Evaluating the Accuracy of Pavement Deterioration Forecasts: Application to United States Air Force Airfields

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Benjamin R. Knost

Graduate Program in Civil Engineering

The Ohio State University

2016

Thesis Committee:

Dr. Rabi G. Mishalani, Advisor

Dr. Mark R. McCord

Dr. Abdollah Shafieezadeh
Abstract

The US Department of Defense is responsible for the design, construction, operation and maintenance of a vast network of pavement infrastructure, including parking, roadways, bridges, and airfields. This study focuses on airfields maintained by the United States Air Force. With budgets shrinking in recent years, it has become increasingly important to maintain airfields at serviceable levels for the lowest possible cost. Air Force civil engineers use a Pavement Management System (PMS) software known as PAVER to support the maintenance decision-making process they undertake on a regular basis. Airfields are represented by a set of contiguous sections expected to exhibit homogeneous condition. The inputs to PAVER are field observations of pavement distress for each section and the outputs are the values of the Pavement Condition Index (PCI) for each section. The PCI values are then used to model and predict deterioration, determine maintenance and repair requirements, and estimate future budgets. It is important for the PCI forecasts to be accurate to ensure that decision makers are making effective infrastructure investment decisions.

The objective of this thesis is to investigate the PCI prediction errors. Historical airfield PCI observations and forecasts at six Air Force installations across the United States are used to compare forecasted PCI values with observed PCI values. The errors in these historical forecasts are then used to develop a forecasting error model, which can be used to correct for systematic errors in the forecasts. Factors such as forecast horizon,
pavement age, condition, climate, and location are considered in developing and estimating the error model using ordinary least squares. Alternative models specifications are tested. The estimated models are also evaluated in terms of their effectiveness in correcting for systematic forecasting errors.

The results reveal that all the previously mentioned factors are statistically significant contributors to forecasting errors. To evaluate the developed relative error models, they are applied to correct the forecast errors of pavement sections in a test set (the records of the test set are not used in estimating the models). The results indicate that forecast error corrections are in general meaningful. The corrections lead to more improvements in the cases of longer forecast horizons, which is particularly encouraging. The implications of the results, directions for future research, and recommendations are discussed.
Acknowledgments

I would like to express my gratitude to my advisor, Dr. Rabi Mishalani for his outstanding insight and advice throughout this project, as well as my committee members, Dr. Mark McCord and Dr. Abdollah Shafieezadeh for their willingness to share their expertise. I would also like to thank Dr. Marilyn Croach for her encouragement in pursuing this degree and her continued support and mentorship through the process, as well as the outstanding professionals at the Air Force Civil Engineer Center for their willingness to share the data and intellectual capital to make this project a success.
Vita

2004..............................................B.S. Civil Engineering, United States Air Force Academy

2006..................................................M.S. Engineering Management, Air Force Institute of Technology

Fields of Study

Major Field: Civil Engineering
Table of Contents

Abstract ........................................................................................................ii

Acknowledgments ......................................................................................iv

Vita ..............................................................................................................v

List of Tables ..............................................................................................viii

List of Figures ...........................................................................................x

Chapter 1: Introduction .............................................................................1
  1.1 Background ..........................................................................................1
  1.2 Motivation ............................................................................................2
  1.3 Scope of Study .....................................................................................6
  1.4 Thesis Organization ............................................................................7

Chapter 2: Data ..........................................................................................9
  2.1 Data Sources .......................................................................................9
    2.1.1 Determining Pavement Condition Index (PCI) .........................11
    2.1.2 Deterioration Forecasting .............................................................14
    2.1.3 Climate Zones .............................................................................16
  2.2 Data Extraction and Integration .........................................................17
    2.2.1 Scope ............................................................................................17
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2 Transcription</td>
<td>19</td>
</tr>
<tr>
<td>2.3 Preliminary Processing</td>
<td>23</td>
</tr>
<tr>
<td>Chapter 3: Model Specification</td>
<td>27</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>27</td>
</tr>
<tr>
<td>3.2 Dependent and Explanatory Variables</td>
<td>27</td>
</tr>
<tr>
<td>3.2.1 Dependent Variables</td>
<td>27</td>
</tr>
<tr>
<td>3.2.2 Explanatory Variables</td>
<td>29</td>
</tr>
<tr>
<td>3.3 Exploratory Analysis and Possible Variable Specifications</td>
<td>34</td>
</tr>
<tr>
<td>3.4 Model Specification Alternatives</td>
<td>43</td>
</tr>
<tr>
<td>Chapter 4: Estimation Results and Model Evaluation</td>
<td>46</td>
</tr>
<tr>
<td>4.1 Summary Statistics of Variables and Outlier Analysis</td>
<td>46</td>
</tr>
<tr>
<td>4.2 General Estimation and Evaluation Approaches</td>
<td>49</td>
</tr>
<tr>
<td>4.3 Model Estimation Results</td>
<td>51</td>
</tr>
<tr>
<td>4.4 Interpretations of Select Models</td>
<td>58</td>
</tr>
<tr>
<td>4.4.1 Interpretation of Piecewise Linear Model</td>
<td>61</td>
</tr>
<tr>
<td>4.4.2 Interpretation of Natural Log Model</td>
<td>68</td>
</tr>
<tr>
<td>4.5 Model Evaluation</td>
<td>69</td>
</tr>
<tr>
<td>4.6 Summary of Estimation Results</td>
<td>75</td>
</tr>
<tr>
<td>Chapter 5: Summary, Future Research, and Recommendations</td>
<td>77</td>
</tr>
<tr>
<td>5.1 Summary</td>
<td>77</td>
</tr>
<tr>
<td>5.2 Future Research and Practical Recommendations</td>
<td>83</td>
</tr>
<tr>
<td>References</td>
<td>86</td>
</tr>
</tbody>
</table>
List of Tables

Table 2.1: Summary of bases and reports selected for analysis ..........................19
Table 2.2: Example dataset entries for bases with three pavement evaluation reports…23
Table 2.3: Example data entries with removed entries stricken through .................26
Table 2.4: Data summary .......................................................................................26
Table 3.1: Correlation coefficients for select variables ........................................35
Table 3.2: Complete list of variables and their definitions ........................................42
Table 4.1: Descriptive statistics ................................................................................47
Table 4.2: Estimation results for initial model specification 1 - Categorical ..............52
Table 4.3: Estimation results for final model specification 1 - Categorical ...............52
Table 4.4: Estimation results for initial model specification 2 - Piecewise linear ......53
Table 4.5: Estimation results for final model specification 2 - Piecewise linear .........54
Table 4.6: Estimation results for initial model specification 3 - Square root ............54
Table 4.7: Estimation results for final model specification 3 - Square root ..............55
Table 4.8: Estimation results for initial model specification 4 - Natural log ............56
Table 4.9: Estimation results for final model specification 4 - Natural log ..............56
Table 4.10: Summary of model alternatives .............................................................57
Table 4.11: Summary of model results with absolute relative error as dependent variable (identical specifications to section 4.3) .................................................................59

Table 4.12: Summary of model results with absolute relative error as dependent variable (with only statistically significant parameters included) ........................................60

Table 4.13: Summary statistics for relative error of test sections .................................71

Table 4.14: Percentages of sections improved by corrected forecasts based on percentile ranges of uncorrected relative error .......................................................74

Table 4.15: Percentages of sections improved by corrected forecast based on forecast horizon .........................................................................................................................75
List of Figures

Figure 1.1: Typical pavement deterioration curve showing corresponding routine maintenance and repair costs ................................................................. 3

Figure 2.1: Example division of rigid and flexible pavement features into sample units. 12

Figure 2.2: Example PAVER deterioration model output ........................................ 15

Figure 2.3: Climate zone map for the United States ........................................ 17

Figure 2.4: Sample section identifiers for taxiway in 2004 and 2008 .................. 21

Figure 3.1: Pairwise scatter plot of select variables ........................................... 36

Figure 3.2: Scatter plot of PCI$_{i0}$ versus Relative Error .................................. 37

Figure 3.3: General form of the categorical function ........................................ 39

Figure 3.4: General form of the piecewise linear function ................................. 41

Figure 4.1: Empirical cumulative distribution function for dependent variable Rel_Error .......................................................................................... 48

Figure 4.2: Empirical cumulative distribution function for dependent variable Rel_Error with upper and lower outlier boundaries shown ........................................ 48

Figure 4.3: Values of relative error for varying values of PCI$_{i0}$ in piecewise linear specification for a reference pavement section ........................................ 66

Figure 4.4: Values of relative error for varying values of PCI$_{i0}$ in both piecewise linear and natural log model specifications ........................................ 69
Figure 4.5: Empirical distribution functions and sample mean values of relative error of the original and corrected forecasts based on the piecewise linear model ………………72
Figure 4.6: Empirical distribution functions and sample mean values of relative error of the original and corrected forecasts based on the natural log model ……………………73
Chapter 1: Introduction

1.1 Background

The Department of Defense is responsible for constructing, operating and maintaining a vast network of pavement infrastructure, including parking, roadways, bridges, and airfields. Maintaining these pavement systems is especially critical in the case of United States Air Force airfields. Since critical aircraft missions are taking off and landing on these airfield pavements every day, it is important for Air Force civil engineers to assess the current and projected future conditions of their pavement systems. With budgets shrinking in recent years, it has become increasingly important to maintain these infrastructure systems at the lowest possible cost.

Pavement Management Systems (PMS) are commonly used by infrastructure agencies and organizations to support the monitoring and decision making functions they oversee and conduct. The use of PMSs has been shown to greatly increase the efficiency and cost effectiveness of management efforts and thus reduce the overall life cycle cost of pavements by allowing decision makers to apply the right actions at the right time (Hadi et al., 2010; Vasquez, 2011; Moazami et al., 2010). The Department of Defense, including Air Force civil engineers, use a PMS developed in the late 1970s by the United States Army Corps of Engineers known as PAVER (Colorado State University, 2015). PAVER uses inputs from field observations of pavement distresses and calculates a
Pavement Condition Index (PCI) for each pavement section observed. These PCI values can then be used to model future deterioration and, thus, predict maintenance and repair requirements to estimate future budget requirements (US Army Corps of Engineers, 2015). Since maintenance and repair activities and budgeting decisions are made based on the PAVER section PCI forecasts, it is important for these forecasts to be accurate in order to ensure that effective decisions are being made.

The purpose of this study is to use historical airfield PCI observations and forecasts at select Air Force installations across the United States to compare forecasted PCI values with observed PCI values. The errors in these forecasts are quantified and, in turn, modeled as a function of variables representing condition at the time of the forecast, design characteristics, environment, and location. Once model specifications and estimations are carried out, the resulting models are used to correct existing forecasts to evaluate the quality of the models.

1.2 Motivation

Pavement deterioration is a topic that has been studied at length. Studies have shown that pavement sections deteriorate over time. Figure 1.1 shows a typical deterioration function. As pavement age increases, its condition worsens, with decreases in condition at a higher rate seen towards the end of the pavements life. It has also been shown that the cost to maintain and repair these pavements is inversely related to the condition. When pavement condition decreases (i.e., worsens), the cost of maintenance increases. And, as the condition worsens at an increasing rate, the maintenance cost rises at an increasing rate (Al-Mansour and Sinha, 1994).
Based on the way in which deterioration manifests in pavements and the economic consequences that result, understanding, modeling, and forecasting this behavior is critical to the efficient maintenance and operations of pavement systems. Researchers have attempted to model pavement deterioration with mixed results (Ramaswamy and Ben-Akiva, 1990). Deterioration is a complex phenomenon with many contributing factors. Most pavement deterioration models consider some representation of condition as the dependent variable such as rutting, cracking, or a generalized condition index as a function of a combination of explanatory variables known to contribute to the deterioration in the form of worsening condition. Observations of the dependent and explanatory variables are used to estimate the parameters of a deterioration model. The success of the model is normally determined by the overall fit.
and the statistical significance of the parameter estimates corresponding to the explanatory variables along with a subjective visual examination of the estimated function as it compares to the sample data—see Hajek et al. (1985) for a comparison of several examples of pavement deterioration models.

In what follows, example deterioration models are presented to illustrate the types of models and issues that have been considered. The original deterioration model adapted for use in PAVER was developed by Shahin et al. (1987). The developers attempted to capture the effects of several explanatory variables by grouping like pavement sections into “families”. Families consist of pavement sections with the same pavement type, traffic use, and loading, thus, they are expected to deteriorate in a similar manner. Within each deterioration family, section PCI is modeled as a function of pavement age only, where a constrained least-squares curve-fitting technique is used to fit a line or curve through available data (Shahin et al., 1987). As a result, each family of pavements at each location, has a separate deterioration model. These models are in turn used to forecast PCI for a pavement section based only on the age and the family the pavement section belongs to.

More recently, studies focused on the deterioration of pavement on Air Force airfields sought to determine what characteristics, if any, contributed to pavement sections deteriorating faster or slower than other similar sections—see Meihaus (2013) and Sahagun (2014). The results suggest that environmental variables, such as rainfall and exposure to the freeze-thaw process, impact the rate of deterioration and that certain climate zones can be established such that pavement sections located within a given zone
experience the same environmental conditions. Overall pavement deterioration as well as the presence of specific types of distresses were shown to vary depending on the climate zone pavement sections are located in.

While Ben-Akiva and Ramaswamy (1990, 1993) also used an index that combines various damage types to represent condition, the developed model included environmental, traffic, loading, design, and routine maintenance and repair action variables into a comprehensive deterioration model. The model estimation also recognizes the presence of a simultaneous relationship between the routine maintenance and repair actions on the one hand and the condition index and other explanatory variables on the other, stemming from the use of data on in-service pavement sections.

In contrast, Archilla and Madanat (2000, 2001) specifically modeled deterioration as measured by rut depth (rather than an index) as a function of a comprehensive set of explanatory variables following a specification derived from mechanistic principles. The model was estimated using accelerated testing data in the absence of the effects of routine maintenance and repair. Similarly, Nakat and Madanat (2008) modeled a specific damage type, in this case cracking, as a function of several explanatory variables.

Even for deterioration modeling efforts that are successful in terms of statistical significance and fit, little has been done to test the ability of deterioration models to accurately forecast future condition. Atypically, in the case of Archilla and Madanat (2000), the established deterioration model was applied to several sample pavement sections that were not used in model estimation to evaluate and demonstrate its effectiveness. The results were examined visually to determine how well the developed
model predicted the observed data for these sections. However, the accuracy of the predictions was not quantified and only a small number of sections were considered.

1.3 Scope of Study

Air Force leaders use a prescribed decision making process that evaluates various routine maintenance and repair, rehabilitation and reconstruction alternatives against each other to determine where limited funds should be spent. PCI values forecasted by PAVER are the heaviest weighted components of the decision making process (Air Force Civil Engineer Center, 2015). The accuracy of these forecasts is critical to making informed and cost effective decisions. However, to date, there has not been any effort to evaluate the existing forecasting methodology or improve upon it.

The study reported in this thesis is concerned with the quantification, modeling, evaluation, and correction of systematic forecasting errors for Air Force airfield pavement sections located in the United States by analyzing observed and forecasted PCI values. The inventory of airfield pavement sections that is currently kept only allows for analyzing the current deterioration rates. While the existing central database does not contain historical data, the archive of historical pavement evaluation reports does. The Air Force Civil Engineer Center maintains copies of previous evaluation reports in portable document file (pdf) format for many Air Force bases around the world. While this archive is not complete, there are numerous reports available that can be used to sufficiently populate a dataset which contains observed PCI values and deterioration rates for families of pavement sections through time. With this information, a forecasted PCI for any pavement section can be calculated for the next time when that section was
observed as documented in a report. Comparing the forecasted to observed values allows for the quantification of forecasting error, which in turn can then be modeled as a function of available explanatory variables. The completed model for forecast error can then be used to correct future forecasts to account for the presence of systematic forecasting error and produce more accurate forecasts in future.

