of points increases. Therefore, when the dimensionality of the design space is especially large, or the number of sample points is especially small (Fig. 2-3 (a)), a low-discrepancy sequence may not necessarily exhibit a low discrepancy (i.e. the distribution may not be very uniform). It remains to be determined whether this is a problem for the PM RSM, for which the number of input variables (the dimensionality) is four and the number of runs available (the number

Figure 2-3: 2D Halton sequence: first 10 (a), first 30 (b), and first 120 points (c).
of sample points) is only 27.

CMAQ Input Combinations

Twenty-seven CMAQ simulations are used for the PM RSM. This includes 24 runs whose input combinations are chosen using a 4-dimensional Halton sequence, plus three additional runs for the cases with all the multipliers set to their minimum values (i.e. the no aviation case), all the multipliers set to 1 (i.e. current levels of aviation), and all the multipliers set to their maximum values (i.e. farthest corner of the design space). The numerical values of the four input emissions multipliers for these 27 CMAQ runs are tabulated in Appendix A.

2.4 Model Formulation

The final step in creating a response surface model is the specification of the mapping of input variables to output variables. The RSM is then constructed by solving a set of equations relating the inputs and outputs of the training runs in order to determine the parameters of the specified mapping. Finally, the RSM can then estimate the behavior of the underlying complex model by using those parameters in combination with any new set of inputs to approximate the resulting outputs. This section contrasts regression and interpolation approaches, and describes the kriging interpolant used in the PM RSM effort.

2.4.1 Regression

Regression methods treat the outputs of a complex process as a function of the inputs, a set of parameters to be estimated, and a random error term. The random error comprises that part of the variation in the outputs of the process that can not be explained using only the inputs and the specified regression model. A typical regression equation has the form

$$ y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_1 x_1^2 + \beta_2 x_1 x_2 + \beta_2 x_2^2 + \ldots + \epsilon, \quad (2.7) $$

where $x_1, x_2, \ldots$ are the input variables, $\epsilon$ is the error term, and $\beta_0, \beta_1, \beta_3, \ldots$ are the model parameters to be estimated. The regression prediction $\hat{y}_0$ at a new location $\bar{x}_0 = (1, x_1, x_2, x_1^2, x_1 x_2, x_2^2, \ldots)$ is then written as

$$ \hat{y}_0 \equiv \hat{y}(\bar{x}_0) = \bar{x}_0 \hat{\beta}, \quad (2.8) $$
where \( \beta^* \equiv (\beta_0^*, \beta_1^*, \beta_2^*, \beta_{11}^*, \beta_{12}^*, \beta_{22}^*, \cdots)^T \) is the optimum set of parameters.

The parameters of the model are optimized to minimize the error term in some predefined sense. For example, ordinary least-squares (OLS) regression—by far the most common type of regression—minimizes the sum of squared errors (SSE) between the values predicted by the model and the actual output values used to construct the model. Formally, OLS solves for the parameters \( \beta^* \) that minimize

\[
\text{SSE} \equiv \sum_{i=1}^{n} (\hat{y}_i - y_i)^2,
\]

where \( \hat{y}_i \) is the output value estimated by the model for input vector \( \bar{x}_i \) using Equation 2.8, \( y_i \) is the actual output value at \( \bar{x}_i \) supplied in the training set, and \( n \) is the number of points in the training set. The optimum parameter vector \( \beta^* \) is then given by

\[
\beta^* = \left( X^T X \right)^{-1} X^T \bar{y},
\]

where \( X = [\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n]^T \) is an \( n \times p \) matrix of trial locations, \( \bar{y} \) is an \( n \times 1 \) vector of observed outputs, \( n \) is again the number of observations in the training set, and \( p \) is the number of \( \beta \) parameters to solve for.

The regression model also provides an estimate of the uncertainty on the predicted output value \( \hat{y}_0 \). The \( 1 - \alpha \) confidence interval for the regression prediction is given by

\[
\hat{y}_0 \pm t \left( 1 - \frac{\alpha}{2}, n - p \right) \sqrt{\frac{\text{SSE}}{n - p} \left[ 1 + \frac{1}{\bar{x}_0^T \left( X^T X \right)^{-1} \bar{x}_0^T} \right]},
\]

where \( \hat{y}_0 \) is the regression prediction from Equation 2.8, and \( t \left( 1 - \frac{\alpha}{2}, n - p \right) \) is the Student’s \( t \)-distribution. The regression prediction interval incorporates both the uncertainties on the \( \beta \) parameters and the uncertainties on the error term \( \epsilon \).

The magnitude of the error term can be reduced by using higher-order equations for the regression model, but there is a tradeoff between the order of the regression model and the statistical significance of the parameters estimated. The higher the order of the equation in the regression model, the more parameters need to be estimated, the fewer degrees of freedom in the model, so the less likely each of those parameters is to be statistically significant. Similarly, the more input variables there are, the more parameters are required. In the extreme case, the model is said to over-fit the data if the number of parameters in
the model approaches the number of sample points. Finally, regression methods require a certain set of assumptions to be satisfied, including that the error be a random variable with zero mean and a normal distribution, that the error random variable have a constant variance, etc. [30] The validity of the regression model deteriorates quickly with departures from these assumptions.

One disadvantage of regression methods for the PM RSM is that the reduced order model represented by the regression parameters can only capture those non-linearities that are assumed to exist in the first place. For example, a second-order regression model can be built with a constant term, a linear term for each dimension, and all combinations of quadratic terms (including all variable interactions, as shown in sample Equation 2.7), but the complex model upon which the regression model is built may exhibit higher-order non-linearities. These higher-order interactions would not be captured by the model and would instead be lumped into the error term. Therefore, the use of a regression model for the RSM requires some prior assumptions for the shape of the response of PM concentrations to each of the input emissions variables.

On the other hand, if such prior knowledge does exist for the dependence of PM on certain inputs, then a regression formulation can enforce that knowledge within the model. That model behavior is guaranteed to persist in un-sampled regions of the design space as well, so regression may be an attractive choice if the design space is only sparsely sampled.

Another disadvantage of the regression technique is that it allows for an error term in the first place. In other words, there is no guarantee that if the same set of inputs is supplied to the regression model as is used for one of the CMAQ runs in the training set, the same outputs would result. Since the goal of the RSM is to mimic the behavior of CMAQ with as little error as possible, regression may or may not be the most appropriate choice for the PM RSM. However, since there are only 27 sample points available in a 4-dimensional design space, a regression model is built in order to compare its performance to that of a kriging interpolation method at off-design locations in that design space.

Figure 2-4 compares such a regression model to a kriging interpolation approach (which is discussed below) for a sample training set with eight points in one dimension.
First-Order OLS Regression and 95% Confidence Bands

Figure 2-4: First-order OLS regression model (a) and kriging interpolation model (b) for a set of eight points in one dimension (the y-axis is the response of interest, so the x-axis represents the only input variable). The solid points are the points in the training set, the dashed curves represent the mean responses of the models, and the dotted curves represent the upper and lower bounds of the 95% confidence bands for each model.

2.4.2 Interpolation

Interpolation methods address the issue of missing the design points by constraining the response surface to pass through each of the points supplied in the training set. While there are several common types of interpolation methods, each with different behaviors between the points given in the training set, they all ensure that the resulting function returns the same outputs as were used to create the function when the same sets of inputs are supplied. More formally, if $y_i$ is the full-model output for input vector $\tilde{x}_i$ and $\hat{y} = f_{\text{interp}}(\tilde{x})$ is an interpolation response surface built using $\tilde{x}_i$ as one of its training points, then $\hat{y}_i = f_{\text{interp}}(\tilde{x}_i)$ is guaranteed to match $y_i$. Possible interpolation techniques include piecewise constant (also termed nearest-neighbor) interpolation, piecewise linear interpolation, and higher-order polynomial interpolations.

Another interpolation method that was originally developed for geostatistics applications is kriging [7][1][65][42][44]. Its ability to recreate complex non-linear surfaces using relatively few training points also make kriging suitable for computer experiments. In particular, if the computer model to be approximated with a response surface is deterministic (i.e. the same outputs result every time the same inputs are used), kriging is often a better choice for mapping inputs to outputs than a regression. This method was also used by the US
EPA in its previous PM response surface modeling work [58]. Therefore, due to both its desirable properties as an interpolant, as well as its previous use in very similar settings, kriging is also selected as another option for mapping inputs to outputs for the PM RSM.

2.4.3 Ordinary Kriging Equations

The kriging estimate of the output value at a new site in the design space is a linear combination of the output values at the training sites. Specifically, the correlation between the output at the new point and the outputs at each of the training points is a function of the Euclidean distances between the new point and the training points. Ordinary kriging models the output value as a constant but unknown mean plus a deviation from that mean that is the realization of a stochastic process with a pre-defined spatial correlation structure. While the correlation structure must be assumed, the parameters of the correlation function and the process variance are estimated from the training data set.

The equations for ordinary kriging are as follows: let \( \mathbf{y} = (y_1, \cdots, y_n, 0)^T \) represent the model outputs in one grid cell for each of the \( n \) runs and \( \hat{y}_0 \) represent the desired estimate of the outputs in that grid cell for a new combination of inputs. Then,

\[
\hat{y}_0 \equiv \hat{y}(\bar{x}_0) = \begin{pmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{pmatrix}^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \bar{w}^T \mathbf{y},
\]

(2.12)

where \( \bar{w} \) is the vector of weights to apply to the previous "observations" (i.e. full model runs), and \( \mu \) is the mean of the previous observations. The weights of this linear combination are then determined according to

\[
\bar{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma(\bar{x}_1, \bar{x}_1) & \cdots & \gamma(\bar{x}_1, \bar{x}_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(\bar{x}_n, \bar{x}_1) & \cdots & \gamma(\bar{x}_n, \bar{x}_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma(\bar{x}_1, \bar{x}_0) \\ \vdots \\ \gamma(\bar{x}_n, \bar{x}_0) \\ 1 \end{pmatrix} = G^{-1} \mathbf{y},
\]

(2.13)

where \( x_1 \) through \( x_n \) are the \( n \) (one for each CMAQ run) \( m \)-dimensional (four dimensions for the PM RSM) input vectors. Here, \( \gamma \)—known as the "variogram function"—is defined
\[
\gamma(x_\alpha, x_\beta) = \sigma^2 \left[ 1 - \rho(x_\alpha, x_\beta) \right],
\]  
(2.14)

where \( \rho(x_\alpha, x_\beta) \) is the correlation function and \( \sigma^2 \) is the variance of the stochastic process. The correlation function defines the behavior of the interpolation between the sampled points, and it is a function of \( d(x_\alpha, x_\beta) \), the Euclidean distance between those points. Some possible correlation functions are listed below [33]:

\[
\rho_{\text{exp}}(d(x_\alpha, x_\beta)) = e^{-\theta d}
\]  
(2.15)

\[
\rho_{\text{gauss}}(d(x_\alpha, x_\beta)) = e^{-\theta d^2}
\]  
(2.16)

\[
\rho_{\text{lin}}(d(x_\alpha, x_\beta)) = \max\{0, 1 - \theta |d|\}.
\]  
(2.17)

Figure 2-5: Some typical correlation functions: exponential (a), Gaussian (b), and linear (c) for \( \theta = 0.5 \) (dotted), \( \theta = 1 \) (dashed), and \( \theta = 3 \) (solid).

