Application of optimization and machine learning techniques in predicting pavement performance and performance-based pavement design

by

Parnian Ghasemi

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Civil Engineering Materials)

Program of Study Committee:
R. Christopher Williams, Major Professor
Vernon R. Schaefer
Derrick K. Rollins
Jeramy C. Ashlock
Azadeh Sheidaei

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

DEDICATION

To my family.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. R. Christopher Williams for his guidance, patience and support throughout this research and the writing of this thesis. I would like to thank my co-advisers Dr. Vernon R. Schaefer and Dr. Derrick K. Rollins and my committee members Dr. Jeramy C. Ashlock and Dr. Azadeh Sheidaei. Their insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I also would like to thank Dr. Wei Hong for his guidance and support. Many thanks to all my beloved friends, Sadaf, Raha, Zahra, Paul, Yuderka, Abhijit, and all of my colleagues in the ISU asphalt lab for their patience, help, support, and true friendship during my study at ISU. I would like to thank all my family without whom I would have never been able to succeed. I am thankful to my parents, Shahnaz and Morteza, and my sister, Kiana, who have always been great source of inspiration and support to me. Last but surely not least, I am deeply grateful to my beloved husband, Mohamad on whose shoulders I stand.

ABSTRACT

Maintenance, rehabilitation, and reconstruction of highway system are the major expenses in a state general expenditure. The emergence of predicting pavement performance and detecting the current state of the pavement health encourages pavement agencies to develop an accurate, efficient, and intelligent model to predict the remaining life of a pavement. Relating pavement condition, surface distresses, and structural properties, to a set of predictors including material properties, traffic loading, environmental factors, etc. through mathematical expressions is called performance modeling. To measure and predict pavement performance, a reproducible, authoritative, and field calibrated condition evaluating system is required. However, in the existence of numerous important predictors and their interrelationships, developing a predictive model for pavement performance is not a trivial task. The present study tackles the problem of developing pavement performance predictive model in two ways. First, a machine learning-based predictive framework is developed based on the laboratory produced performance data. The developed framework is implemented to predict the amount of permanent deformation in asphalt pavement as well as the asphalt pavement dynamic modulus. The developed framework is then used to solve a performance-based pavement design problem along with an evolutionary optimization algorithm. In the second approach, the structural behavior of a gantry crane way pavement at intermodal facilities is investigated by assessing the interactions between pavement, subgrade, and operational loading conditions. The pavement structural response to the crane load is measured through the installed strain gages in the field and used to validate a finite element-based model through an inverse analysis. The validated model is implemented to predict the fatigue life of the pavement structure as well as maintenance, rehabilitation and design recommendation for the existing and new crane way pavement sections.

CHAPTER 1. INTRODUCTION

This research has been motivated by the emergence of predicting pavement performance and detecting the current state of the pavement health. According to the American Association of State Highway and Transportation Officials (AASHTO), pavement performance is the pavement ability to sufficiently serve traffic over time. Maintenance, rehabilitation, and reconstruction of highway system are the major expenses in a state general expenditure Fakhri et al. (2017); Jalali et al. (2019). Therefore, pavement agencies are seeking to develop an accurate and efficient performance model to predict the remaining service life of a pavement. Relating pavement condition, surface distresses, and structural properties, to a set of predictors including material properties, traffic loading, environmental factors, etc. through mathematical expressions is called performance modeling Ghasemi et al. (2019b). To measure and predict pavement performance, a reproducible, authoritative, and field calibrated condition evaluating system is required Mirhosseini et al. (2019). The performance predictive model can be used to provide pavement maintenance, rehabilitation or reconstruction requirement. Several researchers tempted to develop pavement performance predictive models but almost all of the performance models are site specific and also limited to the materials used in the AASHO road test. Predictive modeling is a process that can use data mining tools as well as probability theory techniques to forecast outcomes from a given system, with each model constructed with several predictors likely to influence future results. Once data have been collected for relevant predictors, a statistical model is formulated that may use a simple linear equation, or may rather use a complex structure such as an Artificial Neural Network (ANN) obtained through sophisticated optimization algorithms Rahami et al. (2011a); Aslani and Asla (2010). ANN modeling has the same goal as statistical regression modeling because, both attempt to capture the relationship between a set of model inputs and corresponding outputs Devore (2011) by estimating

Figure 1.1: Rutting happening in an asphalt pavement under the wheelpath.

a set of coefficients that provide the best fit with the data. Among performance related properties of asphalt pavement, rut resistance and dynamic modulus are considered in this study.

1.1 Predicting rut resistance of asphalt pavement

One of the performance related properties of asphalt pavement is its resistance to rutting. Rutting is a term for when permanent deformation or consolidation accumulates in an asphalt pavement surface over time Notani et al. (2019); Ziari et al. (2016). It is typically shown by the wheel path being engraved in the road as indicated in Figure 1.1 Miller et al. (2014).