There are a total of 59 airfields in the continental United States. Data availability and analysis resources allowed for historical reports from six of these installations to be used to extract the data for this study. A portion of this dataset formed the sample set used to estimate forecasting error models, and the remaining portion formed the test set used to evaluate the models. While the models are estimated and evaluated using a dataset specific to these six bases, the model variables are common to all bases in the population. Therefore, the model specifications considered could be used to correct forecasts at any of the 59 bases in the population. Alternatively, if resources are available, similar specifications could be estimated and evaluated using larger datasets and then used to correct forecasts across bases and over time.

1.4 Thesis Organization

The rest of this thesis is organized as follows. Chapter 2 presents in detail the reports that were obtained, the process for extracting data from these reports to develop the dataset, and the preliminary processing of the dataset. Chapter 3 describes the exploratory data analysis, and the development of alternative model specifications. Chapter 4 presents the model estimation results, their interpretations, and the evaluation of their effectiveness in correcting PCI forecasts. Chapter 5 summarizes the study,
provides suggestions for future research, and makes recommendations of practical value for the Air Force in managing their pavement information resources.
Chapter 2: Data

2.1 Data Sources

The data used in this study were provided by the United States Air Force Civil Engineer Center’s (AFCEC) Airfield Pavement Evaluation (APE) team. This organization is responsible for the continual inspection and evaluation of Air Force operated airfields worldwide. Each airfield in the Air Force inventory is scheduled to have a pavement evaluation performed once every four years. Depending on the criticality of the airfield, availability of the inspectors and other mission related factors, the time between evaluations can fluctuate, but evaluations are never performed more frequently than two year intervals and rarely exceed six year intervals.

The APE team has qualified military personnel and necessary equipment to accomplish these inspections, but does not have the capacity to perform all the necessary ones to keep the Air Force inventory current. Due to this limitation, they often use government contractors to perform the routine evaluations within the continental United States and focus Air Force personnel and equipment resources on evaluations overseas whether in friendly territory or in combat zones. Due to evaluations being conducted by different contractors through the years, the evaluation reports vary slightly in format and
content, which presented challenges during the data extraction process. These challenges and subsequent mitigations are discussed in this chapter.

An evaluation consists of an extensive on-site physical inspection of an airfield. During the assessment, the type and severity of pavement distresses and deterioration are observed and recorded. The resulting evaluation report provides base pavements engineers and leadership with a thorough assessment of the current condition of their airfield pavements, projected future condition of pavements, as well as maintenance and repair recommendations. The report includes a summary of the type of pavements present, friction testing results for runway pavements, foreign object damage (FOD) risk assessment, and pavement condition index (PCI) results (Department of Defense, 2004).

Of the elements included in the report, PCI is the most critical. This condition index can be used as a “gradecard” for airfield pavement sections, and is the condition metric used by the Air Force in pavement condition forecasting as well as planning and programming for future maintenance and repair requirements. The observed PCI values and forecasted deterioration rates contained within the evaluation report comprises the source data used in this study.

While evaluations are conducted by a variety of inspectors, the basic evaluation process is the same regardless of whom is performing it. Standardized evaluation procedures are prescribed in Unified Facilities Criteria (UFC) (2004) and American Society of Testing and Materials (ASTM) (2012). The following sections outline the process that evaluators use for dividing the airfield into homogeneous sections for
evaluation, calculating a current PCI for each section, and then establishing a forecasted
deterioration rate for each section.

2.1.1 Determining Pavement Condition Index (PCI)

The following steps are followed to determine the PCI for each section of an
airfield.

1. Divide an airfield into homogeneous sections: Each airfield is defined as a network.
   For example, the airfield at Langley Air Force Base, VA is defined as an airfield network
   named “Langley”. Networks are divided into branches based on primary usage. Every
   airfield has four types of branches: Runways, taxiways, parking aprons, and overruns.
   Within these broad categories there may be multiple branches. For example, if an airfield
   consists of one runway with overruns, two taxiways and two parking aprons, the list of
   branches would likely be: Runway, overrun, taxiway A, taxiway B, East apron, and West
   apron.

   Each branch is then subdivided into sections based on traffic and design
   characteristics. For example, if a runway has portions that are asphalt and portions that
   are concrete, this would necessitate separate sections within the runway branch.
   Likewise, even when the same construction material is used, changes in cross section
   design properties or construction dates leads to the definition of a separate section. The
   airfield network is broken down in branches and then sections in this manner until there
   is a list of sections that are each homogeneous in terms of usage, construction material,
   design characteristics, and construction or rehabilitation date. That is, a designated
   pavement section has consistent pavement layer thickness and materials, was constructed
or rehabilitated at the same time, and is subject to approximately the same traffic loading (Department of Defense, 2004). While most branches have more than one section, there is no requirement for multiple sections within a branch. That is, it is possible to have a branch that is fully defined by one section.

2. Divide sections into sample units: Each section is now divided into sample units. For asphalt pavements, sample units are 5,000 square feet (+ or – 2,000 square feet). For concrete pavements, a sample unit is 20 slabs (+ or – 8 slabs). Figure 2.1 illustrates how sample units are defined for both rigid and flexible pavements.

![Example Division of a Jointed Rigid Pavement Feature Into Sample Units](image1)

![Example Division of Flexible Pavement Feature Into Sample Units](image2)

Figure 2.1: Example division of rigid and flexible pavement features into sample units (Source: ASTM, 2012)
3. **Determine sampling**: Since visual inspection of all sample units across all sections of an airfield is time and cost prohibitive, a statistical sampling technique is utilized to determine how many sample units from each section must be inspected to achieve a 95% confidence level that the PCI determined from the sampled units is equal to the actual PCI of the entire section.

4. **Record distresses**: For each sample unit selected for observation, the type, extent and severity of pavement distress is recorded. The types and severity of pavement distresses are defined and examples given in the regulations pertaining to airfield pavement evaluation (ASTM, 2012; Department of Defense, 2004).

5. **Calculate PCI**: The distresses measured and recorded in step 4 are then used to determine deduct values and these deduct values are subtracted from 100 to give a PCI value for a unit. PCI values range from 0 to 100, where 0 represents a completely failed pavement section and 100 represents a pavement section in perfect condition (ASTM, 2012). Once PCIs are calculated for all sample units, the section PCI value is calculated as the area weighted PCI of the randomly surveyed sample units as follows:

\[
\text{PCI}_s = \frac{\sum_{i=1}^{n} (PCI_i \times A_{ri})}{\sum_{i=1}^{n} A_{ri}}
\]

(2.1)

where:

- \(\text{PCI}_s\) (section PCI) = area weighted PCI of randomly surveyed sample units,
- \(\text{PCI}_i\) = PCI of randomly sampled unit \(i\),
- \(A_{ri}\) = area of randomly sampled unit \(i\),
- \(n\) = number of randomly sampled units surveyed.
This prescribed and repeatable inspection process results in an index that consistently quantifies pavement condition for each section. While PCI calculations can be accomplished manually, most inspectors utilize PMS software to calculate PCI based on their inputted distress observations.

### 2.1.2 Deterioration Forecasting

With current PCI established, forecasting is accomplished using an established process within the PAVER PMS as summarized in what follows.

1. **Define deterioration families:** Forecasting is accomplished by first grouping pavement sections into families. A family is defined as pavement sections that are of like construction material and loading characteristics and that are expected to perform in a similar manner over time (US Army Corps of Engineers, 2015). For example, runway pavements that have an asphalt surface and receive heavy traffic might be grouped into one family, whereas concrete apron pavements might comprise another family. The establishment of families is at the discretion of the inspector and will vary depending on site specific characteristics. For example, if there are no asphalt runway sections, that would not be listed as a family.

2. **Plot PCI versus age within a family:** Once all sections are grouped into families, the observed PCI for each section within that family is plotted against the age of the pavement section. Age is defined as the number of years since the last time the section was at a PCI of 100. If a pavement section was originally constructed in 2000, is now being evaluated in 2016, and it has not been reconstructed or rehabilitated (e.g., resurfaced) since 2000, its age would be 16 years. Likewise, if it was constructed in 2000
and received a rehabilitation or reconstruction in 2015 that brought its surface back to a PCI of 100, then the age at evaluation would be 1 year.

3. **Fit a trendline:** The PCI versus age plot is then used to establish the family deterioration rate. Filters are first applied to omit obvious errors and statistical outliers from the data. With outliers omitted, a constrained least-squares curve-fitting technique is used to fit a regression line to the filtered data (Shahin et al., 1987) where the PCI at age zero is constrained to take the value of 100. The fitted linear function establishes the expected deterioration rate (PCI/yr) for a given pavement family within a given network. An example family deterioration model produced by PAVER is shown in Figure 2.2 with the fitted deterioration model shown as the center line with the top and bottom lines showing the outlier limits.

![Figure 2.2: Example PAVER deterioration model output](image)

With the family deterioration rate determined, the predicted PCI for any pavement section within that family at future time, $t$, is determined as follows:
PCI_{if} = PCI_{i0} – ROD \times (\Delta t) \tag{2.2}

where,

PCI_{if} = Forecasted PCI of pavement section i at time t,

PCI_{i0} = PCI of pavement section i at time of forecast (commonly the latest available value),

ROD = rate of deterioration determined from fitted line, and

\Delta t = number of years between PCI_{i0} and PCI_{if}.

While the deterioration function produced by PAVER software can take a linear or polynomial form, the vast majority of reports utilize a simple linear form to model the degradation of PCI over time. The rare cases when polynomial functions are used occur when there are enough data points within a single family for the PAVER software to fit this type of function to the data. Normally, once PCI versus age is plotted within a family, a simple trend line best fits the available data. All of the evaluation reports that were used in creating the data set for this study are based on a linear deterioration model for forecasting.

2.1.3 Climate Zones

Previous research into the effect of climate on rates of pavement deterioration has been accomplished using Air Force airfield pavement sections as sample data (Meihaus, 2013). This research used historical weather data and geospatial interpolation techniques to establish four distinct climate zones within the United States, and found that there were statistically significant differences in pavement deterioration rates between some of the zones. The four zones were identified as Freeze-Wet, No Freeze-Wet, No Freeze-Dry
and Freeze-Dry based on typical weather patterns. The climate zones are shown in Figure 2.3. These climate zones are used as source data to generate climate related explanatory variables in subsequent sections.

Figure 2.3: Climate zone map for the United States (Source: Meihaus, 2013)

2.2 Data Extraction and Integration

The following sections describe the scope of the pavement evaluation reports used in this study, how data were transcribed from those source documents, and the method used to process the data to ensure validity for the intended purpose.

2.2.1 Scope

While airfield pavement evaluations have been taking place on a consistent basis in the Air Force for the past 15 to 20 years, not all evaluation reports and data have been kept and filed in a central location. The historical data are more complete for some bases than others just as the frequency of data collection is greater at some bases than others. In total, the Air Force manages 102 primary airfields worldwide, and all of the current
and historical pavement evaluation reports available for those airfields are maintained by the Air Force Civil Engineer Center at Tyndall Air Force Base, Florida.

The process of extracting data from reports and manually creating the dataset for this study proved to be time intensive. Therefore, a comprehensive data extraction from all available reports at all Air Force bases was deemed unrealistic. As a result, establishing the scope of the study early on was important. As discussed in chapter 1, the primary goal of this research is to evaluate the quality of the pavement condition forecasts produced by the Air Force. Since most bases are located within the continental United States and previous pavement deterioration research has been done on this sub-set of bases, the scope of this study is restricted to US bases.

Within the United States, previous research has shown that pavement deterioration on Air Force airfields is influenced by four primary climatic zones (Meihaus, 2013). Therefore, it is desirable to select bases within each of these separate climate zones in order to control for the effect of climate. Considering the available reports for bases within each climate zone in the continental US, bases that had at least two previous reports available for analysis and that had large enough airfields to contain a variety of features, construction materials, traffic, and usage are selected. Table 2.1 shows a list of these six bases organized by climate zone and providing the report years considered. Notice, two bases belong to each of the climate zones 1 and 2 and one base belongs to each of climate zones 3 and 4. Moreover, the reports span anywhere from 5 to 9 years across the six bases.
Table 2.1: Summary of bases and reports selected for analysis

<table>
<thead>
<tr>
<th>Climate Zone</th>
<th>Climate Zone 1: Freeze/Wet</th>
<th>Climate Zone 2: No Freeze/Wet</th>
<th>Climate Zone 3: No Freeze/Dry</th>
<th>Climate Zone 4: Freeze/Dry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>Langley AFB, VA</td>
<td>Wright-Patterson AFB, OH</td>
<td>Eglin AFB, FL</td>
<td>Nellis AFB, NV</td>
</tr>
<tr>
<td>Year 3 Report</td>
<td>2014</td>
<td>2013</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>


2.2.2 Transcription

The following process was followed for transcribing data from these reports to create the dataset used in the analysis.
1. Establish list of pavement sections from earliest report (year 1): For each base, information in the earliest report (year 1) was used to create a table in Microsoft Excel with the following variables for each section:

- Network: The pavement network, or base of the corresponding record.
- Branch: The branch name for the record.
- Section: Section identifier for the record.
- PCI_{i0}_Yr: The year in which the PCI evaluation was conducted (i.e., the year of the report).
- Age: Age of the pavement section at PCI_{i0}_Yr.
- PCI_{i0}: The observed PCI of pavement section i at time PCI_{i0}_Yr.
- Family: The deterioration family that the section was classified to belong to.
- ROD: The rate of deterioration (PCI points/year) determined for the corresponding family of the record.

2. Extract information from the subsequent (year 2) report: For each base, the subsequent (year 2) report was now used to add the following variables to the dataset:

- PCI_{ia}_Yr: The year in which the evaluation for report year 2 was conducted.
- PCI_{ia}: The observed PCI of section i at time PCI_{ia}_Yr.

There are a number of years between consecutive reports and thus the evaluators may or may not be the same from report to report. In addition, the report may be accomplished by a different contractor. The guidance for conducting the evaluations remained consistent, but there are some details that are at the evaluator’s discretion and can change between reports. As a result, the defined pavement sections on an airfield are not
necessarily identical between evaluations. Dividing the pavement network into branches and then sections is done at the beginning of each evaluation. Changes in the physical characteristics of the pavement and major construction, re-construction, or rehabilitation applied in the years between reports can change how the homogeneous sections are identified. The naming convention is also at the discretion of the inspector. While most use the previous section identifiers and only modify or change them when necessary, there are instances when the same physical section of pavement is referred to by different identifiers in subsequent reports.

To account for these inconsistencies, the section identifier map included in each report was examined to ensure that the physical pavement section that was observed and given a value for PCI$_{i0}$ in PCI$_{i0}$$_\text{Yr}$, corresponds to the same section for which PCI$_{ia}$ in PCI$_{ia}$$_\text{Yr}$ are recorded. For example, assume a taxiway at a given airfield was evaluated in 2004 using the sections shown in the top part of Figure 2.4. The same taxiway was then evaluated in 2008 with the sections shown in the bottom part of Figure 2.4.

<table>
<thead>
<tr>
<th>Section Identification for 2004 Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1A1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section Identification for 2008 Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1A</td>
</tr>
</tbody>
</table>

Figure 2.4: Sample section identifiers for taxiway in 2004 and 2008

The dataset would contain entries for sections T1A1 and T1A2, with the respective PCIs observed in 2004 recorded as PCI$_{i0}$. The PCI observed for section T1A in 2008 would be used as PCI$_{ia}$ for both the T1A1 and T1A2 entries. In cases where a pavement section
that existed in PCI0_Yr was demolished or re-constructed with different dimensions, construction materials etc., the entry for that section was removed from the dataset. This methodology was applied to each section so that the final dataset contained only observed PCI values that are paired, meaning the PCI index reported was referring to the same portion of pavement across the two reports.

3. **Extract data from next (year 3) report if available:** Bases where three reports are available offered several more data analysis opportunities. Not only could the observed and forecasted PCI values from the year 1 report be compared and analyzed against the current values from the year 2 report, the same analysis could be performed between the year 2 report and the year 3 report as well as between the year 1 report and the year 3 report. In these cases, the same process for populating the dataset is used as described in steps 1 and 2 by adding two additional records for each pavement section. The first (original) record, generated before applying step 3, uses the year 1 report to populate PCI0_Yr and PCI0 and the year 2 report to populate PCI1a_Yr and PCI1a. The second (additional) record, generated as part of step 3, uses the same values for PCI0_Yr and PCI0 (from the year 1 report), but uses values from the year 3 report to populate PCI1a_Yr and PCI1a. The third (additional) record, also generated as part of step 3, uses the year 2 report to populate PCI0_Yr and PCI0 and the year 3 to populate PCI1a_Yr and PCI1a.