Figure 2-6: Some typical variogram functions: exponential (a), Gaussian (b), and linear (c) for \( \theta = 0.5 \) (dotted), \( \theta = 1 \) (dashed), and \( \theta = 3 \) (solid). Here, \( \sigma^2 \) is the variance of the stochastic process, and it is a multiplicative constant in the variogram function (Eq. 2.14).
Both $\theta$ and $\sigma^2$ from Equation 2.14 must be estimated using maximum likelihood methods from the training data set supplied. Figure 2-5 plots the three correlation functions from Eq. 2.15 - 2.17, and Figure 2-6 plots the resulting variogram functions. Kriging also gives an error estimate $\sigma_{OK}^2$, which can be used to create confidence intervals around the estimated $\hat{y}_0$ values:

$$\sigma_{OK}^2 = \left( \begin{array}{c} w_1 \\ \vdots \\ w_n \\ \mu \end{array} \right)^T \left( \begin{array}{c} \gamma(\bar{x}_1, \bar{x}_0) \\ \vdots \\ \gamma(\bar{x}_n, \bar{x}_0) \end{array} \right) = \bar{w}^T \bar{\gamma}$$

Finally, universal kriging extends upon ordinary kriging by allowing for a general linear trend underlying the response surface instead of just a constant mean. Universal kriging is essentially a regression model with a kriging model built on the regression residuals. By incorporating the underlying regression model, universal kriging can improve the performance of the interpolation through the assumption of the general underlying trend in the data.
Chapter 3

RSM Implementation

Having established the conceptual basis for the RSM modeling approach, there are some intricacies that need to be described in the implementation of the PM RSM. This chapter motivates and describes these implementation details. Section 3.1 briefly describes the distinction between the design space multipliers and the emissions multipliers, Section 3.2 discusses the need for individual grid-cell models, Section 3.3 covers the spatial interpolation step required before running the RSM, Section 3.4 describes how the RSM handles uncertainty, and Section 3.5 summarizes the main assumptions and limitations of the current version of the PM RSM.

3.1 Inventory Multipliers vs. Design Space Multipliers

The design space for the PM RSM is set using the four independent variables and their ranges given Section 2.2. However, three of the four independent variables—fuel sulfur content, inventory NOxEI multiplier, and inventory nvPM EI multiplier—do not directly influence the formation of PM. Rather, it is only in conjunction with the fuel burn multiplier that they determine the total aircraft emissions, which then affect PM formation. Although the design space is specified using those four independent variables, CMAQ computes PM concentrations as a function of total aircraft emissions, so the PM RSM needs to be built upon these emissions multipliers in order to mimic the behavior of CMAQ. Therefore, the RSM is built to take in the multipliers of fuel burn, SO2, NOx, and nvPM.
3.2 Grid-Cell Models

The most direct way to build the PM RSM is to treat the concentration of particulate matter in each grid cell as a function of the cell’s row and column in the CMAQ grid, as well as the values of the four inventory multipliers. The resulting PM RSM would consist of a single 6-dimensional regression or kriging model. However, since there are 16,576 (= 148 \times 112) grid cells in the domain and therefore 447,552 (= 16,576 \times 27) 6-dimensional data points to use when building the RSM, the matrices needed to create the kriging model become exceedingly large. In particular, the “variogram matrix” $G$ from Equation 2.13 would be a 447,553 \times 447,553 square matrix. Clearly, this storage requirement precludes the use of a single 6-dimensional kriging model for the PM RSM.

At the other end of the spectrum, the PM RSM can be comprised of 16,576 different 4-dimensional kriging or regression models, one for each grid cell in the domain. This option greatly reduces the storage requirement for the individual models, but it requires computing 16,576 different models. Between these two extremes, the PM RSM can be built using multiple regional 6-dimensional models. However, since the choice of the boundaries of these regional models is arbitrary, and since there is no obvious disadvantage to building individual models for each grid cell, the current version of the PM RSM employs this second option. Although this is not a crucial issue for the regression formulation, both the regression and kriging RSMs are built upon individual grid-cell models for consistency.

The kriging versions of the RSM were implemented using a Matlab toolbox from the Technical University of Denmark. The Design and Analysis of Computer Experiments (DACE) toolbox allows the user to specify different underlying polynomial surfaces, as well as different correlation functions for the stochastic process of the kriging interpolant. The kriging RSMs use a Gaussian correlation function as was done in the previous EPA PM RSM [58], and they compare a 0th order underlying surface to a 1st order surface.

3.3 Spatial Interpolation

Regardless of how the RSM is constructed, both steps of building the RSM and then running it require that the RSM input multipliers be known in each grid cell in the domain. This is not a problem when building the RSM since the CMAQ simulations were all run with constant multipliers across the 325 airports. These multiplier values can simply be extended
to all the other cells in the domain that do not contain airports when building the model. However, when the RSM is run with a new emissions inventory, there is no guarantee that the RSM input multipliers will be constant across all the airports. Therefore, a method of computing the input multipliers in each grid cell is required. Moreover, four separate spatial interpolations are required, one for each of the four independent RSM input variables. Once again, there are several options for this spatial interpolation step.

The most obvious “interpolation” method is to simply average the multiplier values at all the airports in the inventory, and then apply these average multipliers to all the grid cells in the domain. The opposite extreme of this method is to use the multipliers at the single nearest airport for each cell in the CMAQ grid. This latter option is basically a piecewise constant interpolation in two dimensions, or simply a nearest-neighbor interpolation. Between these two extremes, the multiplicative factors at any grid cell in the domain can be computed as the average of the factors at the \( n \) nearest airports. For the first extreme, \( n \) is simply set to 325, while for the other extreme, \( n \) is set to 1.

Another similar option is—for each cell in the grid—to average the multiplicative factors at all the airports within a “radius” \( r \) of the current grid cell. The radius \( r \) can either be specified as a number of grid cells or simply a distance in km. However, the main problem with this approach is that unless \( r \) is set to a very large value (either in grid cells or in km), there may be regions in the grid that are not close enough to any airports. This method could thus break down in the western US where airports are farther apart than, for example, in the northeast.

Two subtleties have not yet been addressed for either of these two methods: the first is the distance from the grid cell of interest to each of the airports in the grid, and the second is the “size” of each airport in the grid. If the first method of averaging the multipliers at the nearest \( n \) airports is employed, the distance from each of those \( n \) airports to the current grid cells needs to be accounted for. This is because, intuitively, the change in PM concentration in any grid cell in the domain is likely to be more correlated to changes in emissions at airports closer to the grid cell, and less correlated to changes at airports farther from it. If \( n \) is set to 1 for the nearest-neighbor approach, then the distance to the nearest airport is not an issue. However, if \( n > 1 \), some inverse distance weighting is required to reduce the influence of airports that are far from the current grid cell relative to the influence of nearer airports.
Similarly, the “size” of the airports needs to be accounted for when \( n > 1 \). If the three nearest airports to a given grid cell are equidistant from that grid cell, but one of those airports is a large international hub while the other two are much smaller regional airports, then the changes in emissions at the large airport are likely to be more influential than the changes at the two smaller airports. Whether the “size” of the airport is measured in flight operations, passenger traffic, or simply total fuel burn, it needs to also be incorporated as a weighting factor to the multipliers at each of the nearest \( n \) airports. Finally, both the distance issue and the size issue also apply to the second approach (with the maximum radius \( r \)). What is required is a flexible way of incorporating these different ideas in a single formulation for the spatial interpolation.

### 3.3.1 Parametric Interpolation

The following parametric interpolation technique is introduced for the RSM: the multiplicative factor for every grid cell in the domain is expressed as a linear combination of the multipliers in the \( n \) nearest grid cells that contain airports. The RSM input multiplier for grid cell \( i \) is

\[
f_i = C_i \sum_{j=1}^{n} w_{ij} f_j,
\]

where \( f_j \) is the multiplicative factor at airport-containing grid cell \( j \), \( w_{ij} \) is the weight on \( f_j \) for grid cell \( i \), and \( C_i \) is a normalization factor. The weight for each airport-cell \( j \) is given by

\[
w_{ij} = \frac{s_j^\alpha}{d_{ij}^\beta},
\]

where \( s_j \) is the “size” of airport-cell \( j \), \( d_{ij} \) is the Euclidian distance between grid cells \( i \) and \( j \), and \( \alpha \) and \( \beta \) are parameters to be set. In the current version of the RSM, the “size” variable is set to the total fuel burn in each airport-containing grid cell, and \( n \) is set to 325 to include all the airports in the inventory. Finally, the normalization constant for grid cell \( i \) is simply

\[
C_i \equiv 1 / \sum_{j=1}^{n} w_{ij}.
\]

The normalization constant ensures that if all the multiplier values at all the airport-containing grid cells are the same, the interpolation method gives that same multiplier
value for the entire domain: if \( f_j = f, \; \forall \) airport-cells \( j \),

\[
f_i = C_i \sum_{j=1}^{n} w_{ij} f = f \; C_i \sum_{j=1}^{n} w_{ij} = f, \; \forall \text{ grid cells } i.
\] (3.4)

Equation 3.4 shows how \( C_i \) guarantees that the inputs to an RSM run with the same emissions inventory as one of the CMAQ simulations will match the inputs used to create the RSM in the first place (see the first paragraph of Section 3.3). Recall that kriging is an interpolant as well (Sec. 2.4), so that if the same inputs are given to a kriging RSM as were used to create its kriging models for every grid cell in the domain, the kriging RSM outputs are guaranteed to match the CMAQ outputs. On the other hand, this is not guaranteed for the regression RSMs.

Table 3.1: Interpolation methods as a function of the \( n, \alpha, \) and \( \beta \) parameters. Note: \( m \) is the total number of points in the training set, so setting \( n = m \) means simply using all the available points as opposed to only a subset of the nearest points.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>Interpolation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m )</td>
<td>1</td>
<td>0</td>
<td>simple weighted average</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>piecewise linear (in one dimension with uniform divisions)</td>
</tr>
<tr>
<td>( m )</td>
<td>0</td>
<td>1</td>
<td>inverse distance</td>
</tr>
<tr>
<td>( m )</td>
<td>0</td>
<td>2</td>
<td>inverse distance squared</td>
</tr>
<tr>
<td>1</td>
<td>--</td>
<td>--</td>
<td>nearest-neighbor (piecewise constant)</td>
</tr>
<tr>
<td>( m )</td>
<td>0</td>
<td>( \infty )</td>
<td>nearest-neighbor (piecewise constant)</td>
</tr>
</tbody>
</table>

Figure 3-1: Common spatial interpolation methods for a data set with 1-dimensional inputs: nearest-neighbor (piecewise constant) (a), piecewise linear (b), and inverse distance (c).

By varying the parameters \( n, \alpha, \) and \( \beta \) appropriately, the parametric interpolation scheme in Equations 3.1 - 3.3 can match many common interpolation techniques. Table 3.1 presents the various parameter settings required to produce these common techniques. Fig-
ure 3-1 depicts a few of these methods for a sample data set with 1-dimensional inputs. Figures 3-2 and 3-3 depict the effects of increasing $\alpha$ and $\beta$, respectively, for a sample data set with 2-dimensional inputs given in Table 3.2. Finally, Appendix B includes similar plots for a sample inventory with the 325 airports located in their appropriate CMAQ grid cells.