To analyze asphalt mixture rut susceptibility, performance testing along with mechanistic-empirical regression-based modeling appear to be a common approach Bashin et al. (2012). To simulate rutting in laboratory a rut resistance index called Flow Number (FN) is defined. In a repeated loading and unloading test FN is the point at which the strain rate starts to increase with loading. A schematic of the FN point is indicated in Figure 1.2. This parameter has demonstrated a strong correlation with rutting that happens in asphalt pavement due to traffic

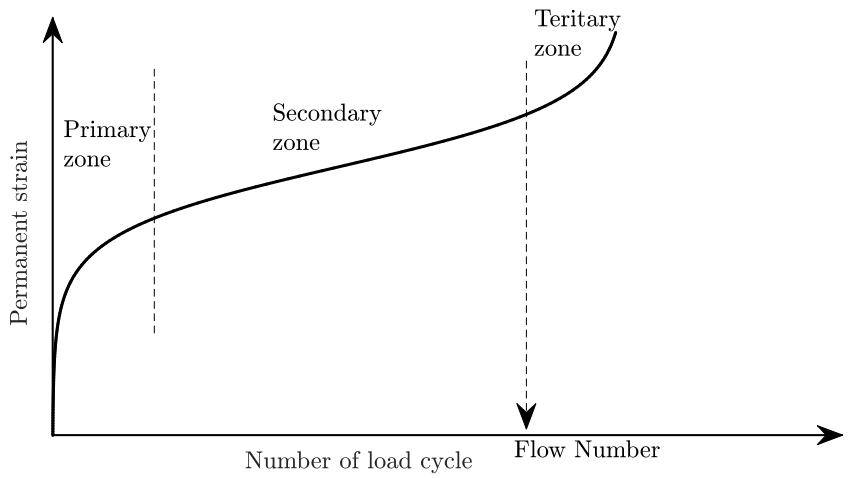

Figure 1.2: Flow number (FN) point.

in field Ghasemi et al. (2018b,a). Therefore, it is selected as the representative of rut behavior of asphalt pavement. In asphalt pavement design procedure, the amount of rutting should generally be limited to 0.4 inches (10.16 mm) regarding the total deformation of a pavement structure. It has been demonstrated that the amount of rutting is a function of binder viscosity, volumetric properties of asphalt mixture, and testing temperature Kaloush et al. (2003); Witczak (2002). Kvasnak et al.Kvasnak et al. (2007), proposed the following properties as the effective factors in rut susceptibility of asphalt mixture: nominal maximum aggregate size (NMAS), voids in mineral aggregate (VMA), percentage aggregate passing through sieve sizes No.4, No.16, No.200, binder grade, binder viscosity, asphalt content, testing temperature, and the number of gyrations. Rodezno et al.Rodezno et al. (2010) specified 12 parameters, i.e., testing temperature, maximum shear stress, normal stress, binder viscosity, percentage aggregate passing through sieve sizes 3/4-inch, 3/8-inch, and No.4, percentage air voids, effective binder content, binder content, VMA, and voids filled with asphalt (VFA) to be important in studying asphalt pavement rutting behavior. It is illustrated by Apeagyei, et al.Apeagyei (2011) that dynamic modulus test results at specific temperature and loading frequencies along with aggregate gradation appears to have strong correlation with FN test results. However, there are some discrepancies on the existence of correlation between rut

susceptibility of asphalt mixture and its dynamic modulus value Birgisson et al. (2004a); Pellinen and Witczak (2002); Timm et al. (2006).

1.2 Predicting Dynamic Modulus of asphalt mixture

Another widely used pavement performance characteristic is dynamic modulus, $|E^*|$, which defines stress-strain relationship of asphalt mixtures under sinusoidal loading. Dynamic modulus represents the stiffness characteristic of asphalt mixture and it has a significant role in Mechanistic-Empirical (M-E) pavement design. Therefore, several researchers have been attempted to predict asphalt mixture dynamic modulus as a function of material's components properties, loading rate, and temperature Ghasemi et al. (2016); Nobakht and Sakhaeifar (2018); Peng et al. (2019). Wiczak et al. developed a predictive model using material components properties including binder viscosity. Andrei et al modified the original Witczak model Andrei et al. (1999). The developed model has then been modified to use binder shear modulus instead of binder viscosity Bari and Witczak (2007). Christensen et al. Christensen Jr et al. (2003) created a predictive model based on the law of mixtures. Alkhateeb et al. Al-Khateeb et al. (2006) created a model from the law of mixtures to be used over a wide range of temperatures and loading frequencies. Sakhaeifar et al.Sakhaeifar et al. (2017) created separate temperature-based models that can predict dynamic modulus over a wide range of temperature. The predictor variables of the aforementioned models are selected from the following list: cumulative percentage aggregate retained on sieve sizes 3/4-inch, 3/8-inch, No.4, and percent aggregate passing the No.200, VMA, VFA, percentage air voids, effective binder content, binder shear modulus ($|G^*|$), and binder phase angle (δ). There are several issues with the existing predictive models for rut resistance and dynamic modulus of asphalt mixture:

- The selected input variables, material properties, are not necessarily accepted as a set of independent variables. Therefore, they might not be appropriate to be used in developing a predictive model. Pairwise correlated input variables can adversely affect the estimation accuracy of their effects on the output. To address this issue, pairwise correlation analysis is conducted in the present study. In case of having several correlated input variables a

multivariate statistical technique called Principal Component Analysis (PCA) is implemented to firstly reduced the dimensionality of the data set and secondly, to eliminate the correlation within the predictors. As a result, a set a orthogonal pseudo-variables is developed to be used in further model development.

- The developed models are not tested against an unseen, new, set of data. Therefore, they might be biased towards the dataset which is used in model development. Therefore, they might be overfitted against the training data. To address this issue, an appropriate cross validation technique is employed, and the developed models are tested against an unseen data set.

- The developed predictive models are based on empirical data. Therefore, they can be prone to extrapolation which is defined as the process of estimating beyond the original observation range. In case of extrapolation, predictive models are subject to major uncertainty and high risk of producing meaningless results. To guard against extrapolation, an n-dimensional hyperspace is found and added to the problem as an extra constraint.

- Due to the nonlinear nature of the pavement performance predictive models, none of the existing models can be implemented in developing a performance-based or optimized pavement design. To indicate an application of the developed predictive framework, a state-of-the-art evolutionary optimization algorithm called Mean-Variance Mapping Optimization, MVMO, is implemented to solve several real-life design related (performance-based design) optimization problems.

1.3 The unique performance predictive framework

Performance predictive models are developed based on a limited amount of empirical data. Therefore, the developed model will not necessarily be a good representative of all the produced asphalt mixtures. For example if a model is developed based on laboratory or field produced data of asphalt mixtures in the State of Arizona, it will not be a good fit for asphalt mixtures in

the State of Iowa. Therefore, for the past several years researchers have been trying to calibrate the existing models, i.e. rutting, fatigue, dynamic modulus, low temperature behavior, etc., and make them more suitable for their local pavements. A better idea will be developing a single performance framework for pavement which substitutes all of the existing models. The predictive framework can be trained based on the local condition and data without further calibration and will be implemented in Mechanistic-Empirical pavement design approach. Using the developed framework will be highly beneficial for asphalt agencies as well as Departments of Transportation (DOT) when a large amount of performance data is available.

1.4 Performance prediction Finite Element-based framework

Other than the aforementioned performance-based modeling approach, the problem of predicting and evaluating pavement performance can be tackled by studying pavement structural response to the applied load. This will result in a framework for decision-making using the information that accommodates and manages inevitable uncertainties. The structural behavior of a pavement can be studied by assessing the interactions between pavement, subgrade, and operational loading conditions. This can be done by installing strain gages in the pavement during or after construction. These strain gages will measure the pavement response to the load in the field and can be used to predict the fatigue life of the pavement structure as well as pavement maintenance and rehabilitation requirements and design recommendation for the existing and new pavement sections. The stain measurements can be used to verify a finite element-based model of pavement section to obtain pavement critical responses which are the maximum tensile stress at the bottom of rigid pavement, maximum compressive stress at top of base course and maximum compressive stress at top of subgrade soil. Finite element analysis is a powerful tool to capture pavement responses and have been used for the past decades to analyze rigid pavement structure Shoukry et al. (2007); Uddin et al. (1995); Sadeghi and Hesami (2018); Huang (1993). The methodology is not new to the pavement design and research community and has been used over the past decades. Traditional pavement-subgrade analysis based on static load and multilayers linear elastic formulation with

infinite dimensions in the horizontal plane and with a semi-infinite subgrade does not allow for dynamic behavior and pavement discontinuities. Besides, the finite element method allows for the dynamic analysis of pavements and the consideration of finite or infinite dimensions of the physical pavement structure. Several finite element programs have been developed exclusively for pavement analysis, e.g. ILLIPAVE and ILLISLAB for flexible and rigid pavement respectively. However, these packages are only capable of performing static analysis.