An example table of records in the case where three reports are available for a base is shown in Table 2.2. The first record is for section A05B where the PCI values are available for 2005 and 2008. The second record is for section A05B where the PCI values are available for 2005 and 2014. And, the third record is for section A05B where
the PCI values are from 2008 and 2014. A similar treatment applies to section A06B.

Once data from all 15 reports were extracted, the dataset contained 1,219 records.

Table 2.2: Example dataset entries for bases with three pavement evaluation reports

<table>
<thead>
<tr>
<th>Network</th>
<th>Branch</th>
<th>Section</th>
<th>PCI$_{i0}$ Yr</th>
<th>Age</th>
<th>PCI$_{i0}$</th>
<th>Family</th>
<th>ROD</th>
<th>PCI$_{ia}$ Yr</th>
<th>PCI$_{ia}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nellis</td>
<td>APARKW</td>
<td>A05B</td>
<td>2005</td>
<td>10</td>
<td>86</td>
<td>Pri PCC Apron</td>
<td>1.06</td>
<td>2008</td>
<td>85</td>
</tr>
<tr>
<td>Nellis</td>
<td>APARKW</td>
<td>A05B</td>
<td>2005</td>
<td>10</td>
<td>86</td>
<td>Pri PCC Apron</td>
<td>1.06</td>
<td>2014</td>
<td>86</td>
</tr>
<tr>
<td>Nellis</td>
<td>APARKW</td>
<td>A05B</td>
<td>2008</td>
<td>13</td>
<td>85</td>
<td>Pri PCC Apron</td>
<td>0.98</td>
<td>2014</td>
<td>86</td>
</tr>
<tr>
<td>Nellis</td>
<td>APARKC</td>
<td>A06B</td>
<td>2005</td>
<td>10</td>
<td>93</td>
<td>Pri PCC Apron</td>
<td>1.06</td>
<td>2008</td>
<td>95</td>
</tr>
<tr>
<td>Nellis</td>
<td>APARKC</td>
<td>A06B</td>
<td>2005</td>
<td>10</td>
<td>93</td>
<td>Pri PCC Apron</td>
<td>1.06</td>
<td>2014</td>
<td>90</td>
</tr>
<tr>
<td>Nellis</td>
<td>APARKC</td>
<td>A06B</td>
<td>2008</td>
<td>13</td>
<td>95</td>
<td>Pri PCC Apron</td>
<td>0.98</td>
<td>2014</td>
<td>90</td>
</tr>
</tbody>
</table>

2.3 Preliminary Processing

The level of routine and preventative maintenance and repair performed on pavement sections has been shown to impact deterioration (Fwa and Sinha, 1986). Such maintenance actions are defined as work that slows deterioration of a pavement section and, as a result, extends the usable life. Work that serves to improve condition or restore it to a previous state is considered rehabilitation or reconstruction (Fwa and Sinha, 1986; Zaniewski and Mamlouk, 1996).

The data used in fitting a deterioration line correspond to sections where a level of routine and preventative maintenance and repair is being applied. Therefore, the effect of this maintenance effort is accounted for in the resulting deterioration model. While exact policies, procedures, and funding levels may vary between locations, using the deterioration already realized in other pavement sections to predict future condition means that the effect of such routine maintenance and repair efforts are captured in the
forecasts. Consequently, any pavement section that displays an increase in PCI over time should be assumed to have undergone rehabilitation or reconstruction. For this reason, assuming that the PCI values are error free, any dataset record where $PCI_i^0 - PCI_i^a < 0$ should be removed from the dataset for the purpose of this study because the section must have received more than routine maintenance, and therefore, the assumption for deterioration forecasting does not hold and, as a result, the forecast does not apply to that section.

However, the PCI values are subject to measurement and sampling errors and, therefore, removing all records with any increase in PCI over time might eliminate valid data. The sampling technique utilized produces results at a 95% confidence level and the PCI assessment is considered to be accurate within $\pm 5$ points (ASTM, 2012). That is, a pavement section may show an increase between reports of up to 5 PCI due to measurement and sampling errors, even if only routine maintenance was applied.

Therefore, only increases in PCI greater than 5 points were assumed to be due to the application of rehabilitation or reconstruction activities and are consequently removed from the dataset. In the case of bases with only two reports, if there is an increase between the year 1 report and the year 2 report, then the associated record is removed. For bases with 3 reports, if there is an increase in PCI greater than 5 between any two of the reports, the corresponding record is removed. In addition, if years 1 and 3 reports exhibit an increase $> 5$ but each of years 1 to 2 and years 2 to 3 reports exhibit an increase $< 5$, all three records would be removed (however, no such cases were reflected in the dataset).
Table 2.3 shows a sample of data records extracted from three evaluation reports for Langley Air Force Base, Virginia. This example relates to three pavement sections, evaluated in 2007, 2010, and 2014. The stricken entries are removed from the dataset due to an increase in PCI that can not be attributed to measurement and sampling errors. That is, $\text{PCI}_{i0} - \text{PCI}_{ia} < -5$. Section A10B showed an increase between the 2007 and 2010 report, but since it is within the margin of measurement error, the entry is kept. Section A18B exhibits increased deterioration (i.e., a reduced PCI value) between 2007 and 2010, but an increase of 9 PCI points between 2007 and 2014 and an increase in 12 points between 2010 and 2014. Therefore, the two records associated with these increases are removed. The PCI of section A13B1 increased from 15 to 100 between 2007 and 2010, clearly indicating a reconstruction during that timeframe. For this reason, this record and the one following it are removed. This section then exhibits increased deterioration (i.e., a reduced PCI value) between 2010 and 2014 allowing the corresponding record to be kept in the dataset. Once the 1,219 records from the original dataset generated by steps 1, 2, and 3 are processed in accordance with this consideration, 861 valid records corresponding to 629 sections remain in the dataset and can be used in the analysis.
Table 2.3: Example data entries with removed entries stricken through

<table>
<thead>
<tr>
<th>Network</th>
<th>Branch</th>
<th>Section</th>
<th>PCI Yr</th>
<th>Age</th>
<th>PCIa</th>
<th>Family</th>
<th>ROD</th>
<th>PCIa Yr</th>
<th>PCIa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A10B</td>
<td>2007</td>
<td>20</td>
<td>92</td>
<td>PCC Apron</td>
<td>0.93181</td>
<td>2010</td>
<td>93</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A10B</td>
<td>2007</td>
<td>20</td>
<td>92</td>
<td>PCC Apron</td>
<td>0.93181</td>
<td>2014</td>
<td>86</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A10B</td>
<td>2010</td>
<td>23</td>
<td>93</td>
<td>PCC Secondary</td>
<td>0.8</td>
<td>2014</td>
<td>86</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A18B</td>
<td>2007</td>
<td>19</td>
<td>69</td>
<td>PCC Apron</td>
<td>0.93181</td>
<td>2010</td>
<td>66</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A18B</td>
<td>2007</td>
<td>19</td>
<td>69</td>
<td>PCC Apron</td>
<td>0.93181</td>
<td>2014</td>
<td>78</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A18B</td>
<td>2014</td>
<td>22</td>
<td>66</td>
<td>PCC Secondary</td>
<td>0.8</td>
<td>2014</td>
<td>78</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A13B1</td>
<td>2007</td>
<td>40</td>
<td>15</td>
<td>APC</td>
<td>1.2177</td>
<td>2010</td>
<td>100</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A13B1</td>
<td>2007</td>
<td>40</td>
<td>15</td>
<td>APC</td>
<td>1.2177</td>
<td>2014</td>
<td>97</td>
</tr>
<tr>
<td>Langley</td>
<td>AWEST</td>
<td>A13B1</td>
<td>2010</td>
<td>43</td>
<td>100</td>
<td>APC</td>
<td>2.2</td>
<td>2014</td>
<td>97</td>
</tr>
</tbody>
</table>

Table 2.4 summarizes the data records used in the analysis according to base, climate zone, and pavement type. Notice, Wright-Patterson has the smallest number of records with 45 and Nellis has the largest at 307. The freeze/dry climate zone has the fewest records with 93, and the no freeze/dry climate zone has the most with 307. And, the majority of records (607 of 861) correspond to rigid pavements.

Table 2.4: Data summary

<table>
<thead>
<tr>
<th>Freeze/Wet</th>
<th>No Freeze/Wet</th>
<th>No Freeze/Dry</th>
<th>Freeze/Dry</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langley</td>
<td>Wright-Patterson</td>
<td>Eglin</td>
<td>MacDill</td>
<td>Nellis</td>
</tr>
<tr>
<td>Flexible</td>
<td>52</td>
<td>19</td>
<td>45</td>
<td>66</td>
</tr>
<tr>
<td>Rigid</td>
<td>157</td>
<td>26</td>
<td>51</td>
<td>45</td>
</tr>
<tr>
<td>Base Total</td>
<td>209</td>
<td>45</td>
<td>96</td>
<td>111</td>
</tr>
<tr>
<td>Climate Total</td>
<td>254</td>
<td>207</td>
<td>307</td>
<td>93</td>
</tr>
</tbody>
</table>
Chapter 3: Model Specification

3.1 Introduction

In this chapter, alternative forecasting models are specified considering several possible dependent and explanatory variables. The premise is that systematic forecasting errors may be related to section characteristics, environment, and forecasting horizon. The overall model building approach consists of three primary steps. First, candidate dependent variables are calculated from the variables available in the dataset. Second, explanatory data analysis is conducted. Third, candidate model specifications are developed. The sections in this chapter describe this process in detail and present the resulting model specifications that are estimated in Chapter 4.

3.2 Dependent and Explanatory Variables

3.2.1 Dependent Variables

The goal is to quantify and model forecasting error in the Air Force’s existing methodology. To start, forecasting error was calculated as:

\[ \text{Error} = \text{PCI}_{if} - \text{PCI}_{ia} \]  

(3.1)

The variable “Error” quantifies the magnitude of forecasting error for each pavement section. There is a nuance that this calculation misses, however. For entries where the pavement section is in better condition, the PCI values of both \( \text{PCI}_{if} \) and \( \text{PCI}_{ia} \) are higher than an entry for a section in worse condition. An error of 5 PCI points for a pavement
section that started at 100 represents an overall error of 5%. The same amount of error in a pavement section with PCI$_{i0}$ of 50, represents a 10% error.

For this reason, the value of error is scaled based on the pavements condition to create a relative error variable:

$$\text{Rel}_\text{Error} = \frac{\text{PCI}_{if} - \text{PCI}_{ia}}{\text{PCI}_{ia}}$$  \hfill (3.2)

The variable Rel$_\text{Error}$ represents the difference between a forecasted value and the true value expressed as a percentage of the true value. This formulation includes both positive and negative values. While a relative error value of, for example, 0.05 and -0.05 both indicate that the forecast was off from the true value by 5% of the true value, in practical terms there is a different consequence to positive and negative forecast errors. Positive values of Rel$_\text{Error}$ indicate an "optimistic" forecast, meaning the section is forecasted to be in better condition than it actually is. In other words, deterioration was under-forecasted. Likewise, a "pessimistic" forecast produces negative value of Rel$_\text{Error}$ and indicates that the section is actually in better condition than forecasted. Deterioration was over-forecasted.

Forecasts are used to produce maintenance and rehabilitation plans and schedules. Forecast error in the optimistic direction results in planning and budgeting for less maintenance and repair than is necessary. Pavements engineers will not request enough funding to cover the work that should be accomplished to maintain the pavement at a specified level. This lack of maintenance will result in increased user costs due to deteriorated pavements as well as more expensive maintenance and repair budgets in the future. Forecast error in the pessimistic direction results in planning and budgeting for
more maintenance and repair than is necessary. Time is exerted and money is spent on providing a maintenance level of effort that is not necessary. For this reason, Rel_Error, with both positive and negative values is used as the primary dependent variable in all model specifications.

Interpreting a relative error model where the signs associated with the parameters have different meaning can be difficult. Intuition could explain why certain characteristics of a pavement section may lead to more or less forecasting error. Determining how these same characteristics influence the optimism or pessimism of a forecast is more challenging. For this reason, the absolute value of relative error is calculated and also considered as a dependent variable ‘ABS_Rel_Error’ for model interpretation purposes only – i.e., not for correcting systematic errors in forecasts as discussed in section 1.1. The results of a model with ABS_Rel_Error as the dependent variable can provide insight into the effect that the explanatory variables have on the amount of forecasting error in absolute terms. Positive parameter values indicate an increase in the overall amount of error and negative values indicate a decrease. Such results could support the interpretation of the estimation results based on Rel_Error as the dependent variable.

3.2.2 Explanatory Variables

The following variables have already been extracted from the source reports as described in Chapter 2 and may serve as candidate explanatory variables in the model:

- Network: The pavement network, or base of the corresponding record.
- Branch: The branch name for the record.
Section: Section identifier for the record.

PCI_{i0\_Yr}: The year in which the PCI evaluation was conducted and rate of deterioration forecasted.

Age: Age of the pavement section at PCI_{i0\_Yr}.

PCI_{i0}: The observed PCI of pavement section i at time PCI_{i0\_Yr}.

Family: The deterioration family to which the record belongs.

ROD: The rate of deterioration (PCI points/year) forecasted for the family of the record.

PCI_{i_{a\_Yr}}: The year in which the PCI was re-evaluated for the corresponding record.

PCI_{i_{a}}: The observed PCI of section i at time PCI_{i_{a\_Yr}}.

PCI_{i_{f}}: Forecasted PCI of pavement section i at time PCI_{i_{a\_Yr}}.

The variables above allow for forecasting error to be quantified and provide possible explanatory variables. However, for the modeling of the error to be effective, considering additional explanatory variables may be useful. The following variables are created as additional candidate explanatory variables because of their possible impact on the rate of deterioration of pavements.

**Forecast Horizon**: Forecasted PCI values depend on the rate of deterioration and the duration until a time in the future when a forecast is desired (e.g., five to ten years). It is expected that this duration will impact the accuracy of the forecast. The rate of deterioration at time PCI_{i0\_Yr} is subject to estimation error and may not represent the rate of deterioration in the future. These errors will continue to magnify the longer the duration of the forecast is. To capture this effect, a variable named ‘Horizon’ was added...
to the dataset. Horizon is defined as the number of years that deterioration is being forecast into the future and is calculated as follows:

\[
\text{Horizon} = \text{PCI}_{ia\ Yr} - \text{PCI}_{i0\ Yr} \tag{3.3}
\]

That is, Horizon is the number of years between the report that the rate of deterioration (ROD) value is extracted from and the report the actual PCI (PCI\(_{ia}\)) value is extracted from.

**Construction Material:** As far back as 1920 when the Bates Experimental Road, located in Illinois, provided the first major empirical testing of pavements, it is been shown that different materials used in pavement construction will exhibit different performance and deterioration characteristics (Yoder and Witczak, 1975). As already discussed, the current forecast method is based on categorizing pavement sections into deterioration families considering construction materials, function, and loading characteristics. If deterioration is forecasted at different rates based on construction material, for example, it is possible that these forecasts reflect different levels of error. Deterioration family is a categorical variable, and the dataset contains sections from a total of 29 different families. To include family in the model, 28 categorical variables would be needed which would make the model cumbersome and not generalizable. While there are 29 specific families across all six bases and 15 evaluation reports, the effect of construction material can be captured with a single variable. All families consist exclusively of either rigid or flexible pavement sections. Rigid and flexible pavements are also evaluated using different distress types and PCI deduction curves in PAVER (ASTM, 2012). With different deterioration characteristics and measurement techniques, it follows that the
type of pavement section being forecasted may contain different types of forecast error. A categorical variable ‘Pav_Type’ was added to the dataset to account for this possible effect. The Pav_Type variable is given a value of 1 for flexible pavement sections and 0 for rigid sections.