Table 3.2: Spatial interpolation data set for Figures 3-2 and 3-3.

<table>
<thead>
<tr>
<th>Column</th>
<th>Row</th>
<th>&quot;Size&quot;</th>
<th>&quot;Multiplier&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>54</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>49</td>
<td>109</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>71</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>104</td>
<td>81</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>139</td>
<td>28</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>

![Figure 3-2](image1.png)

Figure 3-2: Sample spatial interpolation with: $\alpha = 0$ (a), $\alpha = 1$ (b), and $\alpha = 9$ (c). Note, the "airport" in the bottom-right corner is twice as "large" as the other airports, so it’s value dominates as $\alpha$ grows.

![Figure 3-3](image2.png)

Figure 3-3: Sample spatial interpolation with: $\beta = 1$ (a), $\beta = 3$ (b), and $\beta = 9$ (c). Notice that the interpolation approaches a nearest-neighbor scheme as $\beta$ grows.

As a side note, the parametric interpolation method described in this section is just another type of interpolation, as is kriging itself. However, unlike kriging which optimizes...
a parameter to fit the training data, this parametric interpolation is heuristic. It does not adapt to any underlying data, but rather requires the values of its parameters to be pre-specified. Thus it seems that a kriging approach may also be useful for the spatial interpolation step. However, since the CMAQ runs use constant multipliers at all airports, the training set of CMAQ output data contains no information about the relative effects of the changes in emissions at each airport on the concentrations in each grid cell. If the CMAQ simulations had all been performed with varying multiplier values across the domain, a secondary kriging model could be used for the spatial interpolation of multipliers. But since the limited number of CMAQ runs only permits the multipliers to be changed consistently at all airports, the spatial interpolation in the current version of the RSM relies on this parametric interpolation technique.

3.4 Model Uncertainty

As with any model, there are many uncertainties in the PM RSM that need to be addressed. These include uncertainties in the input data, inherent uncertainties in regions of the design space that were not sampled with CMAQ, as well as uncertainties on various parameters used throughout the model. In order to capture the effects of these different types of uncertainty on the RSM outputs, the RSM can be run stochastically using a Monte Carlo approach. Ascertaining the relative effects of the different types of uncertainty can then guide further model development, whether this leads to more work on the RSM input data, more CMAQ runs to reduce the RSM uncertainty, or more research on the effects of air pollution on public health, in order to get a more robust estimate of the health impacts of aviation. This section describes the various types of uncertainty in the RSM and how they are handled within the model.

3.4.1 Emissions Inventories

Aviation emissions inventories are themselves computed using models. Depending on the particular model in use, different airports, different types of operations, different methods of estimating PM precursor emissions, or different landing and take-off profiles. Therefore there is uncertainty in the emissions inventories even for any year in the past, not just for a future scenario. Thus, the RSM needs to account for the uncertainty on its emissions inputs.
The current version of the RSM uses the emissions uncertainty distributions developed by Rojo for a version of the health impact model based on intake fractions [40].

3.4.2 Spatial Interpolation

As described in Section 3.3, the spatial interpolation step is required when running the RSM. However, the interpolation parameters need to be specified before running the RSM, so the choice of parameter values is rather arbitrary. At this time, it is unknown what specific values of \( n, \alpha, \) and \( \beta \) optimize RSM performance since there is no CMAQ data with different multipliers at different airports to compare to. Therefore, the RSM allows the user to input an uncertainty distribution for each of these parameters. The current version of the RSM uses \( n = m = 325, \alpha = \{0, 1\} \) with equal probability, and \( \beta = \{1, 2, 3, 9\} \) with equal probability. The choice of the \( s_j \) weighting variable is also somewhat arbitrary, and the current PM RSM uses the total fuel burn in each airport-cell as the “size” \( s_j \) of that cell.

3.4.3 Regression and Kriging Uncertainty

Both regression and kriging provide error estimates (Eq. 2.11 and 2.18, respectively) as well as mean value estimates. These error estimates can be used to create uncertainty distributions around the mean value estimates, and so running the RSM in Monte Carlo mode allows the RSM output to have an uncertainty distribution.

As shown in Figure 2-4, the kriging uncertainty is zero when the model is run with the same input that was previously used to build the model, is higher for input values between these training set inputs, and is highest for input values outside the intended design space. In other words, the magnitude of the kriging uncertainty is directly proportional to the distance from a new sample point to the previous training points in the design space. Similarly, the the regression prediction interval grows with the distance from the centroid of the input locations.

This observation has a crucial impact on the regression or kriging uncertainty compared to the other types of uncertainty in the RSM: the magnitude of the uncertainty on the RSM outputs is not fixed but is rather a function of the RSM inputs themselves. The implication is that if there is uncertainty on the input inventories, the regression or kriging uncertainty may require different treatment than the other types of uncertainty. For now, the choice
of how to treat the regression or kriging uncertainty is left to the user. A conservative approach that overestimates the uncertainty on the final outputs is to assume that the regression or kriging uncertainty is additive to the input emissions and spatial interpolation uncertainties.

### 3.4.4 Health Impacts

Since the RSM is intended for use in policy analyses in computing the effects of aviation emissions on risks of adverse health impacts, the RSM also incorporates the concentration response functions (CRFs) used by Rojo in the intake fraction model [40]. In order to compute health impacts from changes in PM concentrations, the model also needs population data, baseline incidence data, as well as valuation data to monetize the results. All four of these data types—population, baseline incidences, CRFs, and endpoint valuation—also have varying degrees of uncertainty that need to be accounted for. The population data used in the RSM is different from that in the intake fraction model since the RSM needs the population in each grid cell as opposed to in each county. Therefore the RSM uses the Socioeconomic Data and Applications Center’s Global Rural-Urban Mapping Project (SEDAC-GRUMP) population database\(^1\). The GRUMP database is at a 30 arc-second resolution, and is aggregated up to the coarser CMAQ grid resolution [49] as shown in Figures 3-4 and 3-5. However, the current RSM uses the same baseline incidence, CRF, and valuation data and uncertainty distributions as the Rojo intake fraction model.

### 3.4.5 Sequential Introduction of Uncertainty

Finally, using certain types of uncertainty can drastically increase RSM run times. For example, the spatial interpolation step is very time-consuming, so running the RSM with uncertainty on the spatial interpolation parameters greatly increases run times. Therefore, the RSM introduces the different types of uncertainty sequentially through the impact pathway from changes in emissions inventories to changes in health impacts. Emissions uncertainty is introduced first, then spatial interpolation uncertainty, regression or kriging uncertainty, and finally health impact uncertainty. At each of these steps, the user can either choose a "single value" (meaning no uncertainty) input, or stochastic inputs drawn from an uncertainty distribution.

\(^{1}\)available at http://sedac.ciesin.org/gpw/global.jsp
Figure 3-4: Population in each grid cell in the domain.

Figure 3-5: Log_{10} of population in each grid cell in the domain.
3.5 Assumptions and Limitations

Along with the uncertainty on the input data and model parameters discussed in the previous section, there are also some assumptions that are embedded in the RSM. These are assumptions in the emissions processing, air quality modeling, as well as in the RSM specification, and they impose some limits on the applicability of the RSM. This section summarizes the various assumptions and their implied limitations for the current version of the RSM.

3.5.1 Emissions Processing

LTO Emissions

First, the emissions inventories used for the CMAQ simulations for the current RSM development only account for LTO emissions. That is, only emissions in the first 3,000 ft above ground level are accounted for. The underlying assumption is that cruise emissions—any emissions at an altitude above 3,000 ft—have a negligible impact on PM concentration at the Earth’s surface. However, recent preliminary results show that the impact of cruise emissions may have been significantly underestimated, and may in fact be on the same order of magnitude as the impact of LTO emissions themselves [50][2]. Therefore the current version of the RSM very likely underestimates total PM formation due to aviation emissions, although the extent of the downward bias is not known at this time.

DoE Bounds

The choices of input variable ranges discussed in Section 2.2 are rationalized based on historic trends in the aircraft fleet, as well as projections of future aircraft operations. However, as was discovered when using emissions inventories from a different model as inputs to the RSM, the differences in inventory estimates between different emissions models can be larger than the original ranges specified for the RSM design of experiments (DoE). Specifically, the inventory NOx EI in the AEDT emissions inventory for 2005 is 12% higher than the inventory NOx EI in the CSSI emissions inventory which was used in the CMAQ simulations. But the upper bound for the inventory NOx EI multiplier for the RSM is only 1.1 (i.e. only a 10% increase; see Tbl. 2.4), so running the RSM with AEDT baseline inventories already means sampling a point outside the RSM intended design space. As
is shown in Figure 2-4, the regression or kriging uncertainty can grow quickly outside the model’s design space.

**Emissions Model Assumptions**

The CMAQ simulations for the PM RSM development were performed using emissions inventories computed with the EDMS model. However, the emissions inventories for APMT are computed using the AEDT. While the EDMS inventories from CSSI are for an entire year of operations, use actual taxi times, and include auxiliary power unit (APU) emissions, AEDT only computes a representative day of emissions which is then multiplied by 365, assumes 19 minutes for taxi out and 7 minutes for taxi in for all operations, and does not include APU emissions. Thus, care must be taken when the assumptions on the inventories do not match those on which the RSM was constructed, since the ultimate health impacts computed using the RSM will reflect those differences in the input emissions.

**Representative Airports**

The CSSI emissions inventories used to run CMAQ for the RSM development are for aircraft operations at 325 of the largest airports in the United States. Although these inventories account for 95% of the commercial aviation emissions [39], they may not account for aviation emissions quite properly. Specifically, the vertical distribution of emissions are only currently available for three airports (ATL, ORD, and PVD), so the distributions at the remaining airports are assumed to be similar to one of these three representative vertical profiles. This is probably the least consequential assumption for the emissions processing, so it does not impose any significant restrictions on the usage of the RSM.

**3.5.2 Air Quality Modeling**

**Meteorology**

The CMAQ simulations for the RSM sample points are only for four representative months out of the year. These representative months, borrowed from the previous EPA RSM modeling effort, are assumed to represent the weather patterns for each season. However, the sensitivity of the CMAQ results to the choice of representative months is not fully understood. In fact, the sensitivity of grid cell PM concentrations to meteorology in general
is not well known. The choice of a 5-day run-up period is also consistent with the previous EPA PM RSM work, but it too is rather arbitrary. The CMAQ results would no doubt be different if a 10-day or 15-day run-up period were used, although the significance of these differences is also unknown.

National Level Policies

A further limitation of the RSM is that it is only applicable to policy analyses that do not cause large differences in regional effects. Due to the low number of CMAQ simulations, the input emissions multipliers were altered uniformly across the domain. Therefore the RSM was built upon a training data set which contains no information about the effects of regional policies. This necessitates the spatial interpolation step discussed in Section 3.3, and greatly limits the applicability of the RSM to regional policies. Although the spatial interpolation technique can certainly handle input emissions multipliers that vary greatly across the domain, the uncertainty on the interpolation parameters translates into uncertainty on the input multipliers in those grid cells without airports. Conversely, since most of the 325 airports are located in grid cells with fairly high populations, and since the grid cells with the highest spatial interpolation uncertainty are likely to have lower populations because they are farther from these large airports, it may be the case that the effect of spatial interpolation step on the accuracy of the health impact estimates may not be all that large. The magnitude of this uncertainty is unknown.