The finite element package ABAQUS is available for comprehensive structural pavement response analysis in static as well as dynamic procedures. Moreover, ABAQUS has a variety of material models including linear elastic, nonlinear elastic, viscoelastic etc. Surface to surface contact problems can also be defined and solve in this package. Therefore, in the present study ABAQUS is implemented to simulate the crane way pavement section under the crane load. However, the finite element analysis as the established method with enough efficiency and precision, can simulate the actual performance of the pavement only if the input information is precise enough. With having limited resources, it is not possible to measure every single property in the field. In these cases, the problem has some known as well as unknown properties (e.g. the only field measurement is strain values in concrete pavements). There are two major ways to tackle the problem of finding unknown properties of a system given some, and not all, experimental measurements. In the first approach, one can run the FEM simulation (can also be called the forward problem) of the system using some ranges of the given unknowns and then build a network using the batch of these simulations. Once the network is built given some error threshold one can use it inside an optimization solver to find the unknown properties directly. This approach has been used in a number of studies Ghasemi et al. (2018a). However, this approach is not applicable once we realize that our computational resources are limited because running FEM simulation of such a medium-size model is cumbersome. Another way to tackle this problem is using the inverse optimization approach.

This study is focused on development of numerical models for predicting pavement performance. The issues within the existing predictive models are addressed and a robust performance predictive framework is created based on laboratory and field produced performance data. Artificial neural

network as well as multivariate regression analysis are implemented in developing the proposed framework. The developed framework is not only capable of predicting pavement performance accurately but also it can be used along with a state-of-the-art evolutionary optimization algorithm to develop optimal performance-based pavement design.

1.5 Thesis outline

In chapter 2 rutting susceptibility of asphalt pavement is modeled using principal component pseudo inputs in regression and neural networks. Elimination of correlated inputs and extrapolation in modeling and optimization of permanent deformation is studied in Chapter 3. An n-dimensional hyperspace is defined, found and added as a constraint to the modeling problem. The problem of minimizing the amount of rutting is discussed and solved in this chapter. In Chapter 4 two predictive models for asphalt mixture dynamic modulus are developed and the authority of the developed models are examined. The developed models are used to find the optimal design, corresponding to maximum dynamic modulus value, as well as the design corresponding to a prescribed value of dynamic modulus. The general idea of using a single performance model for asphalt pavement is discussed in Chapter 5. Chapter 6 focuses on developing a finite element-based framework for predicting pavement performance. An Inverse algorithm is implemented to find the unknown material properties. Summary and discussion

CHAPTER 2. MODELING RUTTING SUSCEPTIBILITY OF ASPHALT PAVEMENT USING PRINCIPAL COMPONENT PSEUDO INPUTS IN REGRESSION AND NEURAL NETWORKS

A paper published at the International Journal of Pavement Research and Technology, Elsevier

2.1 Abstract

Permanent deformation is a major load-associated distress occurring in flexible pavement systems and increases with load repetitions affecting road roughness, serviceability, and the international roughness index (IRI). Early detection of rutting is necessary for maintenance and rehabilitation activities, but due to the complex behavior of asphalt mixtures, accurately predicting the permanent deformation of asphalt pavement is difficult. Historically, multivariate regression modeling and recently, artificial neural networks (ANNs) are used widely for material properties prediction. The ability to model accurately the response variable is adversely affected when inputs have pairwise correlations. To overcome this barrier, principal component analysis (PCA), as a dimensionality reduction technique, can be used to produce uncorrelated linear combinations of the original inputs as illustrated in this work using 83 (i.e., samples) laboratory compacted specimens from the State of Wisconsin. Asphalt binder, aggregate, and mix properties are obtained and used as the model inputs. The response parameter is the accumulated strain at the corresponding flow number. Using the developed pseudo inputs from PCA, a multivariate regression and an ANN model are generated and were able to fit the test cases with (r_{fit}) of 0.8 and 0.97 respectively. The developed machine learning-based framework is shown to be a capable tool in estimating the rutting behavior of asphalt mixture.

2.2 Introduction

Permanent deformation (also known as rutting) is one of the most common flexible pavement's distresses affecting road roughness, serviceability, and international roughness index (IRI). Rutting in asphalt mixtures usually occurs in wheel-paths and appears as the longitudinal depressions with small upheavals to the side. This differential consolidation in the pavement profile can cause safety issues Sousa et al. (1991). Early detection of rutting is necessary for maintenance and rehabilitation activities, but due to the complex behavior of asphalt mixtures, accurately predicting the permanent deformation of asphalt pavement is difficult. To determine the amount of permanent deformation, different modeling approaches can be used including empirical, mechanistic-empirical, and mechanistic where the goal is to estimate future performance based on the laboratory test data and the observed distress history of pavement. Mechanistic-empirical, regression-based modeling and performance testing approaches are prevalent in asphalt mixture's rut susceptibility analysis Sousa et al. (1991). Recently, more researchers have concentrated on viscoelastic, viscoplastic, and viscoelastoplastic continuum damage-based modeling to explain the rutting behavior of asphalt mixtures. These models have some limitations including high dependency on the empirical data and requiring accurate characterization of asphalt behavior Bashin et al. (2012). Although permanent deformation of hot mix asphalt (HMA) depends on stiffness of the mixture, deformation cannot be estimated from the stiffness characteristic alone. Many researchers have demonstrated that in order to determine the rutting performance of HMA mixtures, permanent deformation characteristics should be measured directly Zhang et al. (2015). Due to this limitation, researchers have attempted to simulate rutting by using a rutting resistance indicator parameter, entitled flow number (FN), defined as the point where the permanent strain rate reaches a minimum value. This parameter can be measured by a repeated loading and unloading test AASHTO TP 79-13 (2013). The FN has indicated a good correlation with field rutting of asphalt mixtures exposed to different levels of traffic Witczak (2007).