Climate Zone: The primary causes of deterioration in any type of pavement include environmental conditions (Jain et al., 2005). Specifically, the number and severity of freeze-thaw cycles that the pavement is subject to as well as the amount of local precipitation impacts the amount and severity of deterioration (Ben-Akiva and Ramaswamy, 1993; Archilla and Madanat, 2008). In recent years, Air Force pavements engineers have noticed that the majority of pavement distresses observed on airfield pavements were attributable to environmental effects. They commissioned a study into the impact of the environment on pavement deterioration rates across the United States and found that weather data could be used to define different climate zones such that pavements located within these zones experience the same type of environmental phenomenon. The deterioration rates of pavements within these zones were then compared and statistical differences were found between the rates of deterioration of some of the pavement types across zones (Meihaus, 2013). From the climate zone map and definitions presented in section 2.1.3, categorical climate variables can be created to capture the climatic effects on forecasting error. The six bases used to create the dataset were identified on the climate zone map. Four categorical variables were created, one to represent each climate zone. The climate zone variables were given a value of 1 if the base is located in the respective climate zone, and 0 otherwise as follows:
o CZ1_F_W: Categorical variable for climate zone 1, the freeze-wet climate,
   1 = section is in CZ1, 0 = otherwise.

o CZ2_NF_W: Categorical variable for climate zone 2, the no freeze/wet climate,
   1 = section is in CZ2, 0 = otherwise.

o CZ3_NF_D: Categorical variable for climate zone 3, the no freeze/dry climate,
   1 = section is in CZ3, 0 = otherwise.

o CZ4_F_D: Categorical variable for climate zone 4, the no freeze/dry climate,
   1 = section is in CZ4, 0 = otherwise.

**Location:** In addition to the effect that climate has on deterioration, there are also location specific factors that can influence the rates of pavement deterioration and therefore forecasting. Every base airfield is managed by local pavement engineers and subject to local operations and maintenance policies, procedures, and budgets. The effect that localized routine maintenance effectiveness has on deterioration should be accounted for in the deterioration model. However, it will be interesting to test this assumption by using a location variable to quantify any systemic error seen between locations. Since the dataset contains two bases each within CZ1 and CZ2 and one base each in CZ3 and CZ4, there is no need for a base variable for Nellis AFB in CZ3 and Cannon AFB in CZ4. The climate variable already uniquely identifies the location. However, to control for the effect of location within the CZ1 and CZ2 climate zones, a categorical variable is created to represent one base within each of these zones. The second base within each zone will act as the reference base. The following variables are added to the dataset:
o B1_Langley: Categorical variable representing Langley AFB, located in CZ1, 
   1 = Langley, 0 = otherwise.

o B2_Eglin: Categorical variable representing Eglin AFB, located in CZ2, 
   1 = Eglin, 0 = otherwise.

Given that climate zones 3 and 4 are represented by one base each in the dataset, it is 
important to recognize that the previously defined variables CZ3_NF_D and CZ4_F_D 
effectively also represent location variables as they also define Nellis Air Force Base, 
and Cannon Air Force Base, respectively. Therefore, for these two climate zones and 
bases, the effects of the environment and location would be indistinguishable in the 
estimated models.

3.3 Exploratory Analysis and Possible Variable Specifications

The effects of age, condition, forecast horizon, material type, climate, and 
location of a pavement section on forecasting error can be investigated now that the 
dataset includes the possible explanatory variables of interest. Examining these variables 
in relation to the dependent variable as well as their relationships with one another helps 
determining the model specifications to consider. The correlation coefficients among the 
dependent variable Rel_Error and a number of potential explanatory variables is shown in 
Table 3.1. Notice that Pav_Type, Horizon, CZ2_NF_W and CZ3_NF_D have the 
largest absolute correlations with the dependent variable Rel_Error. Other variables all 
have similar fairly low correlations with Rel_Error. Notice that some of the explanatory 
variables have high correlation values, specifically, B1_Langley with CZ1_F_W, and 
B2_Eglin with CZ2_NF_W. This is a result of the categorical variable definition. Since
Langley Air Force base is located in climate zone 1, all records from Langley have a correlation 1.0 with CZ1_F_W. Likewise, Eglin Air Force Base is located in climate zone 2, so all records from Eglin have a correlation 1.0 with CZ2_NF_W.

Table 3.1: Correlation coefficients for select variables

<table>
<thead>
<tr>
<th></th>
<th>Rel_Error</th>
<th>Age</th>
<th>PCIio</th>
<th>Horizon</th>
<th>Pav_Type</th>
<th>CZ1_F_W</th>
<th>CZ2_NF_W</th>
<th>CZ3_NF_D</th>
<th>CZ4_F_D</th>
<th>B1_Langley</th>
<th>B2_Eglin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rel_Error</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCIio</td>
<td>0.04</td>
<td>-0.57</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.15</td>
<td>-0.08</td>
<td>0.09</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.22</td>
<td>-0.14</td>
<td>-0.13</td>
<td>-0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>0.04</td>
<td>0.08</td>
<td>-0.12</td>
<td>-0.34</td>
<td>-0.02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.13</td>
<td>0.14</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.30</td>
<td>-0.36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.20</td>
<td>-0.13</td>
<td>0.06</td>
<td>0.16</td>
<td>-0.24</td>
<td>-0.48</td>
<td>-0.42</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ4_F_D</td>
<td>0.07</td>
<td>-0.12</td>
<td>0.08</td>
<td>0.27</td>
<td>0.00</td>
<td>-0.23</td>
<td>-0.20</td>
<td>-0.26</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.07</td>
<td>0.10</td>
<td>-0.09</td>
<td>-0.26</td>
<td>-0.06</td>
<td>0.88</td>
<td>-0.32</td>
<td>-0.42</td>
<td>-0.18</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>0.06</td>
<td>0.13</td>
<td>0.01</td>
<td>0.09</td>
<td>0.14</td>
<td>-0.23</td>
<td>0.63</td>
<td>-0.26</td>
<td>-0.12</td>
<td>-0.20</td>
<td>1</td>
</tr>
</tbody>
</table>

All the categorical variables take the value of 1 or 0 and, therefore, their graphical depiction with the dependent variable is not likely to be informative. They are included in the potential specifications and the statistical significance of their parameters will determine which ones are worthwhile to maintain in the model. To more closely examine possible relationships the dependent variable has with the continuous variables, pairwise scatter plots allow for some preliminary observations. Figure 3.1 shows pairwise scatter plots among the dependent variable Rel_Error and the explanatory variables Age, Horizon, and PCIio. There appears to be a negative association between Age and PCIio as would be expected and as reflected in the negative correlation of -0.57 shown in Table 3.1. Older pavement sections are expected to be in worse condition. However, Rel_Error and Age do not appear to exhibit a trend as is also reflected in the close to zero
correlation between the two variables shown in Table 3.1. Nevertheless, Age is included in the initial model specifications for the purpose of confirming its parameters statistical significance or lack of. A positive trend appears to exist between Rel_Error and PCI_{i0} even though the correlation is close to zero as shown in Table 3.1.

Figure 3.1: Pairwise scatter plot of select variables

To further evaluate the relationship between Rel_Error and PCI_{i0} and how it may be specified, a separate scatter plot of Rel_Error versus PCI_{i0} is shown in Figure 3.2. Clearly, there is a positive trend. Looking closer at the data, it seems that this relationship behaves differently for lower values of PCI_{i0} than it does for higher values. Accurately capturing the nuance of this relationship might be important in the overall modeling of relative error.
There are a number of model specifications that might be used to capture this relationship. In the evaluation reports from which data were extracted, the PCI is discretized into basic categories to aid in interpretation. The Air Force defines a “critical PCI” threshold of 70, below which either major maintenance and rehabilitation or complete reconstruction is required. A second threshold is set at PCI of 55 where the section is considered to have failed. That is sections with PCI values less than 55 will require complete reconstruction and sections with PCI values between 55 and 70 will require major maintenance and rehabilitation (Air Force Civil Engineer Center, 2015). All sections with PCI > 70 are considered to be in “Good” condition. Sections with PCI from 55 to 70 are considered to be in “Fair” condition, and sections with PCI < 55 are considered to be in “Poor” condition.
With relative error increasing along the spectrum of PCI\textsubscript{0}, one possible model specification would be to apply the Air Force definitions of these broad PCI categories to create a categorical variable for pavement condition to explore how forecast relative error is impacted by Good, Fair, and Poor pavement sections, respectively. Therefore, the following variables were added to the dataset for this purpose:

- **PCI\_Good**: Categorical variable, 1 = pavement section considered “good” by Air Force definition, 0 = otherwise.
- **PCI\_Fair**: Categorical variable, 1 = pavement section considered “fair” by Air Force definition, 0 = otherwise.
- **PCI\_Poor**: Categorical variable, 1 = pavement section considered “poor” by Air Force definition, 0 = otherwise.

The inclusion of the discretized condition variables in the model specification will produce a model of relative Rel\_Error with a stepwise form with respect to PCI\textsubscript{0}. The theoretical form of the categorical function is shown in Figure 3.3 for illustration purposes. All other factors held constant, the predicted relative error will be the same for all sections with PCI values within a category.
An alternative to the purely categorical specification is a piecewise linear one between relative error and PCI$_{i0}$. This function can be formulated by using the existing PCI$_{i0}$ variable as a baseline, and then adding a piecewise linear variable that will modify this relationship depending on the category PCI$_{i0}$ falls within. For this study, pavement sections in the poor category will be treated as the reference group for the piecewise linear function, meaning that sections with PCI < 55 will be represented by the existing variable PCI$_{i0}$ (multiplied by a parameter to be estimated). A piecewise variable ‘PW_Fair’ is added to the dataset to represent sections in the fair category for the piecewise linear specification. PW_Fair is calculated using the following equation:

$$PW_{_\text{Fair}} = PCI_{_\text{Fair}} \times (PCI_{i0} - 54) \quad (3.4)$$

where,

PCI$_{_\text{Fair}} = 1$ when $55 \leq PCI_{i0} \leq 70$, and 0 otherwise.
A piecewise variable ‘PW_Good’ was also added to the dataset to represent sections in the good category for the piecewise linear specification. PW_Good is calculated using the following equation:

\[ PW_{G}ood = PW_{Fair} + (PCI_{Good} \times (PCI_{i0} - 70)) \] (3.5)

where,

\[ PCI_{Good} = 1 \text{ when } PCI_{i0} > 70, \text{ and } 0 \text{ otherwise.} \]

When added to the model, the effect that the parameter \( PCI_{i0} \) has on relative error is either magnified or mitigated by the parameters of \( PW_{Fair} \) and \( PW_{Good} \). The result is a function continuous with respect to \( PCI_{i0} \) with breakpoints that coincide with the PCI category thresholds. Figure 3.4 shows the general form of the piecewise linear function as prescribed by including \( PCI_{i0}, PW_{Fair}, \text{ and } PW_{Good} \) as variables in the specification. Notice, only the slope of the function changes at the breakpoints but the value of relative error does not shift at the breakpoint as it does in the step function of Figure 3.3. This figure demonstrates the conceptual piecewise function and is shown for illustration only. Depending on the parameter values of \( PCI_{i0}, PW_{Fair}, \text{ and } PW_{Good} \), the slope of the lines in the poor, fair, and good categories could vary and take positive or negative values resulting in very different functions.
Both the categorical and piecewise linear specifications are considered to follow the existing Air Force discretization of condition. These categories are representative of when pavement sections require certain levels of effort in terms of maintenance and repair, rehabilitation, or reconstruction and, therefore, necessary for decision making. However, they are arbitrary in terms of the specification of the forecasting relative error model. An alternative is to perform a transformation on the PCI$_{i0}$ variable to create a variable that continuously fits the trend in the raw data without any abrupt changes in level of slope. From visual inspection of Figure 3.2, a square root function or natural logarithm function seem to be candidates that follow the shape the sample data exhibits. These transformations are simple to perform and provide an alternative that produces a continuous function through the data as opposed to one with discrete break points (where in level of slope). Continuous specifications may perform better, especially for PCI$_{i0}$.
values close to one of categorical thresholds and, consequently, are worth consideration when formulating model specification alternatives. Therefore, the following variables are added to the dataset for this purpose:

\[
\text{PCI}_{i0} \_\text{Sqrt} = \sqrt{\text{PCI}_{i0}} \tag{3.6}
\]

\[
\text{PCI}_{i0} \_\text{LN} = \ln \text{PCI}_{i0} \tag{3.7}
\]

Finally, a unique identifier is added to the dataset called ‘Record\_Index’. This variable is added as a simple 1 to n index of the entire dataset so that each record has a unique identifier. Record\_Index represents a pavement section i, observed at time ‘PCI\_Yr\_i0’ and forecasted ‘Horizon’ years into the future. A complete list of the considered variables and their definitions is shown in Table 3.2 for reference.

**Table 3.2: Complete list of variables and their definitions**

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Rel_Error</td>
<td>Difference between a forecasted value and the true value expressed as a percentage of the true value</td>
</tr>
<tr>
<td>ABS_Rel_Error</td>
<td>Absolute value of Rel_Error</td>
</tr>
<tr>
<td><strong>Explanatory Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td>The pavement network, or base of the corresponding record</td>
</tr>
<tr>
<td>Branch</td>
<td>The branch name for the record</td>
</tr>
<tr>
<td>Section</td>
<td>Section identifier for the record</td>
</tr>
<tr>
<td>PCI_Yr_i0</td>
<td>The year in which the PCI evaluation was conducted and rate of deterioration forecasted</td>
</tr>
<tr>
<td>Age</td>
<td>Age of the pavement section at PCI_Yr_i0</td>
</tr>
<tr>
<td>PCI_i0</td>
<td>The observed PCI of pavement section i at time PCI_Yr_i0</td>
</tr>
<tr>
<td>Family</td>
<td>The deterioration family to which the record belongs</td>
</tr>
<tr>
<td>ROD</td>
<td>The rate of deterioration (PCI points/year) forecasted for the family of the record</td>
</tr>
<tr>
<td>PCI_Yr_i0</td>
<td>The year in which the PCI was re-evaluated for the corresponding record</td>
</tr>
<tr>
<td>PCI_i0</td>
<td>The observed PCI of section i at time PCI_Yr_i0</td>
</tr>
<tr>
<td>PCI_if</td>
<td>Forecasted PCI of pavement section i at time PCI_i0_Yr</td>
</tr>
<tr>
<td>Horizon</td>
<td>Number of years that deterioration is being forecast into the future</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>Categorical variable representing the construction material of the pavement section, 1=flexible, 0=rigid</td>
</tr>
</tbody>
</table>

Continued
Table 3.2 Continued:

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory Variables</strong></td>
<td></td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>Categorical variable for climate zone 1, the freeze-wet climate, 1 = record is in CZ1, 0 = otherwise</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>Categorical variable for climate zone 2, the no freeze/wet climate, 1 = record is in CZ2, 0 = otherwise</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>Categorical variable for climate zone 3, the no freeze/dry climate, 1 = record is in CZ3, 0 = otherwise</td>
</tr>
<tr>
<td>CZ4_F_D</td>
<td>Categorical variable for climate zone 4, the no freeze/dry climate, 1 = record is in CZ4, 0 = otherwise</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>Categorical variable representing Langley AFB, located in CZ1, 1 = Langley, 0 = otherwise</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>Categorical variable representing Eglin AFB, located in CZ2, 1 = Eglin, 0 = otherwise</td>
</tr>
<tr>
<td>PCI_Good</td>
<td>Categorical variable, 1 = pavement section considered “good” by Air Force definition, 0 = otherwise</td>
</tr>
<tr>
<td>PCI_Fair</td>
<td>Categorical variable, 1 = pavement section considered “fair” by Air Force definition, 0 = otherwise</td>
</tr>
<tr>
<td>PCI_Poor</td>
<td>Categorical variable, 1 = pavement section considered “poor” by Air Force definition, 0 = otherwise</td>
</tr>
<tr>
<td>PW_Fair</td>
<td>Piecewise variable equal to (PCI₀ – 54) for sections with PCI₀ between 55 and 70</td>
</tr>
<tr>
<td>PW_Good</td>
<td>Piecewise variable equal to (PW_Fair + PCI₀ - 70) for sections with PCI₀ greater than 70</td>
</tr>
<tr>
<td>PCI₀_Sqrt</td>
<td>Square root transformation of the PCI₀ variable (√PCI₀)</td>
</tr>
<tr>
<td>PCI₀_LN</td>
<td>Natural log transformation of the PCI₀ variable (ln PCI₀)</td>
</tr>
<tr>
<td>Section Index</td>
<td>Unique identifier representing a pavement section i, observed at time “PCI₀_Yr” and forecasted “Horizon” years into the future</td>
</tr>
</tbody>
</table>

3.4 Model Specification Alternatives

Based on the explanatory variables available as well as the preliminary analysis of the response data, four model specification alternatives are proposed. Some of the explanatory variables will be common throughout, the major difference between these models is how they specify the relationship between PCI₀ and relative error. The alternatives shown below are the starting points for each model. Depending on the results once the proposed models are estimated, some variables may drop out or be modified.
Model specification 1 is based on using $\text{PCI}_{10}$ as a categorical variable and is given by the following:

$$
\text{Rel Error} = \beta_0 + \beta_1(\text{Age}) + \beta_2(\text{Horizon}) + \beta_3(\text{Pav\_Type}) + \beta_4(\text{B1\_Langley})
+ \beta_5(\text{B2\_Eglin}) + \beta_6(\text{CZ1\_F\_W}) + \beta_7(\text{CZ2\_NF\_W}) + \beta_8(\text{CZ3\_NF\_D}) + \beta_9(\text{PCI\_Poor})
+ \beta_{10}(\text{PCI\_Fair}) + \epsilon \tag{3.5}
$$

All the variables are as defined in Table 3.2. Note, CZ4\_F\_D and PCI\_Good represent the references for the climate and condition state variables, respectively.