3.5.3 RSM Construction

SMATing and Water Apportionment

Since the PM SMATing software is not yet available, the PM concentration data used in the RSM construction is not SMATed. Similarly, since the particle-bound water mass was not computed for the current CMAQ data, the PM data used for the RSM is only for dry PM masses. For both of these reasons, the impacts computed using the PM concentrations from the RSM may not accurately capture the effects of aviation on public health. As mentioned in Section 2.1, this limitation means that the RSM health impact analysis, as is currently implemented, does not follow EPA practice.
3.5.4 RSM Uncertainties in the Context of Paired Analyses

Finally, the PM RSM is intended for use in policy analyses. Its purpose is to compare the health impacts of various emissions scenarios resulting from different aviation policies. Therefore, to compute the incremental change in health impacts due to an aviation policy, the impacts of the baseline emissions scenario must be subtracted from the impacts of the policy emissions scenario. Although the PM RSM estimate of health impacts for any single emissions scenario reflects the various uncertainties described in Section 3.4, the uncertainty of the difference between two emissions scenarios is much lower. This is due to the fact that many of the uncertainties in the RSM are independent of the emissions scenario in question.

Paired Monte Carlo analyses are used to compute the probability distribution of the incremental health impacts for policy analyses. For those uncertainties that are common between the baseline and policy scenarios, the same random draw of an uncertain parameter is used in both the baseline and policy RSM run. For example, if for a particular Monte Carlo run for the baseline scenario a value of the CRF for premature mortality is picked that is 10% higher than its nominal value, that same draw of the CRF value is used for the corresponding Monte Carlo run for the policy scenario. The effect of the paired Monte Carlo approach is to greatly reduce the uncertainty on the differential impact of an aviation policy as compared to the impact estimate for just the baseline or just the policy scenario.
Chapter 4

Results

Several response surface models for PM concentrations are built using the concepts from Chapter 2 and implementation details from Chapter 3. The RSM errors are compared for the various model formulations, and the model with the smallest error is selected. The error decay with additional CMAQ runs is presented next, and finally the RSM is used to estimate the current health impacts of aviation emissions. Section 4.1 describes the RSM error analyses, and Section 4.2 presents the RSM estimates of the current health impacts of aviation, in comparison to both CMAQ and the intake fraction method.

4.1 RSM Error Analysis

This section describes the cross-validation approach used to quantify the RSM errors and their decay with additional CMAQ runs. Three types of RSMs are considered, with the final model selected on the basis of these cross-validation results.

4.1.1 Cross-Validation

The goal of model validation is to ascertain the prediction error of the final RSM constructed. This can be achieved by using the RSM to predict the output for a new combination of inputs that is not included in the training set, and comparing the RSM prediction to the actual CMAQ output for that same combination of inputs.

Leave-one-out or n-fold cross-validation is a common model validation technique that achieves this goal but does not require additional samples that are not used in the training set. It consists of building and testing multiple different models, each of which is built...
upon a different subset of the existing training samples, and computing the prediction error for each model. Specifically, if there are \( n \) samples in the training set, leave-one-out cross-validation involves building \( n \) different models, each of which is built on \( n-1 \) of the available samples and omits 1 sample. Each sample is successively omitted and is then compared to the model prediction for that input in order to get an estimate of the model error. Thus, leave-one-out cross-validation yields \( n \) errors between the model prediction and the actual value, one at each of the \( n \) sample locations.

While this technique only provides an estimate of the model error for the final model, it can easily be extended to analyze error decay with the addition of more sample points. Leave-\( k \)-out cross-validation leaves out \( k \) samples instead of only 1 sample, and constructs the models using the remaining \( n-k \) samples. The total number of models that can be constructed is given by

\[
nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.
\]

Thus, leave-\( k \)-out cross-validation obtains \( k \cdot \binom{n}{k} \) prediction errors, \( k \) errors for each of the \( nC_k \) models built. By successively setting \( k = 1, 2, 3, \ldots \), leave-\( k \)-out cross-validation yields the prediction errors for models built upon \( n-1, n-2, n-3, \ldots \) samples. The behavior of these prediction errors as the number of sample points is increased reveals the decay or convergence of the model errors.

**Error Metric**

Having decided on the validation technique, the actual error metric to compute for each model still needs to be specified. Since there are \( 148 \times 112 = 16,576 \) grid cells in the CMAQ/RSM domain, and therefore 16,576 individual kriging or regression models for each RSM, there are actually 16,576 errors for each comparison of an RSM prediction to an actual CMAQ output. These errors differ both in the absolute and relative sense across the domain, depending on the distances to from each grid cell to the various airports.

Therefore, a summary error metric is desired for each RSM to CMAQ comparison in the cross-validation analysis. Some common options include vector norms and normalized norms. These metrics are given in Eq. 4.2 and 4.3, where \( \vec{x} \) is the vector of errors. As can be seen in their definitions, the standard vector norm accumulates errors, and both the vector norm and the normalized norm are sensitive to outliers. The sensitivity of the norm
to outliers is directly controlled by the $p$ parameter.

\[
\text{vector } p\text{-norm } = \|\vec{x}\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \tag{4.2}
\]

\[
\text{normalized } p\text{-norm } = \left( \frac{1}{n} \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \tag{4.3}
\]

In the extreme case of $p = \infty$, the infinity norm simply returns the maximum absolute value in the vector $\vec{x}$: $\|\vec{x}\|_{\infty} = \lim_{p \to \infty} \|\vec{x}\|_p = \max(|x_i|)$. For $p = 1$, the 1-norm is just the sum of all the elements in $\vec{x}$ and the normalized 1-norm is the arithmetic mean of $\vec{x}$, while for $p = 2$ the normalized 2-norm is also known as the root-mean-square error:

\[
\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \tag{4.4}
\]

Although the 2-norm is typically used in computational modeling applications, the choice of $p$ for the RSM error analysis would still be somewhat arbitrary.

Moreover, the vector norm is only a way of computing a summary metric from a set of errors. The errors themselves can be defined in an absolute sense, a relative sense, or in some other fashion. For the RSM, the absolute error $\varepsilon_i$ in each grid cell for run $i$ can be defined as the difference between the RSM estimate and the CMAQ actual PM concentration in that grid cell:

\[
\varepsilon_i = c_i^{\text{RSM}} - c_i^{\text{CMAQ}}. \tag{4.5}
\]

Similarly, the RSM relative error ($\rho_i$) can then be defined as the absolute error divided by the CMAQ concentration in that grid cell:

\[
\rho_i = \frac{\varepsilon_i}{c_i^{\text{CMAQ}}}. \tag{4.6}
\]

Further, the absolute change in concentration due to aviation emission scenario $i$ ($\alpha_i$) can be defined as the difference in PM concentration between the CMAQ outputs for run $i$ and run 000:

\[
\alpha_i = c_i^{\text{CMAQ}} - c_{000}^{\text{CMAQ}}. \tag{4.7}
\]

With this last definition in mind, the RSM error relative to the contribution from aviation
can be defined as the ratio of absolute error to absolute change in concentration due to aviation, in each grid cell:

$$\psi_i = \varepsilon_i / \alpha_i.$$  \hfill (4.8)

Thus, it becomes clear that once again, several options are available for the specific error metric to use, and that the choice thereof is also rather arbitrary. Each of the error metrics discussed—$$\varepsilon_i$$, $$\rho_i$$, and $$\psi_i$$—may vary greatly across the domain, and for each there may be grid cells whose error values are outliers. As mentioned previously, different vector norms can be used to change the relative impacts of these outliers on the final aggregate metric desired.

However, since the RSM is ultimately used to compute changes in health impacts due to different aviation scenarios, the outlying errors in grid cells with high populations are of more concern than the errors in grid cells with low populations. In light of the several options for the error metric as well as the several options for the vector norm and the complication of population weighting, a more appropriate error metric is needed for the RSM-specific error analysis.

In order to derive the equation for the RSM-specific error metric, the equation for change in impacts due to change in PM concentrations needs to be considered. Restricting the analysis to linear concentration response functions (CRFs), the change in incidence of endpoint $$j$$ due to aviation scenario $$i$$ ($$dI_{i,j}$$) is as follows:

$$dI_{i,j} = \sum_{k=1}^{n} K_j \alpha_{i,k} R_j P_k$$

$$= K_j R_j \sum_{k=1}^{n} (c_{i,k} - c_{000,k}) P_k,$$  \hfill (4.9)

where $$K_j$$ is the linear CRF coefficient for endpoint $$j$$, $$\alpha_{i,k}$$ is the change in concentration in grid cell $$k$$ due to aviation scenario $$i$$ (as defined in Eq. 4.7), $$R_j$$ is the baseline incidence rate for endpoint $$j$$, $$P_k$$ is the population in grid cell $$k$$, and $$n$$ is the total number of grid cells. Note that since the CRF coefficient $$K_j$$ and the baseline incidence rate $$R_j$$ are typically defined as national averages, they are constant across all grid cells $$k$$. The relative RSM error in the change in incidences of endpoint $$j$$ for aviation scenario $$i$$ ($$RE_{i}^{dI}$$) can then be defined as the difference between the RSM estimate of $$dI$$ and the CMAQ actual $$dI$$ divided
by the CMAQ actual $dI$:

$$RE_i^{dl} = \frac{dI_i^{RSM} - dI_i^{CMAQ}}{dI_i^{CMAQ}}$$

$$= \frac{K_j R_j \sum_{k=1}^{n} \left( c_{i,k}^{RSM} - c_{000,k}^{CMAQ} \right) P_k - K_j R_j \sum_{k=1}^{n} \left( c_{i,k}^{CMAQ} - c_{000,k}^{CMAQ} \right) P_k}{K_j R_j \sum_{k=1}^{n} \left( c_{i,k}^{CMAQ} - c_{000,k}^{CMAQ} \right) P_k}$$

$$= \frac{\sum_{k=1}^{n} \left( c_{i,k}^{RSM} - c_{000,k}^{CMAQ} - c_{i,k}^{CMAQ} + c_{000,k}^{CMAQ} \right) P_k}{\sum_{k=1}^{n} \left( c_{i,k}^{CMAQ} - c_{000,k}^{CMAQ} \right) P_k}$$

$$RE_i^{dl} = \frac{\sum_{k=1}^{n} \varepsilon_{i,k} P_k}{\sum_{k=1}^{n} \alpha_{i,k} P_k}$$  \hspace{1cm} (4.10)$$

Note that since both $K_j$ and $R_j$ are constant across grid cells $k$, they cancel in the equation for $RE_i^{dl}$. Therefore, this metric does not depend on the particular endpoint $j$ in question, but rather applies to all endpoints that are computed with a linear CRF. Note also that in the special case when $\varepsilon$ and $\alpha$ are constant across the domain (i.e. if $\varepsilon_{i,k} = \varepsilon_i$, $\alpha_{i,k} = \alpha_i$, $\forall k = \{1, \ldots, n\}$), $RE_i^{dl}$ reduces to $\psi_i$ as defined in Equation 4.8.