The width, path, and severity of the rutting profile depend on the pavement structure, loading, and environmental conditions. During the design procedure, there is generally a limiting criteria

of 0.4 inches (10.16 mm) with the total deformation of the pavement structure and its impacts on the lateral and longitudinal surface profiles. Many researchers have demonstrated that the amount of rutting depends on the mixture volumetric properties, binder viscosity, and testing temperature Kaloush et al. (2003). Asphalt mixture properties, which affect rutting behavior (simulated by FN), were identified more precisely by Kvasnak et al. Kvasnak et al. (2007). They demonstrated that binder grade, binder viscosity, asphalt content, testing temperature, nominal maximum aggregate size (NMAS), voids in mineral aggregate (VMA), percentage aggregate passing from sieve sizes No.4, No.16, No.200, and number of gyrations affect the FN of asphalt mixture. Rodezno et al. Rodezno et al. (2010) represented 12 parameters affecting rutting behavior of asphalt mixtures in the laboratory including testing temperature, maximum shear stress, normal stress, binder viscosity, percentage aggregate passing from sieve sizes 3/4-in, 3/8-in, and No.4, air voids, effective binder content, binder content, VMA, and voids filled with asphalt (VFA). Although, there are some disagreements on the existence of correlation between dynamic modulus of asphalt mixture and its rutting behavior Birgisson et al. (2004a); Pellinen and Witczak (2002); Brown et al. (2002); Timm et al. (2006), Apeagyei Apeagyei (2011) represented that using dynamic modulus test results at specific test temperature and loading frequencies in conjunction with aggregate gradation shows a good correlation with FN test results. According to the existing literatures, the parameters affecting the rutting behavior of asphalt mixtures can be classified into three categories including asphalt properties representing the viscoelastic and viscoplastic behavior, aggregate properties representing the elastic/plastic behavior, and mixture properties.

Early detection of rutting, required for punctual maintenance and rehabilitation activities, provides motivation for the designers to predict the rutting behavior of the asphalt mixtures. Historically multivariate regression modeling and recently, pattern recognition techniques are used widely for material properties prediction Cheng and Titterington (1994b). In the conventional material modeling process, regression analysis is an important tool for building a model. In linear regression analysis, several procedures have been developed for parameter estimation. These methods differ in computational simplicity of algorithms, presence of a closed-form solution, robustness with respect

to heavy-tailed distribution, and theoretical assumptions needed to validate desirable statistical properties. Among these methods, least-square estimation is the simplest and the most common technique. It minimizes the sum of squared residuals, and leads to a closed-form expression for the estimated value of the unknown parameter Devore (2011). Pattern recognition techniques can learn and recognize trends in data contributing to their current widespread use. These techniques learn the pattern from experimental data and design the computational models. One such approach, Artificial Neural Network (ANN), is an interconnected network of many simple processors as shown in Figure 2.1. All ANNs consist of a set of processing units or neurons classified as input, hidden and output neurons. Input neurons receive input from external sources and transfer it to the rest of the network. Hidden neurons receive input and transmit their computed output to the processing units within the network without any outside contact. Output neurons receive the input from the rest of the network that it transforms and sends to external receivers Kartam (1994).

Although ANN can be trained to approximate a non-linear, complicated relationship Kartam, Nabil. Flood (1994); Rahami et al. (2011a); Aslani and Asla (2010); Saltan and Sezgin (2007), similar to multivariate linear regression model and other modeling tools, their ability to accurately predict the response variable highly depends on the quality and properties of input variables May et al. (2011). Cross-correlated inputs adversely affect accurate estimation of their causative effects on the response variable and this impedes the ability of the model to accurately estimate the response variable Rollins et al. (2015). Thus, a pre-processing step is needed to examine the quality and relationship of input variables – a step not commonly practiced by design engineers in this application.

In the presence of correlated input variables, orthogonal variables can be obtained using a dimensionality reduction technique called principal component analysis (PCA). PCA is a multivariate statistical procedure that uses an orthogonal transformation to convert a set of correlated variables into a set of uncorrelated variables called principal components (PCs). The PCs are a set of orthogonal, linear combinations of the original variables within the dataset Jolliffe (2002).