Model specification 2 uses a piecewise linear transformation of $\text{PCI}_{10}$ and is given by the following:

$$
\text{Rel Error} = \beta_0 + \beta_1(\text{Age}) + \beta_2(\text{Horizon}) + \beta_3(\text{PCI}_{10}) + \beta_4(\text{Pav\_Type}) + \beta_5(\text{B1\_Langley})
+ \beta_6(\text{B2\_Eglin}) + \beta_7(\text{CZ1\_F\_W}) + \beta_8(\text{CZ2\_NF\_W}) + \beta_9(\text{CZ3\_NF\_D}) + \beta_{10}(\text{PW\_Fair})
+ \beta_{11}(\text{PW\_Good}) + \epsilon \tag{3.6}
$$

All the variables are as defined in Table 3.2. As in specification 1, CZ4\_F\_D represents the reference for the climate variables.

Model specification 3 uses a square root transformation of $\text{PCI}_{10}$ and is given by the following:

$$
\text{Rel Error} = \beta_0 + \beta_1(\text{Age}) + \beta_2(\text{Horizon}) + \beta_3(\text{PCI}_{10}) + \beta_4(\text{Pav\_Type}) + \beta_5(\text{B1\_Langley})
+ \beta_6(\text{B2\_Eglin}) + \beta_7(\text{CZ1\_F\_W}) + \beta_8(\text{CZ2\_NF\_W}) + \beta_9(\text{CZ3\_NF\_D}) + \beta_{10}(\text{Sqrt\_PCI}_{10})
+ \epsilon \tag{3.7}
$$

All the variables are as defined in Table 3.2. As in previous specifications, CZ4\_F\_D represents the reference for the climate variables.
Model specification 4 uses a natural logarithm transformation of PCI\textsubscript{i0} and is
given by the following:

\[
\text{Rel Error} = \beta_0 + \beta_1(\text{Age}) + \beta_2(\text{Horizon}) + \beta_3(\text{PCI}_{i0}) + \beta_4(\text{Pav\_Type}) + \beta_5(\text{B1\_Langley}) \\
+ \beta_6(\text{B2\_Eglin}) + \beta_7(\text{CZ1\_F\_W}) + \beta_8(\text{CZ2\_NF\_W}) + \beta_9(\text{CZ3\_NF\_D}) + \beta_{10}(\ln(\text{PCI}_{i0})) \\
+ \epsilon
\]

(3.8)

All the variables are as defined in Table 3.2. Again, CZ4\_F\_D represents the reference
for the climate variables.
Chapter 4: Estimation Results and Model Evaluation

4.1 Summary Statistics of Variables and Outlier Analysis

This chapter presents the estimation results of the model alternatives. Before commencing with the approaches followed, the dependent and explanatory variables are summarized and investigated for the possible presence of outliers. The summary statistics of all variables in the dataset are shown in Table 4.1 (The median for the binary categorical variables are not shown as they are not meaningful). Examining the dependent variable, there appears to be a high potential for outliers. Standard convention considers a data point greater than two standard deviations from the mean a statistical outlier. The Rel_Error variable has a standard deviation of 0.1836, a mean of 0.0196, and a range (maximum – minimum) of 2.3138. Clearly, there are data points with Rel_Error values well in excess of two standard deviations from the mean. Using the conventional definition for outliers, the acceptable range for relative error values in this sample is (-0.3476, 0.3868). With a minimum value of -0.5 and maximum value of 1.8138, there are clearly points outside the acceptable limits.
Table 4.1: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Std Dev</th>
<th>Coef of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rel_Error</td>
<td>0.0196</td>
<td>-0.0103</td>
<td>-0.5000</td>
<td>1.8138</td>
<td>0.1836</td>
<td>9.3437</td>
</tr>
<tr>
<td>ABS_Rel_Error</td>
<td>0.0960</td>
<td>0.0545</td>
<td>0.0000</td>
<td>1.8138</td>
<td>0.1577</td>
<td>1.6422</td>
</tr>
<tr>
<td>Age</td>
<td>17.4669</td>
<td>13.0000</td>
<td>0.0000</td>
<td>68.0000</td>
<td>15.2229</td>
<td>0.8715</td>
</tr>
<tr>
<td>PCI_i0</td>
<td>84.8153</td>
<td>90.0000</td>
<td>13.0000</td>
<td>100.0000</td>
<td>16.5283</td>
<td>0.1949</td>
</tr>
<tr>
<td>PCI_ia</td>
<td>78.6597</td>
<td>83.0000</td>
<td>14.0000</td>
<td>100.0000</td>
<td>17.9853</td>
<td>0.2286</td>
</tr>
<tr>
<td>PCI_if</td>
<td>79.0025</td>
<td>84.0000</td>
<td>8.5000</td>
<td>98.0660</td>
<td>16.9552</td>
<td>0.2146</td>
</tr>
<tr>
<td>Horizon</td>
<td>5.5203</td>
<td>6.0000</td>
<td>3.0000</td>
<td>9.0000</td>
<td>1.8829</td>
<td>0.3411</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.2950</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.4563</td>
<td>1.5468</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>0.2950</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.4563</td>
<td>1.5468</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.2404</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.4276</td>
<td>1.7785</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>0.3566</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.4793</td>
<td>1.3441</td>
</tr>
<tr>
<td>CZ4_F_D</td>
<td>0.1080</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.3106</td>
<td>2.8754</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.2427</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.4290</td>
<td>1.7673</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>0.1115</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.3149</td>
<td>2.8245</td>
</tr>
<tr>
<td>PCI_Good</td>
<td>0.8374</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.3692</td>
<td>0.4409</td>
</tr>
<tr>
<td>PCI_Fair</td>
<td>0.0987</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.2985</td>
<td>3.0232</td>
</tr>
<tr>
<td>PCI_Poor</td>
<td>0.0639</td>
<td>-</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.2447</td>
<td>3.8304</td>
</tr>
<tr>
<td>PW_Fair</td>
<td>31.7259</td>
<td>36.0000</td>
<td>0.0000</td>
<td>46.0000</td>
<td>14.0228</td>
<td>0.4420</td>
</tr>
<tr>
<td>PW_Good</td>
<td>17.3252</td>
<td>20.0000</td>
<td>0.0000</td>
<td>30.0000</td>
<td>11.0517</td>
<td>0.6379</td>
</tr>
<tr>
<td>PCI_i0_Sqrt</td>
<td>9.1543</td>
<td>9.4868</td>
<td>3.6056</td>
<td>10.0000</td>
<td>1.0075</td>
<td>0.1101</td>
</tr>
<tr>
<td>PCI_i0_LN</td>
<td>4.4137</td>
<td>4.4998</td>
<td>2.5649</td>
<td>4.6052</td>
<td>0.2560</td>
<td>0.0580</td>
</tr>
</tbody>
</table>

To look closer at the distribution of the Rel_Error variable, an empirical cumulative distribution function (E-CDF) for the relative error is shown in Figure 4.1.

This plot confirms that some of the data points are potential outliers. For example, there are a number of data entries with a relative error greater than one. This means that the PCI forecast for a given pavement section is more than 100% off from what the actual PCI of that section turned out to be when measured. Since this study deals with empirical data collected over time, there is potential that some pavement sections may have experienced unusual deterioration rates due to abnormal loading, traffic pattern changes,
extreme environmental conditions, or faulty construction. Figure 4.2 shows the two standard deviation outlier boundaries added to the E-CDF of Figure 4.1.

Figure 4.1: Empirical cumulative distribution function for dependent variable Rel_Error

Figure 4.2: Empirical cumulative distribution function for dependent variable Rel_Error with upper and lower outlier boundaries shown
Applying the two standard deviation threshold seems to be reasonable. The outlier boundaries fall at what appears to be natural distinguishing values of relative error where the data points outside the plus and minus two standard deviations from the sample mean range deviate markedly from the other data points by exhibiting seemingly large and unrealistic values of relative error. Therefore, records with values less than -0.3476 (sample mean – 2 x standard deviation) or greater than 0.3868 (sample mean + 2 x standard deviation) were removed. In doing so, the sample size decreases from 861 to 834 records.

4.2 General Estimation and Evaluation Approaches

The alternative model specifications as prescribed in section 3.4 are estimated using the regression function in Microsoft Excel. P-values for each explanatory variable were examined and variables with p-values greater than 0.2 were considered not to be statistically significant and removed from the specification sequentially before re-estimating the model. In cases where more than one variable was found not to be statistically significant in the initial estimated model, the variable with the largest p-value is removed first and the model is re-estimated before assessing and removing any additional variables and re-estimating the resulting specification.

Once the final model is determined for each alternative specification, the alternatives are narrowed down based on statistical significance of the explanatory variables, overall model fit and interpretability. The final candidate models are then fully interpreted, and the effectiveness of these models in producing corrected forecasts is evaluated. To perform meaningful evaluation of the final models, a sample of the records
in the dataset not used to estimate the parameters of the model is needed. Estimating model parameters on sample data and then using the same data to evaluate the model’s effectiveness may result in false confidence in the model.

Therefore, a statistical analysis technique known as out-of-sample validation is used. The model parameters are estimated using a subset of the available sample called the training set. The portion of the available data not used in model estimation is called the test set (or hold-out sample), and is used to assess model effectiveness (Picard and Berk, 1990). Treating the test sample as a “new” dataset, applying the estimated model to these data, and analyzing the accuracy of the forecasts gives insight into how the model will generalize to other independent datasets.

In studies with very large sample sizes, hold-out samples are typically split so that half of the data are used in model estimation and the other half in evaluation. The smaller the available sample size for estimation, the greater impact on the reduced predictive power of the model because more valuable information would be excluded from the estimation process. In studies with extremely small sample sizes, this problem eliminates out-of-sample evaluation as a viable technique all together (Hawkins et al., 2002). The sample size in this study is in between these two extremes and, therefore, a hold-out sample consisting of 20% of the records in the dataset was used. That is, 80% of the records of the available dataset are used to estimate the models and the remaining 20% is used to evaluate the effectiveness.

With a total sample size of 834 records, 667 records are selected at random and used in model estimation. The remaining 167 records are used as the test set for model
evaluation. To ensure the selection is made randomly, a variable named ‘RAND’ is added to the dataset. The random number generator in Excel was used to create a random number uniformly distributed between 0 and 1 for each record. Once generated, the dataset was sorted based on the value of RAND from smallest to largest. The first 667 records are used as the training set and the remaining 167 are used as the test set.

4.3 Model Estimation Results

For each of the four alternative model specifications, the estimation results of the full initial specification are presented, followed by the results of the final specification once the variables parameter estimates that are found not to be statistically significant are removed. All regression model estimations were performed using the regression tool in Microsoft Excel.

Table 4.2 shows the estimation results obtained for the initial categorical model as specified in equation 3.5. The parameter of the variable Age is the only one found not to be statistically significant. Therefore, Age is removed from the specification and the model is re-estimated. The results are shown in Table 4.3. This final model specification has variables with parameters that are all statistically significant at the 80% confidence level (i.e., test size, $\alpha = 0.2$).
Table 4.2: Estimation results for initial model specification 1 - Categorical

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.102314</td>
<td>4.86E-08</td>
<td>5.52083</td>
</tr>
<tr>
<td>Age</td>
<td>0.00012</td>
<td>0.67285</td>
<td>0.422429</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.01215</td>
<td>4.32E-09</td>
<td>-5.95173</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.029593</td>
<td>0.00085</td>
<td>3.351349</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.093666</td>
<td>3.06E-07</td>
<td>5.17345</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.03249</td>
<td>0.031185</td>
<td>-2.15935</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.13008</td>
<td>5.86E-10</td>
<td>-6.28816</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.020423</td>
<td>0.198185</td>
<td>1.288042</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06227</td>
<td>8.98E-07</td>
<td>-4.96019</td>
</tr>
<tr>
<td>PCI_Poor</td>
<td>-0.09788</td>
<td>1.53E-08</td>
<td>-5.73043</td>
</tr>
<tr>
<td>PCI_Fair</td>
<td>-0.04972</td>
<td>0.000352</td>
<td>-3.59282</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2186$

Table 4.3: Estimation results for final model specification 1 - Categorical

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.103832</td>
<td>1.67E-08</td>
<td>5.714823</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.01217</td>
<td>4.05E-09</td>
<td>-5.96264</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.028539</td>
<td>0.000792</td>
<td>3.371429</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.093788</td>
<td>2.89E-07</td>
<td>5.184078</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.03207</td>
<td>0.032946</td>
<td>-2.13726</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.1296</td>
<td>6.22E-10</td>
<td>-6.27836</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.021556</td>
<td>0.167998</td>
<td>1.38019</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06201</td>
<td>9.51E-07</td>
<td>-4.94864</td>
</tr>
<tr>
<td>PCI_Poor</td>
<td>-0.09482</td>
<td>1.48E-09</td>
<td>-6.13355</td>
</tr>
<tr>
<td>PCI_Fair</td>
<td>-0.04771</td>
<td>0.000258</td>
<td>-3.67388</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2195$

Table 4.4 shows the estimation results for the piecewise linear model as specified in equation 3.6. The parameter of the PW_Good variable in this estimation is not found to be statistically significant. Given the specification of the piecewise linear model as reflected in the addition of the variables PW_Fair and PW_Good (see equations 3.4 and 3.5), this result indicates that there is no difference between “Fair” pavement sections and
“Good” pavement sections in terms of the rate of change of the relative error with respect to PCI$_{i0}$. Moreover, the statistical significance of the parameter of the PW_Fair variable indicates that there is a difference in the rate of change of the relative error with respect to PCI$_{i0}$ for pavement sections that are in the “Poor” category versus the “Fair” or “Good” categories. After PW_Good is removed from the model, the Age parameter is still not found to be statistically significant and, as a result, is removed from the specification. The estimation results for the final piecewise linear specification are shown in Table 4.5. This final model specification has variables with parameters that are all statistically significant at the 80% confidence level (i.e., test size, $\alpha = 0.2$).