Thus, the $RE_i^{dl}$ error metric, defined specifically for the PM RSM with linear CRFs, is more meaningful for the purposes of the RSM than arbitrarily using $\varepsilon_i$, $\rho_i$, or $\psi_i$ alone. It also avoids the somewhat arbitrary choice of a specific vector norm to use when aggregating the 16,576 individual errors into a single summary metric. Finally, it takes into account the relatively higher importance of errors in grid cells with high populations. Although the RSM can handle non-linear CRFs for which $RE_i^{dl}$ would not be the most appropriate error metric, the current version only includes the linear CRFs encoded in the previous intake fraction model. Therefore, it is this $RE_i^{dl}$ error metric that is used in the RSM cross-validation error analyses.

### 4.1.2 Final Model Selection

Cross-validation methods are typically used to select the most appropriate model for a specific application. As such, several candidate RSM formulations are proposed and examined before the final RSM is selected. As discussed in Section 2.4, regression models can be computed quickly, but they may not return the training set outputs exactly, and they require some assumptions about the underlying behavior of the output of interest. Depending on
how well the regression model explains the variation observed in the training data, those assumptions may or may not be appropriate. On the other hand, interpolation models such as kriging do return the training set outputs exactly. However, they also require a set of assumptions that influence how well the model behaves away from the design sites used to create the model. In order to determine the final RSM model formulation, three types of models are compared using the cross-validation approach.

The first model is an ordinary kriging model. The PM concentration in each grid cell \( \hat{Y}_1(\mathbf{x}) \) is modeled as constant mean \( \beta_0 \) with Gaussian deviations from that mean as a function of the specific combination of input multipliers supplied \( Z(\mathbf{x}) \):

\[
\hat{Y}_1(\mathbf{x}) = \beta_0 + Z(\mathbf{x}).
\]  

(4.11)

The second model is a universal kriging model that assumes a linear underlying surface as opposed to a constant value. This model incorporates the assumption that since aviation emissions account for only a small fraction of total emissions, the response of PM concentrations is nearly a linear function of the aviation emissions. Thus the PM concentration \( \hat{Y}_2(\mathbf{x}) \) is modeled as a 4-dimensional linear surface \( \mathbf{x}^T \beta \) with Gaussian deviations from that underlying surface \( Z(\mathbf{x}) \):

\[
\hat{Y}_2(\mathbf{x}) = \mathbf{x}^T \beta + Z(\mathbf{x}).
\]  

(4.12)

The third and final model considered is an ordinary-least-squares (OLS) regression model. This model relies solely on the assumption of linearity, and does not force the response surface to pass through each of the points in the training set exactly. The PM concentration \( \hat{Y}_3(\mathbf{x}) \) is modeled as the 4-dimensional linear surface \( \mathbf{x}^T \beta \) that minimizes the sum of squared errors between the RSM predicted and CMAQ actual PM concentrations:

\[
\hat{Y}_3(\mathbf{x}) = \mathbf{x}^T \beta.
\]  

(4.13)

The leave-one-out cross-validation is performed for each of these three model formulations. The prediction errors are computed at each of the 27 design sites for each of the three models, and are presented in Figures 4-1 and 4-2. As can be seen in the figure, the kriging model with a linear underlying surface performs much better than the kriging model with only a sample mean. The largest relative error for the model with the constant assump-
Figure 4-1: $RE^d_i$ for each of the 27 CMAQ runs for the leave-one-out cross-validation for the kriging models with constant (dots) and linear (stars) underlying surfaces. The root-mean-square of the $RE^d_i$ metric across the 27 runs is 0.09981 for the kriging model with a constant underlying surface (dots) and 0.01314 for the kriging model with a linear underlying surface (stars).

Figure 4-2: $RE^d_i$ for each of the 27 CMAQ runs for the leave-one-out cross-validation for the kriging model with linear underlying surface (stars) and linear regression model (squares). The root-mean-square of the $RE^d_i$ metric across the 27 runs is 0.01314 for the kriging model with a linear underlying surface (stars) and 0.01085 for the linear regression model (squares).
tion is 44.958%, with a root-mean-square error of 9.98%. The largest relative error for the model with the linear assumption is 4.86%, and the root-mean-square error is 1.314%. This reflects the fact that if most of the variation in the training data set can be explained by a linear underlying surface, the remaining deviations from that underlying surface are much smaller than the deviations form the sample mean. Thus the Gaussian deviations imposed by the kriging model are much smaller for the model with the linear underlying surface, and the observed prediction errors are smaller at new locations in the design space. On the other hand, the response of the kriging model with the constant underlying surface tends to return to the sample mean as the distance from the previous trial sites increases. The specific output values supplied in the training set are fit using Gaussian deviations from the sample mean, but as the distance from the design sites increases, the effects of those deviations diminishes. This leads to large errors away from the initial trial sites if there is any general trend in the data other than a constant mean.

However, a more surprising result is that the linear regression model performs better than the kriging model with the linear underlying surface. Although kriging is an exact interpolant that returns the CMAQ output exactly at each of the trial sites, the purpose of the cross-validation analysis is to investigate the expected prediction errors at new, yet untested locations in the design space. The results of the leave-one-out cross-validation in Figure 4-2 show that the regression model performs slightly better than the kriging model in this regard. The maximum error for the regression model is roughly 4.346%, and the root-mean-square error is 1.085%. This finding may be due to the relatively few data points in the model training set. Although a statistically significant linear regression model can be obtained with 27 points in four dimensions (i.e. with 22 degrees of freedom), 27 points may not be enough for a robust 4-dimensional kriging model. The performance of the model in regions of the design space away from the initial trial sites may be compromised in order to enforce the interpolant quality of the kriging model at those initial trial sites.

Thus, due to the slightly lower prediction errors, reduced complexity, and shorter run-time of the regression model compared to the kriging model, the linear regression is chosen as the final RSM formulation. As a side note, the results for the regression formulation are for models built with only 26 observations. The regression residuals are slightly smaller for the model built upon all 27 CMAQ simulation, as presented next in the error decay analysis.
4.1.3 Error Decay with Additional CMAQ Runs

A leave-k-out cross-validation analysis is performed for the OLS PM RSM, with \( k = \{1, 2, 3, 4, 5\} \). As shown in Equation 4.1, the number of possible models grows very quickly as \( k \) increases. Thus the full cross-validation is only performed for \( k = \{1, 2\} \), and a subset of 135 models is used for \( k = \{3, 4, 5\} \). Table 4.1 summarizes the maximum number of models \( (nC_k) \), the actual number of models constructed, and the resulting number of errors computed for each value of \( k \).

Figure 4-3 illustrates the number of times each of the 27 CMAQ runs is sampled for the \( k = \{3, 4, 5\} \) cases. The \( k = \{3, 4, 5\} \) cases correspond to RSMs built using subsets of 24, 23, and 22 of the CMAQ runs, respectively. Since the specific subset of combinations used in the \( k = \{3, 4, 5\} \) cases is chosen at random from the full set of possible combinations of CMAQ runs, there is no guarantee that each CMAQ run is left out the same number of times as every other CMAQ run. Indeed, the distribution on the number of times each run is left out in the leave-k-out cross-validation is non-uniform, as is shown in Figure 4-3. These non-uniform distributions caused by the restriction to a subset of all the possible RSMs for \( k = \{3, 4, 5\} \) therefore add some uncertainty on the results of their respective leave-k-out cross-validations. Moreover, there is no uncertainty on the results for \( k = \{1, 2\} \) (i.e. RSMs with 26 and 25 CMAQ runs, respectively) since all possible RSMs are built for these cases and each CMAQ run is left out the same number of times.

With this caveat in mind, Figure 4-4 presents the reduction of the root-mean-square of the \( RE_i^{df} \) error metric, computed across the 27 CMAQ runs, as the number of CMAQ runs used to build the RSM is successively increased. Figures 4-2 and 4-1 shows both the prediction errors for the leave-k-out cross-validations, as well as the regression residuals for each RSM built. The root-mean-square of the regression residuals is always lower than the root-mean-square of the cross-validation prediction errors since the regression residuals are computed for all 27 trial sites for each model built, while the cross-validation prediction errors are only computed for those runs omitted from each model built. The results for \( k = \{3, 4, 5\} \) (i.e. models with 24, 23, and 22 trial sites, respectively) are increasingly more uncertain since a smaller and smaller proportion of the total possible combinations of models is created as \( k \) is increased (Tbl. 4.1). The resulting uncertainty is more obvious for the plot of cross-validation prediction errors, and less so for the regression residuals. This
Table 4.1: Cross-validation combinations

<table>
<thead>
<tr>
<th>Number of CMAQ runs left out ((k))</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum possible combinations (\binom{n}{k})</td>
<td>1</td>
<td>27</td>
<td>351</td>
<td>2,925</td>
<td>17,550</td>
<td>80,730</td>
</tr>
<tr>
<td>Number of RSMs created</td>
<td>1</td>
<td>27</td>
<td>351</td>
<td>135</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>Percent of maximum</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>4.62%</td>
<td>0.77%</td>
<td>0.17%</td>
</tr>
<tr>
<td>Number of errors computed</td>
<td>0</td>
<td>27</td>
<td>702</td>
<td>405</td>
<td>540</td>
<td>675</td>
</tr>
</tbody>
</table>

Figure 4-3: Histograms for the number of times each CMAQ run is left out of an RSM, and hence the number of errors that are computed for each CMAQ run, for the leave-\(k\)-out cross-validation with: \(k = 3\) (a), \(k = 4\) (b), and \(k = 5\) (c).

RSM Error Convergence

Figure 4-4: Convergence of the root-mean-square of the \(RE^U\) metric; computed only for the CMAQ runs left out of each successive leave-\(k\)-out cross-validation (squares), and across all CMAQ runs for each cross-validation (circles).
is again due to the fact that the regression residuals are computed for all 27 trial sites, not just for those sites that are omitted from each model, so the effects of the uncertainty on the prediction errors are dominated by the regression residuals at the remaining trial sites.

Figure 4-4 shows that the RSM errors are indeed reduced with the addition of more CMAQ simulations. However, the errors are already on the order of only 1%, so the marginal benefit of even more CMAQ runs may not be too great. The reduction of the regression residuals is fairly linear, with a reduction of just over 0.01% with each additional run. The trend in the cross-validation prediction errors is less certain. Finally, although the root-mean-square of the regression residuals is only around 0.098% for the final RSM built, the prediction error at new trial sites not yet simulated in CMAQ will probably be closer to the 1.085% prediction errors observed in the cross validation analysis.

Figure 4-5 depicts the $R^2$ values across the domain for the final RSM built. As discussed in Section 3.2, the RSM is a collection of 16,576 individual grid-cell models. The $R^2$ value for the majority of the grid cells is greater than 0.95. In other words, for the vast majority of the domain, over 95% of the variation in the PM concentrations is explained with the linear regression model. The worst-performing portion of the domain is the north-west corner.
This is most likely due to the boundary effects close to the edge of the modeling domain. However, due to both the low variations in PM concentrations and the low populations in those grid cells, the relatively poor performance of the RSM in those grid cells has little impact on the aggregate $RE_t^d$ error metric.