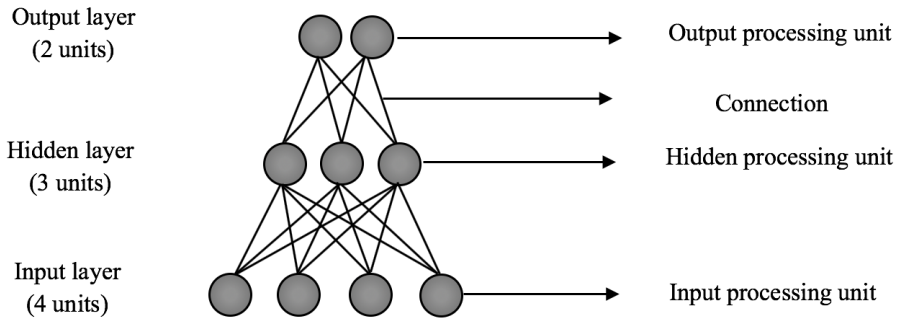

Figure 2.1: Example of an architecture for ANN.

The present study focuses on developing a machine learning-based framework to reduce a large set of correlated input variables to a set of uncorrelated input variables to model the accumulated strain of asphalt mixtures using ANN and multivariate regression structures.

The rest of the paper is organized in the following way. Experimental materials and methods are presented in Section 2.3. Section 2.4 covers the pre-processing step for input variables evaluation. PCA is described in detail in Section 2.5. The proposed modeling methods, PCR and PCNN, are presented in Sections 2.6, followed by results, discussion, and conclusions in Sections 2.9 and 2.10 respectively.

2.3 Experimental materials and methods

2.3.1 Materials Sampling and Collection

The materials used in the present study were sampled at the plant site directly from the back of trucks after they had been loaded out, in accordance with ASTM Standard D979 and D3665. In addition to the mix, the asphalt binder was also sampled for each pavement section. Maximum theoretical specific gravity (G_{mm}) was measured in accordance with AASHTO T209/ASTM D2041 for two 1250 g split samples for each job. The G_{mm} was used to determine the volumetric properties of the specimens. Eighty-three specimens from 21 different HMA mixtures collected from different projects in the State of Wisconsin, were compacted in the laboratory using a Pine AFGC125X Superpave Gyratory Compactor (SGC) that can produce specimens in the dimensions of roughly

150 mm in diameter by 170 mm in height. Specimens were compacted to 4.0%, 7.0%, and 10.0% air voids. The bulk specific gravity was determined in accordance with AASHTO T166/ASTM D2726.

2.3.2 Dynamic Modulus Testing

A 100-mm diameter by 150 mm high cylindrical specimen was cored, trimmed, and prepared for the dynamic modulus test. The specimens were tested under a repeated sinusoidal compressive stress at an effective test temperature of 36.6 °C and at four loading frequencies including 25, 10, 1, and 0.1 Hz in unconfined conditions. The effective test temperature for all of the laboratory tests (36.6 °C) was selected based on climatic condition of the Midwestern parts of the United States, and was considered the temperature at which permanent deformation would occur, which is equivalent to a seasonal correction throughout the year. A Universal Testing Machine (UTM 100) was used to conduct the testing with a temperature controlled testing chamber. In accordance with AASHTO TP 79-13 the test was conducted from higher to lower frequencies to mitigate the amount of deformation that is induced upon specimens.

2.3.3 Flow Number Testing

After conducting the dynamic modulus tests, the same specimens were used for performing the FN test under a repeated haversine compressive stress at a single effective temperature. The UTM 100 machine was used to perform the tests, with a temperature-controlled testing chamber. The load was applied for a duration of 0.1 second and a dwell period of 0.9 second. No confining pressure was used and the axial stress is the deviator stress (600 kPa). The FN test is conducted at the effective test temperature of 36.6 °C. The reason why the accumulated strain at the FN is selected as the response variable, and not the FN itself, is that the FN is just an indicator parameter of rutting resistance, and to relate laboratory data to the AASHTO design procedure, which uses strain and not FN as the rutting criteria.

2.3.4 Complex Shear Modulus Testing

To obtain binder shear properties, complex shear modulus test was conducted at the same effective test temperature of 36.6 °C and same frequencies including 25, 10, 1, and 0.1 Hz. The test was conducted in accordance with ASTM D755-09.

2.4 Pre-processing step: input variables selection strategy

Selecting input variables is a fundamental and crucial task in identifying the optimal functional form of statistical models. Accurate modeling of the output requires a set of input variables sufficiently high in information content that maps to the output space. The difficulty of selecting a parsimonious set of input variables arises due to the following reasons: 1. the number of available variables can be very large; 2. high correlations exist between input variables, and; 3. variables that are unknown to be weakly related or unrelated to the response Fodor (2002).