Table 4.4: Estimation results for initial model specification 2 – Piecewise linear

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.27882</td>
<td>2.44E-05</td>
<td>-4.25063</td>
</tr>
<tr>
<td>Age</td>
<td>0.000342</td>
<td>0.267373</td>
<td>1.110077</td>
</tr>
<tr>
<td>PCI$_{i0}$</td>
<td>0.006293</td>
<td>6.62E-06</td>
<td>4.542333</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.01259</td>
<td>7.14E-10</td>
<td>-6.25575</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.033129</td>
<td>0.000187</td>
<td>3.756968</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.089224</td>
<td>8.76E-07</td>
<td>4.965312</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.0323</td>
<td>0.029953</td>
<td>-2.17545</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.1249</td>
<td>1.78E-09</td>
<td>-6.10272</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.021504</td>
<td>0.168556</td>
<td>1.378383</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.05937</td>
<td>2.26E-06</td>
<td>-4.7713</td>
</tr>
<tr>
<td>PW_Fair</td>
<td>-0.00557</td>
<td>0.028355</td>
<td>-2.19721</td>
</tr>
<tr>
<td>PW_Good</td>
<td>0.000478</td>
<td>0.769985</td>
<td>0.292515</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2435$
Table 4.5: Estimation results for final model specification 2 – Piecewise linear

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.25738</td>
<td>9.35E-06</td>
<td>-4.46678</td>
</tr>
<tr>
<td>PCl0</td>
<td>0.006061</td>
<td>8.89E-08</td>
<td>5.408928</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.0126</td>
<td>6.54E-10</td>
<td>-6.27004</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.029927</td>
<td>0.000355</td>
<td>3.590152</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.09115</td>
<td>4.14E-07</td>
<td>5.114388</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.03205</td>
<td>0.03075</td>
<td>-2.16496</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.12533</td>
<td>1.35E-09</td>
<td>-6.14944</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.024258</td>
<td>0.114937</td>
<td>1.57848</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.05989</td>
<td>1.61E-06</td>
<td>-4.84123</td>
</tr>
<tr>
<td>PW_Fair</td>
<td>-0.00517</td>
<td>7.08E-05</td>
<td>-3.99918</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. R² = 0.2443

The square root model was estimated as specified in equation 3.7 and the results are shown in Table 4.6. Again, the Age parameter is not found to be statistically significant and is removed from the specification and re-estimated. The estimation results for the final specification are shown in Table 4.7. This final model specification has variables with parameters that are all statistically significant at the 80% confidence level (i.e., test size, α = 0.2).

Table 4.6: Estimation results for initial model specification 3 - Square root

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.78964</td>
<td>1.06E-05</td>
<td>-4.43897</td>
</tr>
<tr>
<td>Age</td>
<td>0.000294</td>
<td>0.340796</td>
<td>0.953286</td>
</tr>
<tr>
<td>PCl0</td>
<td>-0.01253</td>
<td>7.66E-10</td>
<td>-6.24405</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.00869</td>
<td>0.001182</td>
<td>-3.25746</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.033205</td>
<td>0.000175</td>
<td>3.774636</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.092649</td>
<td>3.03E-07</td>
<td>5.175295</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.0333</td>
<td>0.024732</td>
<td>-2.25074</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.12739</td>
<td>7.34E-10</td>
<td>-6.2512</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.02101</td>
<td>0.177378</td>
<td>1.350324</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06126</td>
<td>9.63E-07</td>
<td>-4.94604</td>
</tr>
<tr>
<td>PCl0_Sqrt</td>
<td>0.176372</td>
<td>6.55E-05</td>
<td>4.017798</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. R² = 0.2467

54
Table 4.7: Estimation results for final model specification 3 - Square root

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.80256</td>
<td>7.17E-06</td>
<td>-4.52505</td>
</tr>
<tr>
<td>PCI$_{i0}$</td>
<td>-0.01255</td>
<td>7E-10</td>
<td>-6.25871</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.00933</td>
<td>0.000329</td>
<td>-3.61074</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.030497</td>
<td>0.000269</td>
<td>3.663333</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.093989</td>
<td>1.88E-07</td>
<td>5.266745</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.03289</td>
<td>0.02647</td>
<td>-2.22426</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.12743</td>
<td>7.22E-10</td>
<td>-6.2536</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.023463</td>
<td>0.126719</td>
<td>1.529107</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06149</td>
<td>8.75E-07</td>
<td>-4.96545</td>
</tr>
<tr>
<td>PCI$_{i0}$_Sqrt</td>
<td>0.184252</td>
<td>2.2E-05</td>
<td>4.274031</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2468$

The natural log model is estimated as specified in equation 3.8 and the results are shown in Table 4.8. Like the square root specification above, Age is not found to be statistically significant and is removed from the specification. Pavement age has been shown to be a significant contributor to deterioration in the literature (see George et al., 1989) so not finding it statistically significant in any of these model specifications is interesting. As discussed in section 3.3 and shown in Table 3.1, there appears to be a strong association and high correlation between pavement age, Age, and PCI$_{i0}$. Both support the conclusions of previous research that age impacts the amount of deterioration and since both Age and PCI$_{i0}$, as well as derivations of PCI$_{i0}$, were included in all model specifications, not finding Age statistically significant may be attributed to its high correlation with PCI$_{i0}$. The final estimation results for the natural log specification are shown in Table 4.9.
Table 4.8: Estimation results for initial model specification 4 - Natural log

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.15696</td>
<td>1.29E-05</td>
<td>-4.39521</td>
</tr>
<tr>
<td>Age</td>
<td>0.000289</td>
<td>0.348812</td>
<td>0.937568</td>
</tr>
<tr>
<td>PCIi0</td>
<td>-0.01255</td>
<td>7.02E-10</td>
<td>-6.25847</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.00315</td>
<td>0.014208</td>
<td>-2.45853</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.033099</td>
<td>0.000182</td>
<td>3.764057</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.092681</td>
<td>2.95E-07</td>
<td>5.180628</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.0333</td>
<td>0.024666</td>
<td>-2.25179</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.12726</td>
<td>7.42E-10</td>
<td>-6.24938</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.021289</td>
<td>0.171527</td>
<td>1.36881</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06096</td>
<td>1.05E-06</td>
<td>-4.92852</td>
</tr>
<tr>
<td>PCIi0_Ln</td>
<td>0.342506</td>
<td>4.74E-05</td>
<td>4.09562</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2474$

Table 4.9: Estimation results for final model specification 4 - Natural log

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.18502</td>
<td>6.95E-06</td>
<td>-4.53163</td>
</tr>
<tr>
<td>PCIi0</td>
<td>-0.01258</td>
<td>6.4E-10</td>
<td>-6.2735</td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.00353</td>
<td>0.003839</td>
<td>-2.90139</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.030433</td>
<td>0.000275</td>
<td>3.657736</td>
</tr>
<tr>
<td>B1_Langley</td>
<td>0.093993</td>
<td>1.85E-07</td>
<td>5.270599</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.03289</td>
<td>0.026381</td>
<td>-2.22558</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>-0.12729</td>
<td>7.32E-10</td>
<td>-6.25138</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.023711</td>
<td>0.122543</td>
<td>1.546178</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>-0.06116</td>
<td>9.62E-07</td>
<td>-4.9462</td>
</tr>
<tr>
<td>PCIi0_Ln</td>
<td>0.357302</td>
<td>1.57E-05</td>
<td>4.351121</td>
</tr>
</tbody>
</table>

No. of Observations = 667, Adj. $R^2 = 0.2476$

A summary table of the final estimation results for all four model specifications are shown in Table 4.9. Notice that the adjusted $R^2$ values are low for all model specifications. As noted in the exploration analysis of section 3.3, the empirical in-service data reflects a high degree of variability which can explain the relatively low degree of fit. The number of significant predictors in each model, however, with most being significant at the 95% or 99% confidence levels (test size, $\alpha = 0.05$ or 0.01)
indicates that the models are worthwhile. Amongst the four model specifications, the
categorical model has a lower adjusted $R^2$ than all the others. Therefore, this
specification is eliminated from further consideration and analysis. Also note that the
sign and magnitude of the parameter values for the square root model and the natural log
model are almost identical. This similarly suggests that considering both specifications is
unnecessary. The natural log model had a slightly higher value of adjusted $R^2$ and,
therefore, this specification is kept for further analysis and the square root specification
will no longer be considered.

Table 4.10: Summary of model alternatives

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Categorical</th>
<th>Coef</th>
<th>P-Val</th>
<th>Piecewise</th>
<th>Coef</th>
<th>P-Val</th>
<th>Square Root</th>
<th>Coef</th>
<th>P-Val</th>
<th>Natural Log</th>
<th>Coef</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.1038</td>
<td>1.67E-08</td>
<td>-0.2573</td>
<td>9.35E-06</td>
<td>-0.8026</td>
<td>7.17E-06</td>
<td>-1.185</td>
<td>6.95E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCI₀</td>
<td>-</td>
<td>-</td>
<td>0.0061</td>
<td>8.89E-08</td>
<td>-0.0126</td>
<td>7E-10</td>
<td>-0.0126</td>
<td>6.4E-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizon</td>
<td>-0.0122</td>
<td>4.05E-09</td>
<td>-0.0126</td>
<td>6.54E-10</td>
<td>-0.0093</td>
<td>0.000329</td>
<td>-0.0035</td>
<td>0.003839</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.0285</td>
<td>0.000792</td>
<td>0.0299</td>
<td>0.000355</td>
<td>0.0305</td>
<td>0.000269</td>
<td>0.0304</td>
<td>0.000275</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B₁_Langley</td>
<td>0.0938</td>
<td>2.89E-07</td>
<td>0.0912</td>
<td>4.14E-07</td>
<td>0.0940</td>
<td>1.88E-07</td>
<td>0.0940</td>
<td>1.85E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B₂_Eglin</td>
<td>-0.0321</td>
<td>0.032946</td>
<td>-0.0321</td>
<td>0.03075</td>
<td>-0.0329</td>
<td>0.02647</td>
<td>-0.0329</td>
<td>0.026381</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ₁_F_W</td>
<td>-0.1296</td>
<td>6.22E-10</td>
<td>-0.1253</td>
<td>1.35E-09</td>
<td>-0.1274</td>
<td>7.22E-10</td>
<td>-0.1273</td>
<td>7.32E-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ₂_NF_W</td>
<td>0.0216</td>
<td>0.167998</td>
<td>0.0243</td>
<td>0.114937</td>
<td>0.0235</td>
<td>0.126719</td>
<td>0.0237</td>
<td>0.122543</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ₃_NF_D</td>
<td>-0.0620</td>
<td>9.51E-07</td>
<td>-0.0599</td>
<td>1.61E-06</td>
<td>-0.0615</td>
<td>8.75E-07</td>
<td>-0.0612</td>
<td>9.62E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCI_Poor</td>
<td>-0.0948</td>
<td>1.48E-09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCI_Fair</td>
<td>-0.0477</td>
<td>0.000258</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PW_Fair</td>
<td>-</td>
<td>-</td>
<td>-0.0052</td>
<td>7.08E-05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCI₀_Sqrt</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.1843</td>
<td>2.2E-05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCI₀_Ln</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.3573</td>
<td>1.57E-05</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$= 0.2195</td>
<td>Adj. $R^2$= 0.2443</td>
<td>Adj. $R^2$= 0.2468</td>
<td>Adj. $R^2$= 0.2476</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of Observations = 667

While the square root and natural log specifications have the highest adjusted $R^2$
among the four specifications, the square root model is eliminated in favor of the
piecewise linear model because of the nature of the two specifications. Both the
categorical and piecewise models directly reflect the Air Force’s definitions of good, fair
and poor pavements based on the PCI$_{10}$ value. This feature may be useful in terms of interpretability and relatability to practitioners in the field as they are used to analyzing pavement sections in terms of their categorical condition state, rather than a specific PCI value. Therefore, it is desirable to consider the better of these two models, namely the piecewise linear one.

Both piecewise linear (specification 2) and natural log (specification 4) model alternatives exhibit significant parameters and similar adjusted $R^2$ values, yet one reflects the discrete definition of condition states and the other does not. Therefore, interpreting the estimation results of both specifications as well as using the test set for evaluating the effectiveness in correcting forecast errors are worthwhile and are presented in the subsequent sections.

4.4 Interpretations of Select Models

While the objective is to understand and model the relative error with regard to true PCI values as discussed in section 3.2.1, the interpretation of the association of the various explanatory variables with forecasting error regarding the sign of the error is challenging. Therefore, to support the interpretation, the absolute value of relative error, ABS_Rel_Error, is modeled and interpreted. Doing so allows for interpreting the association of the explanatory variables with the magnitude of the errors. Following the interpretations regarding the magnitude of error, the models estimated in section 4.3 are interpreted regarding the sign of the forecasting error.

Estimating the absolute relative error following the explanatory variable specifications of the piecewise linear and natural log alternative models resulted in the
parameters and p-values shown in Table 4.11. Not surprisingly, not all parameters are found to be statistically significant. The variables associated with the parameters that are not found to be significant are removed from the specifications and the models are re-estimated. The results are shown in Table 4.12. Notice, the signs and magnitudes of the parameters of the remaining variables are consistent with those of the corresponding relative error model estimation results. Naturally, it is possible that other variables, if included, may exhibit statistically significant parameters. Such an investigation is left for future research.

Table 4.11: Summary of model results with absolute relative error as dependent variable (identical specifications to section 4.3)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Piecewise</th>
<th></th>
<th>Natural Log</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef</td>
<td>P-Val</td>
<td>Coef</td>
<td>P-Val</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.1723</td>
<td>3.06E-05</td>
<td>0.3308</td>
<td>0.0768</td>
</tr>
<tr>
<td>PCI₀</td>
<td>-0.0023</td>
<td>0.0035</td>
<td>-0.0006</td>
<td>0.4995</td>
</tr>
<tr>
<td>Horizon</td>
<td>0.0069</td>
<td>1.73E-06</td>
<td>0.0069</td>
<td>1.78E-06</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.0434</td>
<td>7.3E-13</td>
<td>0.0434</td>
<td>8.13E-13</td>
</tr>
<tr>
<td>B₁_Langley</td>
<td>-0.0133</td>
<td>0.2966</td>
<td>-0.0137</td>
<td>0.2835</td>
</tr>
<tr>
<td>B₂_Eglin</td>
<td>-0.0313</td>
<td>0.0031</td>
<td>-0.0313</td>
<td>0.0032</td>
</tr>
<tr>
<td>CZ₁_F_W</td>
<td>0.0374</td>
<td>0.0102</td>
<td>0.0377</td>
<td>0.0098</td>
</tr>
<tr>
<td>CZ₂_NF_W</td>
<td>0.0345</td>
<td>0.0017</td>
<td>0.0346</td>
<td>0.0017</td>
</tr>
<tr>
<td>CZ₃_NF_D</td>
<td>0.0186</td>
<td>0.0355</td>
<td>0.0187</td>
<td>0.0345</td>
</tr>
<tr>
<td>PW_Fair</td>
<td>0.0010</td>
<td>0.2823</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PCI₀ Ln</td>
<td>-</td>
<td>-</td>
<td>-0.0624</td>
<td>0.2874</td>
</tr>
<tr>
<td>Adj. R² = 0.2219</td>
<td></td>
<td></td>
<td>Adj. R² = 0.2219</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.12: Summary of model results with absolute relative error as dependent variable (with only statistically significant parameters included)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Piecewise Coef</th>
<th>Piecewise P-Val</th>
<th>Natural Log Coef</th>
<th>Natural Log P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.1338</td>
<td>4.47E-12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PCIi0</td>
<td>-0.0015</td>
<td>2.03E-19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Horizon</td>
<td>0.0067</td>
<td>2.82E-06</td>
<td>0.0063</td>
<td>3.674E-05</td>
</tr>
<tr>
<td>Pav_Type</td>
<td>0.0439</td>
<td>4.13E-13</td>
<td>0.0511</td>
<td>1.46E-15</td>
</tr>
<tr>
<td>B2_Eglin</td>
<td>-0.0321</td>
<td>0.0024</td>
<td>-0.0298</td>
<td>0.0078</td>
</tr>
<tr>
<td>CZ1_F_W</td>
<td>0.0259</td>
<td>0.0071</td>
<td>0.0325</td>
<td>0.0015</td>
</tr>
<tr>
<td>CZ2_NF_W</td>
<td>0.0341</td>
<td>0.0019</td>
<td>0.0354</td>
<td>0.0024</td>
</tr>
<tr>
<td>CZ3_NF_D</td>
<td>0.0176</td>
<td>0.0454</td>
<td>0.0217</td>
<td>0.0233</td>
</tr>
</tbody>
</table>

Adj. R² = 0.2218 Adj. R² = 0.1210

To support the interpretation, a pavement section is defined to serve as a reference. With most variables being categorical in nature, it is convenient to have the reference section take on the reference value of the categorical variables. Doing so facilitates interpretation by having the default value of the categorical variables be set to 0 for the reference section. One can then interpret that a pavement section in a certain climate zone, in a certain location, or of a certain pavement type is expected to have more or less absolute and relative error in accordance with the parameter value for that categorical variable. Therefore, the reference section used for these interpretations is a rigid pavement section at Cannon Air Force Base located in climate zone 4 (freeze-dry climate). This leaves only Horizon and the PCIi0 related variables in the model to be considered. The median value of four years for Horizon is used to represent the reference section. The coefficient of the PCIi0 related variables are interpreted on the full range of possible values.
4.4.1 Interpretation of Piecewise Linear Model

Climate Variables: Recall, as discussed in section 3.2.2, climate zones 3 and 4 (represented by CZ3_NF_D and CZ4_F_D) have only one base each in the dataset and, as a result, the climate zone variable for the sections in these climate zones, as defined in section 3.2.2, effectively also represents location. Consequently, in these cases the climate zone and location variables are indistinguishable. Therefore, interpretation provided here regarding climate is restricted to climate zone 1 and 2 (represented by the variables CZ1_F_W and CZ2_NF_W) where two locations are represented in each.