4.2 Health Impacts of Current Levels of Aviation

In this section, the PM RSM is used to compute the health impacts of current aviation levels. Again, as stated in Section 2.1, the health impact calculations in the PM RSM do not follow EPA practice. The PM concentration data upon which the RSM is built is not SMATed and does not include particle-bound water. With this caveat in mind, the RSM estimates of the health impacts are also compared to those from the previous iF method.

4.2.1 Monte Carlo Comparison of RSM to iF

Both the PM RSM and the iF method are employed using two aviation emissions inventories computed using EDMS. The first inventory is the RSM baseline inventory, while the second inventory corrects for an error in the nonvolatile primary PM emissions calculation related to the engine bypass ratios. The annual total emissions are shown in Table 4.2. Tables 4.3 and 4.4 compare the RSM and iF estimates of incidences and valuations of health impacts due to aviation for the pre-BPR correction inventory, while Tables 4.5 and 4.6 repeat the comparisons for the post-BPR correction inventory. All four tables are restricted to two significant digits, due to the inherent uncertainty on the results.

Table 4.2: Pre- and post-BPR correction emissions inventories from CSSI used to compare the RSM to the iF model.

<table>
<thead>
<tr>
<th>[metric tons]</th>
<th>nvPM</th>
<th>volPM-sulf</th>
<th>volPM-org</th>
<th>NOx</th>
<th>SOx</th>
<th>VOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>preBPR</td>
<td>850</td>
<td>70</td>
<td>136</td>
<td>78,900</td>
<td>10,400</td>
<td>32,400</td>
</tr>
<tr>
<td>postBPR</td>
<td>275</td>
<td>70</td>
<td>136</td>
<td>78,900</td>
<td>10,400</td>
<td>32,400</td>
</tr>
</tbody>
</table>

As can be seen in the tables, the estimates for the health impacts after the BPR correction are between 85% and 90% of the impacts estimates before the correction for the iF method, and between 80% and 85% for the RSM. Although the corrected inventory contains only about one third of the primary, non-volatile PM (nvPM) emissions contained in
the baseline inventory, Rojo [40] shows that primary PM emissions only account for about 15% of the health impacts of aviation. Therefore, a decrease in impacts of just over 10% between the corrected inventory and the baseline inventory is expected for the iF method. This difference is only slightly larger for the RSM estimated impacts, perhaps due to the methodological differences between the RSM and the iF method.

Table 4.3: Comparison between the iF and RSM estimated incidences of health impacts due to the pre-BPR correction (i.e. RSM baseline) emissions inventory.

<table>
<thead>
<tr>
<th>preBPR incidences</th>
<th>iF 5%</th>
<th>mean</th>
<th>95%</th>
<th>RSM 5%</th>
<th>mean</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature mortality -</td>
<td>310</td>
<td>560</td>
<td>910</td>
<td>170</td>
<td>260</td>
<td>370</td>
</tr>
<tr>
<td>Long-term exposure (adults age 30+)</td>
<td>2.1</td>
<td>3.7</td>
<td>5.8</td>
<td>1.1</td>
<td>1.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Premature mortality -</td>
<td>140</td>
<td>230</td>
<td>330</td>
<td>80</td>
<td>100</td>
<td>130</td>
</tr>
<tr>
<td>Long-term exposure (infants age &lt; 1 yr)</td>
<td>95</td>
<td>160</td>
<td>250</td>
<td>54</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>100</td>
<td>160</td>
<td>240</td>
<td>60</td>
<td>75</td>
<td>91</td>
</tr>
<tr>
<td>Hospital admissions - respiratory</td>
<td>190</td>
<td>320</td>
<td>480</td>
<td>110</td>
<td>150</td>
<td>190</td>
</tr>
<tr>
<td>Hospital admissions - cardiovascular</td>
<td>190,000</td>
<td>220,000</td>
<td>320,000</td>
<td>79,000</td>
<td>100,000</td>
<td>120,000</td>
</tr>
</tbody>
</table>

Table 4.4: Comparison between the iF and RSM estimated valuations of health impacts due to the pre-BPR correction (i.e. RSM baseline) emissions inventory.

<table>
<thead>
<tr>
<th>preBPR valuations [millions of US dollars, year 2000]</th>
<th>iF 5%</th>
<th>mean</th>
<th>95%</th>
<th>RSM 5%</th>
<th>mean</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature mortality -</td>
<td>810</td>
<td>3,100</td>
<td>6,200</td>
<td>410</td>
<td>1,400</td>
<td>2,600</td>
</tr>
<tr>
<td>Long-term exposure (adults age 30+)</td>
<td>5.6</td>
<td>20</td>
<td>40</td>
<td>3.1</td>
<td>9.5</td>
<td>17</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>10</td>
<td>80</td>
<td>260</td>
<td>5.6</td>
<td>38</td>
<td>130</td>
</tr>
<tr>
<td>Hospital admissions - respiratory</td>
<td>1.6</td>
<td>3.2</td>
<td>6.3</td>
<td>0.86</td>
<td>1.5</td>
<td>2.8</td>
</tr>
<tr>
<td>Hospital admissions - cardiovascular</td>
<td>1.9</td>
<td>3.8</td>
<td>7.4</td>
<td>1.1</td>
<td>1.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Emergency room visits for asthma</td>
<td>0.05</td>
<td>0.1</td>
<td>0.2</td>
<td>0.03</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>Minor Restricted Activity Days</td>
<td>4.2</td>
<td>11</td>
<td>20</td>
<td>2.2</td>
<td>5.1</td>
<td>8.4</td>
</tr>
</tbody>
</table>

More importantly, the RSM mean estimates are roughly 45% of the iF mean estimates. Due to rounding, the apparent lower bound of the 95% confidence intervals for the RSM is as high as 60% of the iF lower bound, and the upper bound of the intervals for the RSM is as low as 35% of the iF upper bound. Thus, the confidence intervals seem to be smaller for the RSM than for the iF method. However, this may not be a statistically significant result, since the apparent variation in the ratio of RSM to iF estimates is at least partly due to their restriction to two significant digits.

On the other hand, the reasons for the difference between the RSM and iF estimates (i.e. the mean value of the ratio itself), are two-fold. The first reason is that the iF method
was developed using surface level emissions that were not specific to aviation. Thus, the impacts attributed to aviation in the iF are larger than those estimated by the RSM since the emissions are assumed to all be at the ground level (instead of some emissions aloft as for the RSM) and since they impact a larger population (iFs are not constrained to airport emissions sources only). This likely causes the iF estimate of the health impacts of aviation to be high.

Table 4.5: Comparison between the iF [41] and RSM estimated incidences of health impacts due to the post-BPR correction emissions inventory.

<table>
<thead>
<tr>
<th>postBPR</th>
<th>iF</th>
<th>RSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature mortality -</td>
<td>260</td>
<td>140</td>
</tr>
<tr>
<td>Long-term exposure (adults age 30+)</td>
<td>1.8</td>
<td>0.94</td>
</tr>
<tr>
<td>Premature mortality -</td>
<td>1.8</td>
<td>0.94</td>
</tr>
<tr>
<td>Long-term exposure (infants age &lt; 1 yr)</td>
<td>5.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>120</td>
<td>69</td>
</tr>
<tr>
<td>Hospital admissions - respiratory</td>
<td>80</td>
<td>46</td>
</tr>
<tr>
<td>Hospital admissions - cardiovascular</td>
<td>90</td>
<td>52</td>
</tr>
<tr>
<td>Emergency room visits for asthma</td>
<td>160</td>
<td>93</td>
</tr>
<tr>
<td>Minor Restricted Activity Days</td>
<td>120,000</td>
<td>68,000</td>
</tr>
</tbody>
</table>

Table 4.6: Comparison between the iF [41] and RSM estimated valuations of health impacts due to the post-BPR correction emissions inventory.

<table>
<thead>
<tr>
<th>postBPR</th>
<th>iF</th>
<th>RSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature mortality -</td>
<td>1.8</td>
<td>0.94</td>
</tr>
<tr>
<td>Long-term exposure (adults age 30+)</td>
<td>4.8</td>
<td>2.5</td>
</tr>
<tr>
<td>Premature mortality -</td>
<td>1.8</td>
<td>2.5</td>
</tr>
<tr>
<td>Long-term exposure (infants age &lt; 1 yr)</td>
<td>36</td>
<td>7.7</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>9</td>
<td>4.6</td>
</tr>
<tr>
<td>Hospital admissions - respiratory</td>
<td>1.3</td>
<td>0.73</td>
</tr>
<tr>
<td>Hospital admissions - cardiovascular</td>
<td>1.7</td>
<td>0.98</td>
</tr>
<tr>
<td>Emergency room visits for asthma</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Minor Restricted Activity Days</td>
<td>3.6</td>
<td>1.8</td>
</tr>
</tbody>
</table>

The second reason is that the RSM emissions processing does not follow EPA practice. Although the effect of SMATing the data on the final health impacts is ambiguous, adding the particle-bound water mass to the current, dry PM masses used in the RSM would certainly increase the estimate of the health impacts of aviation. Thus, the RSM estimate is likely to be low.

Finally, if aviation emissions at cruise altitudes do indeed have a significant impact on health impacts as is suggested by some recent work [50][2], both the iF and the RSM estimates would be lower than the actual, total impacts of aviation.
4.2.2 RSM Point Estimate Comparison to CMAQ

Next, the RSM is used to compute the point estimates of changes in PM concentrations and subsequent changes in health risks due to the baseline (pre-BPR correction) emissions inventories. Although the post-BPR correction inventories provide a better estimate of actual aviation emissions, only the pre-BPR correction inventories were simulated in CMAQ. Thus, to allow for a direct comparison of the point estimates using the RSM vs. CMAQ, the RSM is used with the pre-BPR correction inventories.

The point estimates assume the emissions inventories are known with certainty (i.e. no uncertainty distributions on the inputs). As a result the spatial interpolation uncertainty becomes irrelevant as well, since the inventory multipliers—defined as the ratio of the input emissions to the baseline emissions—are equal to one at each airport and therefore across the entire domain, regardless of the $\alpha$ and $\beta$ parameters, as shown in Equation 3.4.

The RSM point estimate of change in adult premature mortality across the domain due to current aviation emissions is 215. This compares favorably to the point estimate of 216 using the CMAQ results directly, and agrees with the -0.5% RSM error for the baseline CMAQ simulation (run 999) shown in Figure 4-2. However, the RSM point estimate is 17% lower than the mean estimate from the probabilistic Monte Carlo run presented in Table 4.3. This larger difference is due to the upward mean bias in the input emissions uncertainty distributions, consistent with the previous LAQ model [40][41].