Mathematical modeling is the process of mathematically relating measured input variables to output variables. The modeler selects a mathematical structure and a process for estimating unknown model parameters. For a general model structure, let its expectation be represented as $\eta_i = f(\mathbf{x}_i, \theta)$, where η_i is the expected value of the response (i.e., output) at the i^{th} sampling time, $i = 1, \ldots, n$; $\mathbf{x}_i$ is the vector of input values at the i^{th} sampling time; and θ is the vector of unknown model parameters with $\theta = [\theta_1, \ldots, \theta_q]^T$ Let the element of its Jacobian Matrix, $\mathbf{J}^{n \times q}$, in the i^{th} row and j^{th} column be $\frac{\partial \eta_i}{\partial \theta_j}$ i.e. $\mathbf{J} = \frac{\partial \eta_i}{\partial \theta_j}$, Note that, the j^{th} column represents θ_j and its column vector represents the change in the response space as θ_j changes for the set of experimental conditions. If two columns such as j and k are orthogonal, then their correlation coefficient is zero. More specifically, if these two columns are orthogonal, the information to estimate θ_j is decoupled (i.e., separate or independent) from the information to estimate θ_k, and vice versa. The advantage of this attribute is that it strengthens causative relationships of inputs on the output and maximizes parameter accuracy, and thus, estimation accuracy for the modeled output. Correlated columns in the Jacobian Matrix arise from pairwise correlation of inputs. Thus, for a given set of experimental data, to minimize standard error and maximize the accurate mapping of input behavior into the

Table 2.1: Original Input Variables

Variable	Identity	Values in the database				Selection based on	
		Min.	Max.	Ave.	Std. Dev.	literature	Current research
x_1	Binder %	3.4	6.6	5.093	0.773	*	
x_2	G*	210801	1163560	612180	265903.652		*
x_3	NMAS	12.5	25	15.922	3.761	*	
x_4	Passing 3/4”	81.3	100	98.554	4.085	*	
x_5	Passing 1/2”	38.3	98.8	87.13	15.199	*	
x_6	Passing 3/8”	34.1	89.9	76.34	15.099	*	
x_7	Passing #4	26.2	72.5	56.248	13.741	*	
x_8	Passing #8	17.5	54	42.249	10.512		*
x_9	Passing #16	14.2	47.4	32.178	8.74	*	
x_{10}	Passing #30	9.6	39.1	23.196	7.004		*
x_{11}	Passing #50	5.7	18.6	12.022	3.179		*
x_{12}	Passing #100	3.7	9.8	6.187	1.424		*
x_{13}	Passing #200	2.8	8.5	4.322	1.115	*	
x_{14}	VMA	10.323	21	16.452	2.502	*	
x_{15}	VFA	46.45	91.719	65.189	9.062	*	
x_{16}	Va%	1.019	9.825	5.868	2.088	*	
x_{17}	E*	395.7	2299.4	869.41	411.524		*

response space, this work seeks to minimize pairwise correlation of the inputs that are used to model the response behavior.

According to the literature Kaloush et al. (2003); Rodezno et al. (2010), the rutting behavior of an asphalt mixture can be explained by its components' properties . The properties and their ranges used in the present study are indicated by Table 2.1. As previously mentioned, the binder properties describe the viscous behavior and the aggregate properties describe the elastic behavior. The complex shear modulus of asphalt binder is selected to describe the shear relaxation behavior and the dynamic modulus of asphalt mixture is selected as the demonstration of material stiffness used in mechanistic-empirical pavement design.

The seventeen aforementioned properties are expressed as input variables to predict the accumulated strain value at the FN. Using JMP statistical software, the pairwise correlation matrix of input variables is calculated and presented by Table 2.2. Results with absolute value of 0.5 and higher are in bold and red text.