Positive parameter values in Tables 4.11 and 4.12 indicate that, all else equal, the Freeze_Wet and No Freeze_Wet climates are associated with more absolute relative error in forecasts. Wet conditions introduce more complexity to the deterioration phenomenon whether in the presence of freeze-thaw or not. Therefore, it is reasonable to expect that predicting deterioration for these two dimensions is more prone to error.

Considering the climate variables in the relative error model, the signs of the corresponding parameters as seen in Tables 4.11 and 4.12 for both the piecewise linear and natural log models indicate that, all else equal, pavement sections in the freeze-wet climate zone (CZ1_F_W) are associated with more error in the pessimistic direction, while no freeze-wet climates (CZ2_NF_W) are associated with more error in the optimistic direction. This nuance is interesting. In regard to environmental characteristics, there does not appear to be a common trait that leads to an optimistic or pessimistic forecast. One possible explanation to this result is that while climactic factors do impact the amount of forecast error (absolute relative error), the pessimistic or
optimistic nature of the error may be more influenced by other location specific factors. Interestingly, considering the absolute value of the parameters, the freeze-wet climate zone (CZ1_F_W) is associated with larger forecasting errors than the no freeze-wet climate zone (CZ2_NF_W), which may be attributable to the more complex nature of deterioration in the presence of freeze-thaw cycles.

**Location Variables:** Specific sign or magnitude for the location parameters cannot be anticipated in an a priori manner given what is known about the bases. The absolute relative error model indicates that Eglin Air Force Base (one of the two in the no freeze-wet climate zone) is associated with more accurate forecasts in a statistically significant manner, all else equal. There could be a number of location specific factors that influence this result.

Regarding the relative error, Langley Air Force Base (one of the two in the freeze-wet climate zone) and Eglin Air Force Base are associated with statistically significant parameters indicating that the magnitude and the pessimistic or optimistic nature of the relative forecast error are associated with location. The signs of the two parameters are opposite. The forecasts at Langley are optimistic, all else equal, and at Eglin are pessimistic, all else equal. Moreover, considering the absolute value of the parameters, the errors are larger for Langley than for Eglin. However, no specific knowledge about the bases is available to offer any explanations or interpretations. Nevertheless, site specific factors influence their error, and could lead to forecasts that are either optimistic or pessimistic.
**Horizon:** A larger forecast horizon is associated with a larger amount of absolute relative error in the forecast, as indicated by a positive and statistically significant parameter values in Tables 4.11 and 4.12. This result is intuitive as the further into the future one attempts to forecast, the more unreliable that forecast will be. Negative parameter values in the relative error model seen in Table 4.10 indicate that increasing forecast horizon is associated with increasing error in the pessimistic direction. Considering the descriptive statistics of the dataset that the models are estimated on (see Table 4.1), the sample mean relative error of the forecasts is negative. It would make sense then, that if the existing forecast tendency is to produce pessimistic forecasts, increasing the forecast horizon would lead to an increasing amount of pessimism in that forecast.

**Pavement Type:** Positive parameter values in Tables 4.11 and 4.12 indicate that flexible pavement sections, all else equal, produce more absolute relative error in forecasting than do rigid sections. This result is not surprising since flexible pavement sections, especially on airfields, are difficult to evaluate and are susceptible to larger amounts of measurement error. As a result, the estimation of the rate of deterioration, ROD, used for forecasting is expected to reflect a higher degree of estimation error.

The fact that the parameter value is positive and significant in the relative error model indicates that this increase in error is on the optimistic side. That is, flexible pavement sections on Air Force airfields seem to be deteriorating at a faster rate than the Air Force is forecasting. This may be attributed to recent decreases in the operations and maintenance budgets available to base engineers to maintain these pavement sections. Forecasted deterioration is based on the historical rate at which similar pavement sections
have deteriorated over time. Preventative and routine maintenance has been applied to the sections over time and its effect is, therefore, captured in the forecasted deterioration rate. When available resources to perform such preventative and routine maintenance decrease, the maintenance effort reflected in the deterioration rate would be greater than the effort actually applied over the forecast horizon. As a result, pavement sections are expected to exhibit greater rates of deterioration than they did in the past. Therefore, the forecasts based on deterioration rates determined from data on sections that received more routine maintenance efforts are likely to reflect better condition levels than actually materializes under less routine maintenance efforts.

The fact that this increase in relative error is seen in flexible pavement sections but not in rigid pavement sections can be explained by the fact that flexible pavements are more maintenance intensive than rigid pavements. Rigid pavements have a greater initial cost at construction, but they require less routine maintenance and their expected life is greater. Flexible pavements are less expensive to construct but have shorter life expectancy and rely more heavily on sustained routine maintenance effort. It follows then, that a decrease in overall maintenance effort at an airfield with both flexible and rigid pavements would have a greater impact on the deterioration rates of flexible pavements.

**PCI\textsubscript{i0} related variables:** Based on parameter values in Tables 4.11 for the piecewise linear model, the magnitude of absolute error of the forecast for a section with PCI\textsubscript{i0}=1 decreases by 0.0023 for each additional point of PCI\textsubscript{i0} value. Once the section crosses the threshold of PCI\textsubscript{i0} = 55 into the “fair” category, the absolute relative error value continues
to decrease by 0.0023 for each $\text{PCI}_{i0}$ point due to the $\text{PCI}_{i0}$ parameter, and also increases by 0.0010 due to the PW_Fair parameter. The end result is that past the $\text{PCI}_{i0}=55$ threshold, the absolute relative error decreases by a net 0.0013 for each point of increase in $\text{PCI}_{i0}$ for the model estimation results shown in Table 4.11. In the case of the estimation results shown in Table 4.12, the piecewise linear model reduces to a purely linear model as the variable PW_Fair is removed due to the fact that its parameter is found to be insignificant as seen in Table 4.11.

These results indicate that the better condition the pavement section is in when the forecast is made, the more accurate the forecast will be, which makes intuitive sense. In general, pavement sections with higher PCI values are newer, and are likely to be on the low rate of deterioration portion of the typical degradation curve. Forecasts for such sections are more reliable as there is a low chance for these sections to enter the period of accelerated degradation during the forecast horizon. Pavement sections of low PCI value have a much higher potential to degrade quickly and nonlinearly and thus forecasts based on a constant rate of deterioration are prone to be unreliable.

Considering the parameters of the relative error model in Table 4.10, all other factors held constant, the relative error values on the low end of the $\text{PCI}_{i0}$ range are pessimistic and increase with increasing $\text{PCI}_{i0}$. This positive linear relationship continues, but is modified by the PW_Fair variable once $\text{PCI}_{i0}$ crosses into the “Fair” category, producing a smaller positive slope due to the negative PW_Fair parameter that has an absolute value less than the $\text{PCI}_{i0}$ parameter. Figure 4.3 shows the piecewise linear specification as a function of $\text{PCI}_{i0}$ for the reference section described in section
4.4. Notice, the increasing relative error eventually leads to an optimistic forecast. The switch from pessimistic to optimistic forecasts as $\text{PCI}_{i0}$ increases is also seen for sections other than the reference case of Figure 4.3. That is, pavement sections in poor condition are forecasted to deteriorate faster than what they actually do while pavements in good condition are forecasted to deteriorate slower than what they actually do.

![Rel_Error vs. PCI_{i0}](image)

Figure 4.3: Values of relative error for varying values of $\text{PCI}_{i0}$ in piecewise linear specification for a reference pavement section

This result is at first glance counter-intuitive considering the existing forecast methodology. Due to its reliance on a simple linear deterioration rate for each family, it would seem that pavement sections in poorer condition would be more likely to have optimistic forecasts. These lower $\text{PCI}_{i0}$ sections are likely in or nearing the stage in their lifecycle where deterioration occurs rapidly and at greater rates than pavement sections in better condition. The linearity assumed for the forecast methodology would lend itself to missing this marked increase in deterioration leading to forecasted PCI values greater
than actual PCI values. A possible reason for the relative error model estimation results not conforming to this expectation is that the pavement sections observed on Air Force airfields are not actually deteriorating in accordance with the natural deterioration curve depicted in the literature. Based on numerous pavement evaluation reports used to extract the data populating the dataset used in this study, it is evident that deterioration at most locations and in most families is occurring in a similar manner to the example shown in Figure 2.2. That figure does not demonstrate any trend where the deterioration follows a greater rate for lower PCI values. The linear trend line seems to be a good fit for the data points available.

A possible explanation for why this deterioration pattern is prevalent lies in the Air Force corporate operations, maintenance and repair policies. Prioritization of requirements and allocation of funds across the Air Force is driven by a “worst-first” strategy. Infrastructure in the worst condition gets priority. Since the maintenance and repair requirements always exceed available resources, the pavement sections in better condition do not receive much attention until they deteriorate to the point where they “make the cut” in terms of funds distribution. While in this study any pavement sections that underwent rehabilitation or reconstruction are removed from the dataset, routine maintenance is assumed to occur as part of normal operations. Given a limited annual budget, the tendency is to focus the bulk of the routine maintenance effort on the pavement sections in poorer condition to try and maintain them at an acceptable level, and put off routine maintenance on sections that are currently in better condition. When another evaluation is performed three or four years later, the poorer sections are likely to
outperform the forecast due to the amount of resources expended to keep them from deteriorating further, while better sections of pavement, which have been neglected, are likely to deteriorate at a faster rate than forecasted.

4.4.2 Interpretation of Natural Log Model

Climate, location, horizon and pavement type: All the parameters of these variables have the same sign and relative magnitude as those in the piecewise linear model, both in terms of absolute relative error and relative error. The two models behave similarly in regard to these variables and can be interpreted in a similar manner.

PCI<sub>i0</sub> related variables: The general relationship between PCI<sub>i0</sub> and relative error in the natural log specification is similar to that in the piecewise linear specification. In terms of absolute relative error, pavement sections with greater values of PCI<sub>i0</sub> exhibit less forecast error. In terms of relative error, to visualize the relationship with PCI<sub>i0</sub>, Figure 4.4 builds upon Figure 4.3 by including the natural log specification versus relative error through the full range of possible PCI<sub>i0</sub> values for the same reference section. For a given section with PCI<sub>i0</sub> = 1 the relative error in the natural log model is -1.24, as compared to the relative error value from the piecewise linear model given the same other explanatory variable values is -0.30. From this point, the natural log model is continually increasing but at a decelerating rate until it reaches a maximum relative error value of .06 for a section with PCI<sub>i0</sub> = 100.

The natural log function is markedly lower than the piecewise linear one over the lower end of the PCI<sub>i0</sub> scale as seen in Figure 4.4. Once PCI<sub>i0</sub> reaches approximately 25 to 30, the two models produce similar results. For pavement sections in the "Good"
category, the results are almost identical. For pavement sections that are in the "Fair" category or near the transition between the "Fair" and "Poor" categories occurring at PCI$_{i0}$ = 55, the relative errors associated with the models are slightly different due to the continuous nature of the natural log and the piecewise nature of the piecewise linear model. Since the PCI threshold of 55 separating the “Fair” and the “Poor” categories is arbitrary, the natural log model may produce better results in the transition period between “Fair” and “Poor”.

Figure 4.4: Values of relative error for varying values of PCI$_{i0}$ in both piecewise linear and natural log model specifications

4.5 Model Evaluation

The predictive power of the model specifications can now be tested using the test set, which was not included in estimating the models. Expected values of relative error for each of the 167 records in the test set are calculated using both the piecewise linear
and natural log models by applying the values of the explanatory variables corresponding to each record to the two models. The resulting relative error values (Rel_Error_i) are in turn applied to the original forecasted PCI values to produce corrected PCI forecasts as follows:

\[ \text{PCI}^C_{if} = (\text{Rel}_\text{Error}_i \times \text{PCI}_{if}) + \text{PCI}_{if} \]  

(4.1)

where,

\( \text{PCI}^C_{if} \) = corrected forecast,

\( \text{Rel}_\text{Error}_i = \) relative error calculated by applying the models to each record in the test set, and

\( \text{PCI}_{if} \) = original uncorrected forecast.

The corrected forecasts are then compared to the actual observations to establish the relative error of the corrected forecasts for each record based on equation 3.2 as follows:

\[ \text{Rel}_\text{Error}^C_i = \frac{\text{PCI}^C_{if} - \text{PCI}_{ia}}{\text{PCI}_{ia}} \]  

(4.2)

where,

\( \text{Rel}_\text{Error}^C_i \) = relative error of the corrected forecast, and

\( \text{PCI}_{ia} \) = the observed PCI of section i.

The relative errors of the corrected forecasts are analyzed and compared to the relative errors of the original uncorrected forecasts for each record to quantify the effectiveness of the models in terms of producing corrected PCI forecasts.

Table 4.13 shows a comparison of summary statistics for relative error of the 167 test sections when using the original uncorrected forecasts, corrected forecasts using the
piecewise linear model, and the corrected forecasts using the natural log model. The sample mean (average) relative error across the 167 records decreases when the forecast is corrected with either model specification. While the piecewise linear model changed the average relative error from positive to negative, in absolute terms it is closer to zero than the uncorrected forecasts. However, an evaluation based only on the average error could be misleading because the relative errors for individual records could be much larger in absolute value terms for the corrected forecasts than the original uncorrected forecasts.

Table 4.13: Summary statistics for relative error of test sections

<table>
<thead>
<tr>
<th></th>
<th>Uncorrected Rel_Error</th>
<th>Corrected Rel_Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Piecewise Linear</td>
</tr>
<tr>
<td>Mean</td>
<td>0.00508</td>
<td>-0.00035</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.10744</td>
<td>0.09736</td>
</tr>
<tr>
<td>Min</td>
<td>-0.34194</td>
<td>-0.32886</td>
</tr>
<tr>
<td>10th percentile</td>
<td>-0.08989</td>
<td>-0.08282</td>
</tr>
<tr>
<td>25th percentile</td>
<td>-0.05532</td>
<td>-0.05722</td>
</tr>
<tr>
<td>75th percentile</td>
<td>0.05309</td>
<td>0.02885</td>
</tr>
<tr>
<td>90th percentile</td>
<td>0.13325</td>
<td>0.11710</td>
</tr>
<tr>
<td>Max</td>
<td>0.37381</td>
<td>0.32361</td>
</tr>
<tr>
<td>Total Range</td>
<td>0.71575</td>
<td>0.65248</td>
</tr>
<tr>
<td>90th - 10th Range</td>
<td>0.22314</td>
<td>0.19992</td>
</tr>
<tr>
<td>75th - 25th Range</td>
<td>0.10841</td>
<td>0.08607</td>
</tr>
</tbody>
</table>

To address this limitation, the distributions of uncorrected and corrected relative errors across the 167 records are considered. The standard deviation of the corrected relative errors based on both models is less than that of uncorrected relative errors, and the ranges of values are also less for both models. These results indicate that applying the model corrections to the original forecasts narrows the distribution of the forecast
relative errors. Moreover, the most pessimistic and most optimistic corrected forecasts are closer to the true values than the corresponding original uncorrected forecasts.