Figure 4-6 plots the changes in ambient PM concentrations across the domain attributable to aviation activity in the United States. These changes in concentration combined with the population data presented in Figure 3-4 yield the increase in population exposure to PM due to aviation shown in Figure 4-8. Figure 4-8 shows that the majority of the impacts are localized in a small number of grid cells, with the largest contribution in Los Angeles County. This finding also agrees with previous air quality modeling results [40][46]. Note that Figure 4-8 is not specific to any health endpoint; rather the sum of the increase in exposure across each grid cell in the domain (i.e. the sum of the values in each grid cell in Fig. 4-8(a)), multiplied by the background incidence rate and the CRF for a specific health endpoint, yields the change in incidences of that endpoint across the domain, as shown previously in Equation 4.9.
Change in PM Concentration due to Aviation Emissions [micrograms per cubic meter]

Min = -1.209696e-003 @ (31,52)
Max = 5.990398e-002 @ (22,46)
Mean = 1.147391e-003

Figure 4-6: Change in PM concentration [μg/m³] due aviation emissions in each grid cell in the domain.

Log_{10} Change in PM Concentration due to Aviation Emissions

Min = -3.626291e+000 @ (1,16)
Max = -8.432246e+000 @ (65,108)
Mean = -3.455515e+000

Figure 4-7: Log_{10} of change in PM concentration [μg/m³] due aviation emissions in each grid cell in the domain.
Population x Change in Concentration

Min = -7.363018e+001 @ (31,52) Max = 2.050433e+005 @ (23,46) Mean = 1.463071e+002

Figures 4-8: Product of population and change in PM concentration across the domain. This population exposure plot is the cell-by-cell product of the plots in Figures 3-4 and 4-6.

Log₁₀(Population x Change in Concentration)

Min = 2.976222e-004 @ (41.67) Max = -Inf @ (1,1) Mean = -Inf

Figures 4-9: Log₁₀ of population times change in PM concentration. This population exposure plot is the cell-by-cell sum of the plots in Figures 3-5 and 4-7.
4.2.3 Impact Apportionment to Input Emissions

Finally, the RSM formulation lends itself to a straightforward analysis of the relative contributions of each input emissions species to PM formation. Since the RSM is a linear regression model, each of its \( \beta \) parameters represents the sensitivity of the PM concentration to one of the input emissions multipliers. The four \( \beta \) parameters (for the fuel burn, \( \text{SO}_X \), \( \text{NO}_X \), \( \text{nvPM} \) input multipliers) are depicted across the domain in Appendix D. As previously described in Equations 2.1 - 2.6, the fuel burn multiplier actually represents changes in \( \text{volPM-org} \) and \( \text{VOC} \) emissions, while the \( \text{SO}_X \) multiplier affects both \( \text{volPM-sulf} \) and \( \text{SO}_2 \) emissions.

As can be seen in Figures D-1, D-4, and D-5 as compared to Figures D-2 and D-3, the effects of \( \text{volPM-org} \), \( \text{VOC} \) and \( \text{nvPM} \) emissions are more localized than those of \( \text{SO}_X \), \( \text{volPM-sulf} \), and \( \text{NO}_X \) emissions. Although \( \text{SO}_X \), \( \text{volPM-sulf} \), and \( \text{NO}_X \) emissions have a greater impact on the average change in PM\(_{2.5}\) concentration across the domain, this does not translate directly into their dominance of health impacts (Tbl. 4.7), precisely because of their increased dispersion into grid cells with lower populations.

The RSM is used again to provide point estimates of the impacts of the four emissions multipliers. This time, the corrected (post-BPR) inventories are used since they provide a more accurate estimate of current aviation emissions. This is no longer a comparison to CMAQ results, but rather a best-estimate of the relative impacts of each emission type on public health. The point estimate for adult premature mortality is 180 and the estimate of the total health costs is $1.025 billion. These totals are attributed to the RSM input multipliers in Table 4.7. As mentioned previously, these point estimates are lower than the mean estimates from the Monte Carlo simulations shown in Tables 4.5 and 4.6 because of the upward bias in the emissions uncertainty distributions, but the proportional contributions of each of the four inputs in Table 4.7 still hold due to the linearity of the RSM.

Table 4.7: Apportionment of impacts to the fuel burn, \( \text{SO}_X \), \( \text{NO}_X \), and \( \text{nvPM} \) emissions from aviation. Note that the fuel burn stands in for \( \text{volPM-org} \) and \( \text{VOC} \) emissions, while the \( \text{SO}_X \) column represents simultaneous changes in \( \text{SO}_2 \) and \( \text{volPM-sulf} \) emissions.

<table>
<thead>
<tr>
<th></th>
<th>fuel burn</th>
<th>( \text{SO}_X )</th>
<th>( \text{NO}_X )</th>
<th>( \text{nvPM} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean change in PM conc [( \mu g/m^3 )]</td>
<td>1.78 ( \times 10^{-2} )</td>
<td>5.11 ( \times 10^{-2} )</td>
<td>3.38 ( \times 10^{-2} )</td>
<td>4.34 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>Adult premature mortality</td>
<td>54</td>
<td>54</td>
<td>50</td>
<td>21</td>
</tr>
<tr>
<td>Total health costs [$, year-2000]</td>
<td>$310 mil</td>
<td>$310 mil</td>
<td>$288 mil</td>
<td>$117 mil</td>
</tr>
<tr>
<td>Proportional population exposure</td>
<td>30%</td>
<td>30%</td>
<td>28%</td>
<td>11%</td>
</tr>
</tbody>
</table>
Chapter 5

Conclusions

This chapter concludes the thesis by briefly reviewing the approach, implementation, and results of the PM RSM effort, and providing guidance for future development. Section 5.1 summarizes the results of this thesis, while Section 5.2 suggests areas of future research regarding the air quality impacts of aviation.

5.1 Summary

This thesis presented the response surface model methodology, and developed the rationale for the choices reflected in the current PM RSM. The variables of interest were selected, their ranges set to define the RSM design space, and a low-discrepancy sequence sampling technique was chosen to allow for the flexibility of including additional CMAQ simulations. Once the CMAQ simulations were completed, three model formulations were implemented and compared: one kriging model with only a constant underlying assumption, another kriging model with a linear underlying surface, and a linear regression model. An error metric appropriate for the specific intended use of the RSM was developed. A cross-validation analysis was used to select the final formulation, and an error convergence analysis was performed to investigate the effect of additional CMAQ simulations.

The linear regression model implemented showed a lower prediction error than both of the kriging models. Although the kriging models return the CMAQ results when the same inputs are supplied while the regression model does not, the cross-validation analysis showed that the kriging models over-fit the data as compared to the regression model. This is probably because the 27 CMAQ simulations may not provide enough coverage of the
4-dimensional RSM design space. Thus the higher prediction errors of the kriging models may be due to the requirement of those model to pass exactly through the points in the training set, which compromises their performance away from those points. While the linear regression model does not pass through the training data points, it exhibits lower prediction error for new regions of the design space. The regression model is also more attractive due to its simplicity and shorter runtimes as compared to the kriging model, so it is selected as the formulation for the PM RSM. The prediction error is also shown to be reduced with the addition of new CMAQ data points, with a relative prediction error of the final model of roughly 1%.

Finally, the linear regression PM RSM is used to estimate the effects of current aviation levels in the United States. The model predicts 210 premature mortalities due to the changes in ambient PM concentrations caused by aviation emissions below 3,000 ft., with a 95% confidence interval between 140 and 290. The total cost of the health impacts of aviation is estimated at roughly $1.21 billion, with a 95% confidence interval between $370 million and $2.15 billion (year 2000 US dollars). Of these total impacts, 30% are found to stem from emissions of VOCs and volatile PM from organics, another 30% from emissions of SO$_2$ and volatile PM from sulfur, 28% from NO$_x$ emissions, and about 11% from non-volatile PM emissions.

The estimates of the total impacts are just under half of the estimates provided by the intake fraction method [41]. Due to the omission of particle-bound water in this version of the RSM, the actual impacts of aviation LTO emissions are likely higher than currently estimated by the model. In addition, the inclusion of aviation cruise emissions may further increase the estimates of the health impacts of aviation in the United States.

### 5.2 Future Work

As discussed in Section 3.5, there are a couple of major limitations in the applicability of the current version of the RSM to health impact analyses. Additionally, there are a couple of extensions of the current model that would provide a more accurate estimate of the true impacts of aviation emissions. As such, the following areas of future research are suggested.
5.2.1 Current CMAQ Data

The first few areas for future work all involve only the current model with the current CMAQ simulations. The first and most important steps are to SMAT the simulated data and to add the particle-bound water mass. This would align the RSM methodology with EPA practice. Although the PM SMATing software was not yet available for the PM RSM, and the proper water apportionment method has not yet been identified, both of these steps are necessary for a future version of the PM RSM.

Next, the spatial interpolation step required before running the RSM is an additional source of uncertainty not present in the intake fraction model. Therefore, the effect of the uncertainty on the spatial interpolation parameters on the RSM outputs needs to be investigated. As the spatial interpolation is the most time-consuming step of an RSM run, if it turns out the interpolation uncertainty is insignificant compared to the inventory or CRF uncertainties, then the spatial interpolation may be streamlined to improve RSM performance.

Another specific area that needs to be addressed is the treatment of the regression uncertainty when running a Monte Carlo simulation with the RSM. The relationship between the input emissions and spatial interpolation uncertainties and the regression uncertainty needs to be well defined, either to justify the current practice of treating the uncertainties as additive, or to determine how the uncertainties should interact in a future version of the RSM.

Finally, although the current RSM runtimes are much shorter than CMAQ runtimes, running the RSM in Monte Carlo mode can take several hours. Therefore, another area of improvement for the current RSM is the treatment of input and parameter uncertainty. One possible approach is to work with the cumulants of a probability distribution rather than with individual random draws from that distribution [70]. If the model formulation is simple enough, the method of combined cumulants may be substantially faster than Monte Carlo simulation.

5.2.2 Future Air Quality Simulations

Looking beyond the current CMAQ data set, there several long-term improvements that can be incorporated into the PM RSM. The first is widening the design space to better
accommodate the differences between emissions models. Given the rather large uncertainty on the emissions inputs to the PM RSM, as well as the differences between the EDMS and AEDT inventories, it seems prudent to widen the ranges on the RSM input variables accordingly. Second, with the inclusion of further CMAQ simulations, a fifth input variable could be added to the RSM. Namely, including the inventory HC EI (Eq. 2.3 and 2.6) will complete the RSM’s capability of characterizing changes in the emissions of primary PM and secondary PM precursors. Thus its inclusion will allow a new version of the RSM to be sensitive to changes in emissions it currently cannot detect. A third improvement that can be accomplished with the addition of a significant number of new CMAQ simulations is the inclusion of some spatial variation in the emissions multipliers for the CMAQ runs. This would provide some insight on the spatial effects of differential changes in emissions, which could make a new version of the RSM more applicable to regional policy analyses.

Next, and perhaps most importantly, the effect of cruise emissions on surface concentrations of PM needs to be fully investigated and characterized. If the impact of cruise emissions is significant as compared to the impact of LTO emissions, a future version of the RSM must certainly include cruise emissions as well. Several of the assumptions of the current PM RSM will need to be revisited with the inclusion of cruise emissions. The spatial interpolation scheme may need to be revised or replaced as the spatial distribution of cruise emissions needs to be characterized. The linearity of the response of PM concentrations to aviation emissions may also no longer hold, so the model formulation will need to be reconsidered. However, the inclusion of cruise emissions may turn out to be critical to fully account for all the health impacts of aviation.

Finally, a long term goal is to develop RSMs for other parts of the world. As part of the APMT software suite, the air quality analysis should be reproduced for Europe and Asia in order to capture the majority of global LTO emissions. This effort will involve identifying appropriate emissions and air quality models for the regions of interest, or calibrating the EDMS/CMAQ combination for use in these regions. Aviation and background emissions, initial and boundary concentration levels, as well as meteorological data will all be required for these parts of the world in order to replicate the PM RSM effort for the United States.
Appendix A

Design of Experiments

Table A.1: Combinations of design space input multipliers. Points RSM002-RSM025 are determined using a 4-dimensional Halton sequence. These design space multipliers are then combined according to Equations 2.1 through 2.6 to give the CMAQ input emissions multipliers at each airport.

<table>
<thead>
<tr>
<th>Name</th>
<th>Fuel Burn</th>
<th>Fuel Sulfur Content</th>
<th>NOx Inventory EI</th>
<th>EI nvPM Inventory EI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM000</td>
<td>0</td>
<td>0.025</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>RSM999</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RSM001</td>
<td>2.5</td>
<td>5</td>
<td>1.1</td>
<td>3.6</td>
</tr>
<tr>
<td>RSM002</td>
<td>0.5</td>
<td>0.735714286</td>
<td>0.833333333</td>
<td>1.925</td>
</tr>
<tr>
<td>RSM003</td>
<td>1</td>
<td>1.446428571</td>
<td>0.966666667</td>
<td>1.0875</td>
</tr>
<tr>
<td>RSM004</td>
<td>1.5</td>
<td>2.157142857</td>
<td>0.744444444</td>
<td>2.7625</td>
</tr>
<tr>
<td>RSM005</td>
<td>2</td>
<td>2.86757143</td>
<td>0.877777778</td>
<td>0.66875</td>
</tr>
<tr>
<td>RSM006</td>
<td>0.1</td>
<td>3.578571429</td>
<td>1.011111111</td>
<td>2.34375</td>
</tr>
<tr>
<td>RSM007</td>
<td>0.6</td>
<td>4.289285714</td>
<td>0.788888889</td>
<td>1.50625</td>
</tr>
<tr>
<td>RSM008</td>
<td>1.1</td>
<td>0.126530612</td>
<td>0.922222222</td>
<td>3.18125</td>
</tr>
<tr>
<td>RSM009</td>
<td>1.6</td>
<td>0.837244898</td>
<td>1.055555556</td>
<td>0.459375</td>
</tr>
<tr>
<td>RSM010</td>
<td>2.1</td>
<td>1.547959184</td>
<td>0.714814815</td>
<td>2.134375</td>
</tr>
<tr>
<td>RSM011</td>
<td>0.2</td>
<td>2.258673469</td>
<td>0.848148148</td>
<td>1.296875</td>
</tr>
<tr>
<td>RSM012</td>
<td>0.7</td>
<td>2.969387755</td>
<td>0.981481481</td>
<td>2.971875</td>
</tr>
<tr>
<td>RSM013</td>
<td>1.2</td>
<td>3.680102041</td>
<td>0.759259259</td>
<td>0.878125</td>
</tr>
<tr>
<td>RSM014</td>
<td>1.7</td>
<td>4.390816327</td>
<td>0.892592593</td>
<td>2.553125</td>
</tr>
<tr>
<td>RSM015</td>
<td>2.2</td>
<td>0.228061224</td>
<td>1.025925926</td>
<td>1.715625</td>
</tr>
<tr>
<td>RSM016</td>
<td>0.3</td>
<td>0.93877551</td>
<td>0.803703704</td>
<td>3.390625</td>
</tr>
<tr>
<td>RSM017</td>
<td>0.8</td>
<td>1.649489796</td>
<td>0.937037037</td>
<td>0.3546875</td>
</tr>
<tr>
<td>RSM018</td>
<td>1.3</td>
<td>2.360204082</td>
<td>1.070370370</td>
<td>2.0296875</td>
</tr>
<tr>
<td>RSM019</td>
<td>1.8</td>
<td>3.070918367</td>
<td>0.72962963</td>
<td>1.1921875</td>
</tr>
<tr>
<td>RSM020</td>
<td>2.3</td>
<td>3.781632653</td>
<td>0.862962963</td>
<td>2.8671875</td>
</tr>
<tr>
<td>RSM021</td>
<td>0.4</td>
<td>4.492346939</td>
<td>0.996296296</td>
<td>0.7734375</td>
</tr>
<tr>
<td>RSM022</td>
<td>0.9</td>
<td>0.32951837</td>
<td>0.774074074</td>
<td>2.4484375</td>
</tr>
<tr>
<td>RSM023</td>
<td>1.4</td>
<td>1.040306122</td>
<td>0.907407407</td>
<td>1.6109375</td>
</tr>
<tr>
<td>RSM024</td>
<td>1.9</td>
<td>1.751020408</td>
<td>1.040740741</td>
<td>3.2859375</td>
</tr>
<tr>
<td>RSM025</td>
<td>2.4</td>
<td>2.461734694</td>
<td>0.818518519</td>
<td>0.5640625</td>
</tr>
</tbody>
</table>
Appendix B

Spatial Interpolation

Figure B-1: Parametric spatial interpolation (Sec. 3.3) for fuel burn multiplicative factors at the airports in the CMAQ domain for $\alpha = \{0, 1, 2\}$ and $\beta = \{1, 4, 16\}$. 
### Appendix C

#### Airport Mapping to Representative Profiles

<table>
<thead>
<tr>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
<th>Apt</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
<td>ABE</td>
<td>PVD</td>
</tr>
<tr>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
<td>ABH</td>
<td>PVD</td>
</tr>
<tr>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
<td>ABQ</td>
<td>PVD</td>
</tr>
<tr>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
<td>ABY</td>
<td>PVD</td>
</tr>
<tr>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
<td>ACK</td>
<td>PVD</td>
</tr>
<tr>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
<td>ACY</td>
<td>PVD</td>
</tr>
<tr>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
<td>ADC</td>
<td>PVD</td>
</tr>
<tr>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
<td>ADD</td>
<td>PVD</td>
</tr>
<tr>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
<td>AEX</td>
<td>PVD</td>
</tr>
<tr>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
<td>AFW</td>
<td>PVD</td>
</tr>
<tr>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
<td>AGC</td>
<td>PVD</td>
</tr>
<tr>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
<td>AGS</td>
<td>PVD</td>
</tr>
<tr>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
<td>ALS</td>
<td>PVD</td>
</tr>
<tr>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
<td>ANC</td>
<td>PVD</td>
</tr>
<tr>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
<td>AOC</td>
<td>PVD</td>
</tr>
<tr>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
<td>APA</td>
<td>PVD</td>
</tr>
<tr>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
<td>APC</td>
<td>PVD</td>
</tr>
<tr>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
<td>APF</td>
<td>PVD</td>
</tr>
<tr>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
<td>ASE</td>
<td>PVD</td>
</tr>
<tr>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
<td>ATL</td>
<td>PVD</td>
</tr>
<tr>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
<td>ATW</td>
<td>PVD</td>
</tr>
<tr>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
<td>AUS</td>
<td>PVD</td>
</tr>
<tr>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
<td>AVL</td>
<td>PVD</td>
</tr>
<tr>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
<td>AVP</td>
<td>PVD</td>
</tr>
<tr>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
<td>AZO</td>
<td>PVD</td>
</tr>
<tr>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
<td>BCT</td>
<td>PVD</td>
</tr>
<tr>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
<td>BDL</td>
<td>PVD</td>
</tr>
<tr>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
<td>BDR</td>
<td>PVD</td>
</tr>
<tr>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
<td>BID</td>
<td>PVD</td>
</tr>
<tr>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
<td>BIF</td>
<td>PVD</td>
</tr>
<tr>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
<td>BFL</td>
<td>PVD</td>
</tr>
<tr>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
<td>BFM</td>
<td>PVD</td>
</tr>
<tr>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
<td>BGM</td>
<td>PVD</td>
</tr>
<tr>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
<td>BGR</td>
<td>PVD</td>
</tr>
<tr>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
<td>BHB</td>
<td>PVD</td>
</tr>
<tr>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
<td>BHM</td>
<td>PVD</td>
</tr>
<tr>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
<td>BIL</td>
<td>PVD</td>
</tr>
<tr>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
<td>BIS</td>
<td>PVD</td>
</tr>
<tr>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
<td>BIV</td>
<td>PVD</td>
</tr>
<tr>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
<td>BJC</td>
<td>PVD</td>
</tr>
<tr>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
<td>BKL</td>
<td>PVD</td>
</tr>
<tr>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
<td>BLG</td>
<td>PVD</td>
</tr>
<tr>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
<td>BLM</td>
<td>PVD</td>
</tr>
<tr>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
<td>BNS</td>
<td>PVD</td>
</tr>
<tr>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
<td>BOW</td>
<td>PVD</td>
</tr>
<tr>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
<td>BOS</td>
<td>PVD</td>
</tr>
<tr>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
<td>BPT</td>
<td>PVD</td>
</tr>
</tbody>
</table>

Table C.1: Mapping of 325 airports in domain to 3 representative profiles
Appendix D

Grid Beta Plots

Figures D-1 - D-4 on the next four pages depict the slopes of the PM$_{2.5}$ response surface along each of its input dimensions, for each grid cell in the domain. The fuel burn input multiplier represents changes in volPM-org and VOC emissions (if the other three inputs are held constant), as shown in Equations 2.1 - 2.6. Similarly, the SO$_X$ input multiplier represents changes in SO$_X$ and volPM-sulf emissions, while the NO$_X$ and nvPM multipliers represent changes in NO$_X$ and nvPM emissions directly. Since the RSM is a linear model, and since the input variables are multiples of current aviation emissions, these response surface slopes are also explicitly the contribution from each emission type to current PM$_{2.5}$ concentration. Finally, Figure D-5 presents the same plot for the post-BPR correction nvPM emissions. Note that the post-BPR correction nvPM emissions are roughly a factor of three lower than those used for the RSM baseline, so their corresponding impact on changes in PM$_{2.5}$ concentrations is also lower. The spatial interpolation used to create Figure D-5 had both the $\alpha$ and $\beta$ parameters set to 1.
Figure D-1: Response of PM concentrations to changes in volPM-org and VOC emissions across the domain.
Figure D-2: Response of PM concentrations to changes in SO\textsubscript{X} and volPM-sulf emissions across the domain.
Figure D-3: Response of PM concentrations to changes in NO$_X$ emissions across the domain.
Figure D-4: Response of PM concentrations to changes in nvPM emissions across the domain.
Figure D-5: Response of PM concentrations to changes in corrected nvPM emissions across the domain. Note the smaller scale, as well as minimum, maximum, and mean values on the top plot vs. those in Figure D-4.
Bibliography


[19] ICAO. Committee on Aviation Environmental Protection Sixth Meeting. Economic Analysis of NOx Emissions Stringency Options CAEP/6-IP/13, International Civil Aviation Organization, Committee on Aviation Environmental Protection, Forecasting and Economic Analysis Support Group, Montreal, Canada, February 2004.