To directly investigate the distributions of the relative errors, empirical cumulative distribution functions for the relative error of the uncorrected forecasts and the corrected forecasts based on the piecewise linear model across the 167 test set records are shown in Figure 4.5. Empirical cumulative distribution functions for the relative error of the uncorrected forecasts and the corrected forecasts based on the natural log model across the 167 test set records are shown in Figure 4.6. These figures provide a valuable visual confirmation of the conclusions discussed previously. Namely, for either model, the distribution of the relative error of the corrected forecasts is narrower than that of the original uncorrected forecasts. The sample mean of the relative errors are also shown in Figures 4.5 and 4.6 where one could see that the sample mean values of the relative errors of the corrected forecasts are closer to zero than the sample mean value of the relative errors of the uncorrected forecasts.
Figure 4.5: Empirical distribution functions and sample mean values of relative error of the original and corrected forecasts based on the piecewise linear model

Figure 4.6: Empirical distribution functions and sample mean values of relative error of the original and corrected forecasts based on the natural log model

In addition to comparing the standard deviation and range summary statistics and the empirical cumulative distribution functions, each record in the test set is analyzed on an individual basis to determine if the original forecast is improved by either of the model corrections. Table 4.14 shows the results of the corrected forecasts in terms of the percentage of sections that are closer to the true value than the original forecasts. Results of this analysis are shown in total as well as broken down into percentile based categories of the test set records. These results show that 53.29% of the total records exhibit improved forecast corrections based on both the piecewise linear and natural log models.

In addition, more records with higher values of relative error in the original forecasts exhibit improved forecasts. For example, 75.76% of the records with
uncorrected forecast relative errors falling in either the upper or lower 10th percentile exhibit improved corrected forecasts based on both models. Moreover, most of the records that do not exhibit improved corrected forecasts already have uncorrected forecasts close to the true values. For example, only 36.90% and 35.71% of the records with uncorrected forecast relative errors falling in the 25th to 75th percentile range exhibit improved corrected forecasts based on the piecewise linear and natural log models, respectively. These results are consistent with the summary statistics above which saw not only the average relative error moving closer to the truth, but the range of values across all sections narrowing.

Table 4.14: Percentages of records improved by corrected forecasts based on percentile ranges of uncorrected relative error

<table>
<thead>
<tr>
<th>Data sub-set</th>
<th># of Records</th>
<th>Piecewise Linear</th>
<th>Natural Log</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>167</td>
<td>53.29</td>
<td>53.29</td>
</tr>
<tr>
<td>Lower 10th percentile</td>
<td>16</td>
<td>87.50</td>
<td>87.50</td>
</tr>
<tr>
<td>Upper 10th percentile</td>
<td>17</td>
<td>64.71</td>
<td>64.71</td>
</tr>
<tr>
<td>Outer 20th percentile</td>
<td>33</td>
<td>75.76</td>
<td>75.76</td>
</tr>
<tr>
<td>Lower 25th percentile</td>
<td>42</td>
<td>66.67</td>
<td>69.05</td>
</tr>
<tr>
<td>Upper 25th percentile</td>
<td>41</td>
<td>73.17</td>
<td>73.17</td>
</tr>
<tr>
<td>Outer 50th percentile</td>
<td>83</td>
<td>69.88</td>
<td>71.08</td>
</tr>
<tr>
<td>Inner 50th percentile</td>
<td>84</td>
<td>36.90</td>
<td>35.71</td>
</tr>
</tbody>
</table>

Finally, the percentage of records with improved forecast corrections are analyzed in relation to the forecast horizon. The results are shown in Table 4.15. The forecast corrections are clearly more effective as the forecast horizon increases. With forecast horizon having a statistically significant parameter in the two relative error models, this result should not be surprising. Greater forecast horizons lead to an increase in the...
amount of forecast relative error, thus, correcting this error would be more effective as the horizon increases.

Table 4.15: Percentages of records improved by corrected forecasts based on forecast horizon

<table>
<thead>
<tr>
<th>Horizon</th>
<th># of Records</th>
<th>Piecewise</th>
<th>Natural Log</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>30</td>
<td>33.33</td>
<td>36.67</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>48.28</td>
<td>51.72</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>53.85</td>
<td>53.85</td>
</tr>
<tr>
<td>6</td>
<td>37</td>
<td>56.76</td>
<td>56.76</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>56.67</td>
<td>53.33</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>92.86</td>
<td>85.71</td>
</tr>
</tbody>
</table>

4.6 Summary of Estimation Results

Both of the model specifications that were interpreted and evaluated provide very similar signs, magnitudes, and statistical significance in the parameters. The model evaluation analysis shows that both specifications have value in correcting forecasts produced by the existing methodology implemented by PAVER. The natural log model demonstrated a slightly better overall fit, and the best average improvement in relative error when used to correct forecasts. While this model was estimated with a dataset which did not contain any records with values of PCI$_{i0}$ lower than 20, based on the form of the specification shown in Figure 4.4, records with PCI$_{i0}$ values close to 0 would likely result in unrealistic values of model relative error.

The piecewise linear model provides ease of interpretability due to its tie to the existing practice of considering three categories of condition in the field: good, fair, and poor. For Air Force practitioners who are used to referring to pavements in poor, fair and
good condition, as well as applying certain maintenance and repair, rehabilitation, and reconstruction actions based on these categories of condition, tying model interpretation and the use of the forecasts in maintenance decision making application to these condition categories is intuitive. It allows for a relatable model that is easily understood by an intended audience.
Chapter 5: Summary, Future Research, and Recommendations

5.1 Summary

In this study, airfield pavement condition forecasting error was investigated. Specifically, forecasting error models were developed and evaluated using data from US Air Force bases. The Air Force uses a curve fitting technique to establish the deterioration rates for families of airfield pavements. These rates are in turn used to forecast future pavement section condition to support maintenance, repair, rehabilitation and reconstruction decision making. The accuracy of this forecasting method had not previously been tested.

This study uses historical pavement evaluation data to compare forecasted pavement condition to observed condition at six Air Force airfields across the US and quantify the forecasting error of the current methodology. The error was then modeled using multiple linear regression with explanatory variables that are likely to impact pavement deterioration and therefore accuracy of forecasts. Four model specification alternatives were developed and estimated. Two of these specifications, the piecewise linear and natural log, were identified for further evaluation, interpreted, and evaluated using a hold-out sample from the original dataset.
The results showed that the model specifications were worthwhile due to the statistical significance of the parameters. While the adjusted $R^2$ values are relatively low for all specifications, this result is in large part due to an in-service dataset with a high degree of variability in the dependent variable. For each of the piecewise linear and natural log model specifications, eight of the ten parameters are statistically significant at the 99% confidence level (i.e., test size, $\alpha = 0.01$). One parameter is significant at the 95% confidence level (i.e., test size, $\alpha = 0.05$) and one parameter is significant at the 87% confidence level (i.e., test size, $\alpha = 0.13$). The statistical significance of the parameters corresponding to the explanatory variables indicates that factors such as construction material, pavement condition, climate, extent of the forecast horizon, and location contribute to the degree of systematic forecast error.

Interpretations for the piecewise linear and natural log model specifications are provided. The absolute relative error as the dependent variable was first considered to support the interpretation of the results of the relative error models. More specifically, absolute error models allow for the investigation of the magnitude of the error irrespective of whether it is optimistic (forecasted better condition than the true value) or pessimistic (forecasted worse condition than the true value). As expected, the forecast horizon showed a positive association with absolute error while the pavement condition at time of forecast showed a negative association with absolute error. Flexible pavement sections also showed greater degrees of absolute forecast error than rigid sections. The freeze-wet and no freeze-wet climates demonstrated higher degrees of absolute error, and
the location variables showed that the location of the pavement section had a significant association with the magnitude of forecast error.

When considering the results of the model using relative error as the dependent variable, some important nuances to the interpretation are noted. While forecast horizon continued to be a significant predictor of relative error, the greater the horizon, relative error increased in the negative direction indicating more pessimistic forecasts. One conclusion that can be drawn from this result is that the Air Force’s deterioration forecasts using the current methodology tend to be, on average, pessimistic. That is, more deterioration is forecasted than actually occurs. This result is attributed to the manner in which routine maintenance and repair decisions are made as discussed subsequently.

In the relative error interpretation, flexible pavement sections are found to increase the relative error. That is, the existing methodology forecasts less deterioration in flexible pavements than has actually occurred. One possible explanation for this result is that Air Force maintenance budgets have been decreasing in recent years. Fewer resources lead to less maintenance effort applied, and thus resulting in more deterioration. This trend occurs specifically in flexible pavement sections due to a greater need for routine maintenance and repair for such sections as compared to rigid pavement sections.

The categorical variables which represented location and climatic conditions showed statistically significant results in terms of both absolute relative error and relative error. Due to limitations in the data sample, specific interpretations regarding the no freeze-dry and freeze-dry climate zones are not possible due to the indistinguishability
between climate and location. Based on statistical significance of the parameters, however, the freeze-wet and no freeze-wet climate zones both showed greater amounts of forecast error in absolute terms, which is consistent with findings in prior research. Regarding the relative error, environmental phenomenon could not be linked to the optimistic or pessimistic nature of the forecast error. However, the larger magnitude of the absolute value of the parameter associated with the freeze-wet climate with respect to the absolute value of the parameter associated with the no freeze-wet climate may be attributed to the more complex nature of deterioration in the presence of freeze-thaw cycles. The parameters of the two location variables are statistically significant and have opposite signs. While it is not clear what location traits are behind these results, it is likely that site specific factors impact the rate of deterioration and the accuracy of the deterioration forecasts.

In both model specifications that are interpreted, the condition of the pavement at the time of forecast is a significant contributor to the forecast error. That is, considering the coefficient values for PCI$_{i0}$ and its transformations in both the piecewise linear and natural log models, when all other variables are held constant, relative error increases with PCI$_{i0}$. A reference pavement section was established to aid in the interpretation. For that reference section, lower values of PCI$_{i0}$ resulted in pessimistic forecasts with relative error increasing through the range of PCI$_{i0}$ with higher values resulting in optimistic forecasts.

Based on the accelerated deterioration seen in older pavements and the Air Force’s adoption of simple linear deterioration lines dependent on age that are fit to in-
service data, it would seem that the forecast for pavement sections with lower PCI values should tend to be optimistic. Take for instance a situation where all of the pavement sections in a particular deterioration family are in relatively good condition. If current pavement sections are all in the first 20 years of an expected 50 year lifespan, the PCI and age data will likely produce a linear deterioration model that fits the current observed PCI values well over the first 20 years on the age scale. However, if one applies this deterioration rate to forecast PCI over the life of the pavement section in this family, the projection will not capture the critical point in the deterioration curve when PCI starts to drop markedly at an increasing rate. This leads to an optimistic forecast, which in turn will likely result in suboptimal deferred maintenance and an increase in maintenance and repair costs.

However, the estimation results discussed previously contradict the described theoretical expectation. The explanation behind this apparent contradiction involves the maintenance and repair policies that are currently in place for the Air Force. The PCI value of a given pavement section is the primary factor that drives project scoring, which is used to evaluate competing requirements, prioritize effort, and determine which infrastructure requirements will receive attention in the presence of limited funds. In light of tight budget constraints, a “worst first” approach is in place where the greatest emphasis is placed on the assets that are in the worst condition. The deterioration that is being analyzed in this study is occurring in the presence of routine and preventative maintenance and repair actions only. Pavement sections that undergo rehabilitation or reconstruction are not considered when the deterioration rate is determined and,
therefore, are excluded from the analysis of this study. However, the worst first management mentality affects routine maintenance and repair decisions as well.

When a pavement evaluation is completed at a base, the base’s civil engineers use the resulting inventory to generate a “to-do” list of which pavements are in the worst condition, and which areas pose highest risks to the mission. Engineers then apply the bulk of the maintenance effort on the worst pavement sections and put off preventative maintenance on sections in good condition. The deterioration rates are determined by fitting a deterioration line to the observed data at the time. When the next evaluation is performed and the actual PCI values of sections are assessed and recorded again, it is reasonable to expect sections that were originally in the worst condition when previously forecasted, to outperform the forecast given the maintenance attention they have received. It is also reasonable to expect sections that were in the best condition when the forecast was made, to underperform in light of the deferred maintenance associated with such sections. That is, the majority of the maintenance effort has been put into the poor sections and thus they show better performance relative to the expected deterioration.

Two of the more promising relative error models developed and estimated in this study were evaluated using a test set of records not used in the estimation. The original historical forecasts and the corrected forecasts are compared to observed PCI values to evaluate the effectiveness of the error models and resulting corrections. The results show that applying either model specification to correct the original forecasts produced relative errors whose sample means are closer to zero. It was also shown that the range of the relative errors for the test set records is reduced by applying the model corrections,
essentially narrowing the distribution of the relative errors and centering it closer to zero. Furthermore, original forecasts that exhibit relatively high (positive) or low (negative) relative errors are improved the most by the corrections.

Both model specifications evaluated are meaningful in both the statistical significance of the parameters as well as their impact on improving the accuracy of forecasts. The natural log model has a slightly better overall fit and improvements in the corrected forecasts across the 167 test set records. The piecewise linear model is valuable regarding its interpretability and relatability to field practitioners given its reliance on the thresholds used in practice in categorizing condition.

5.2 Future Research and Practical Recommendations

This study is based on a dataset comprising pavement sections from six bases. It would be worthwhile to estimate and evaluate forecasting error models using a larger, more comprehensive dataset representing a larger number of bases; possibly all bases. Expanding on the scope of this study to include multiple bases in each climate zone would allow for meaningful interpretations to be made regarding the parameter values of all climate and location variables.

The results of this study indicate that there may be site specific factors that contribute to forecast error. Capturing explanatory variables that reflect such factors would be worthwhile. Work would need to be done to develop a way to quantify local factors such as annual budget levels, years of experience of the operations personnel, and the presence or absence of particular maintenance and repair policies and standard operating procedures. Once such variables are defined and quantified, they could be
explored in developing improved forecasting error models. In doing so, explanatory variables that capture interactions among pertinent variables could be explored.

This study focused on modeling relative error where the difference between an optimistic and pessimistic forecast is distinguishable, as those types of errors have different consequences in terms of maintenance and repair decision making. A model using absolute value of relative error was estimated as part of this study to support the interpretation of the relative error model estimation results. The absolute relative error model could be fully developed to model the magnitude of forecast error, regardless of whether the error is optimistic or pessimistic.

There is a desire within the Air Force engineering community to transition the current infrastructure investment approach to better capture the cost-benefit analysis of project alternatives and invest more resources in proactively keeping assets in good condition as opposed to waiting until they reach a poor enough condition to rise to the top of the priority list when they will cost more to repair. While this study explored the accuracy of airfield pavement condition forecasts, which form critical inputs to investment decision making, it did not directly consider the effect of forecast corrections on decision outcomes. A follow-on study could use the models developed in this study to correct PCI forecasts, and then quantify the effect of doing so on maintenance, repair, rehabilitation and reconstruction decision making. In turn, changes in the existing decision making approach could be explored where cost-benefit analysis and the total cost of ownership could be considered.
Currently a large database of the entire Air Force pavement inventory exists, but is only updated with the most current PCI observations. Past observations and corresponding deterioration rates can only be found in copies of the evaluation reports themselves. If the current database began to be populated with PCI observations that were dated, as well as the updated corresponding deterioration rates, in time, a comprehensive database would be available. Such a database could prove effective in supporting large, comprehensive empirical studies aimed at improving all aspects of pavement management, namely monitoring, deterioration modeling, condition forecasting, and investment decision making.
References