Table 2.2: Correlation Matrix for the Input Variables

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	1	0.33	-0.73	0.6	0.71	0.63	0.47	0.45	0.46	0.45	0.41	0.43	0.42	0.61	-0.18	0.35	-0.43
x_2	0.33	1	-0.11	0.08	-0.03	-0.23	-0.38	-0.32	-0.18	-0.06	0.01	0.23	0.35	0.02	-0.02	0.02	0.24
x_3	-0.73	-0.11	1	-0.64	-0.76	-0.69	-0.51	-0.41	-0.39	-0.32	-0.25	-0.30	-0.39	-0.53	0.13	-0.30	0.37
x_4	0.6	0.08	-0.64	1	0.88	0.78	0.58	0.51	0.42	0.32	0.25	0.23	0.24	0.46	-0.21	0.32	-0.45
x_5	0.71	-0.03	-0.76	0.88	1	0.95	0.77	0.71	0.61	0.49	0.41	0.39	0.36	0.56	-0.33	0.44	-0.51
x_6	0.63	-0.23	-0.69	0.78	0.95	1	0.92	0.86	0.75	0.61	0.49	0.30	0.16	0.53	-0.27	0.38	-0.49
x_7	0.47	-0.38	-0.51	0.58	0.77	0.92	1	0.95	0.84	0.69	0.49	0.11	-0.14	0.43	-0.16	0.26	-0.37
x_8	0.45	-0.32	-0.41	0.51	0.71	0.86	0.95	1	0.95	0.83	0.59	0.11	-0.15	0.42	-0.16	0.26	-0.41
x_9	0.46	-0.18	-0.39	0.42	0.61	0.75	0.84	0.95	1	0.96	0.74	0.18	-0.11	0.40	-0.16	0.25	-0.41
x_{10}	0.45	-0.06	-0.32	0.32	0.49	0.61	0.69	0.83	0.96	1	0.84	0.28	-0.06	0.35	-0.15	0.22	-0.37
x_{11}	0.41	0.01	-0.25	0.25	0.43	0.49	0.49	0.59	0.74	0.84	1	0.59	0.17	0.21	-0.20	0.20	-0.25
x_{12}	0.43	0.23	-0.30	0.23	0.39	0.30	0.11	0.11	0.18	0.28	0.59	1	0.82	0.24	-0.28	0.28	-0.15
x_{13}	0.42	0.35	-0.39	0.24	0.36	0.16	-0.14	-0.15	-0.11	-0.06	0.17	0.82	1	0.29	-0.30	0.32	-0.22
x_{14}	0.61	0.02	-0.53	0.46	0.56	0.53	0.43	0.42	0.40	0.35	0.21	0.24	0.29	1	-0.62	0.83	-0.71
x_{15}	-0.18	-0.02	0.13	-0.21	-0.33	-0.27	-0.16	-0.16	-0.16	-0.15	-0.20	-0.28	-0.30	-0.62	1	-0.94	0.52
x_{16}	0.35	0.02	-0.30	0.32	0.44	0.38	0.26	0.26	0.25	0.22	0.20	0.28	0.32	0.83	-0.94	1	-0.63
x_{17}	-0.43	0.24	0.37	-0.45	-0.51	-0.49	-0.37	-0.41	-0.41	-0.37	-0.25	-0.15	-0.22	-0.71	0.52	-0.63	1

As it can be seen in the table, the absolute value of 273 results are above 0.1 which shows that most of the inputs are highly correlated with 41 of them greater than 0.5 indicating that several inputs are highly correlated. Therefore, to predict the response variable accurately, we used principal component analysis (PCA) to obtain a much smaller set of pseudo variables that are uncorrelated Fodor (2002).

2.5 Principal component analysis (PCA)

Mathematically, PCA is defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate (or the first principal component), the second greatest variance on the second coordinate, and so on. It can be considered as fitting an n-dimensional ellipsoid to the data, where each axis of the ellipsoid represents a principal component. If some axis of the ellipsoid is small, then the variance along that axis is also small, and by omitting that axis

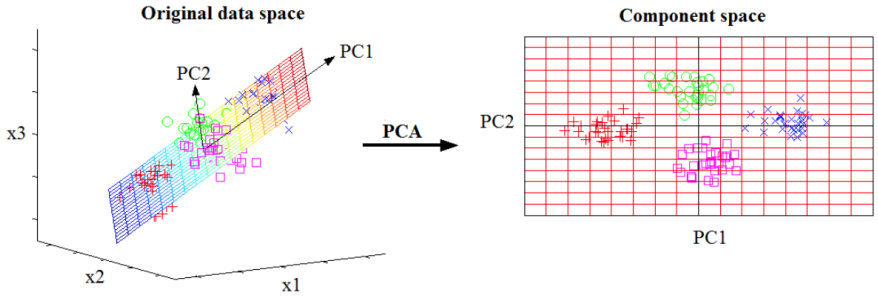

Figure 2.2: Schematic of the PCA transformation.

and its corresponding principal component from our representation of the data set, we lose only a commensurately small amount of information. Often, this operation can be thought of as revealing the internal structure of the data in a way that best explains the variance within the data. If a multivariate dataset is visualized as a set of coordinates in a high-dimensional data space (1 axis per variable), PCA can supply the user with a lower- dimensional picture, a projection of this object when viewed from its most informative viewpoint Jolliffe (2002). A schematic of this transformation for three inputs is presented in Figure 2.2. This is done using only the first few principal components so that the dimensionality of the transformed data is reduced.

Considering a data set $\mathbf{X}$, PCA reduces the dimension of $\mathbf{X}$ by expressing the p (p =17) original variables $(x_1, \ldots, x_p)$ as d new pseudo-variables (principal components, PCs), where $d < p$. The PCs are a set of orthogonal (i.e., uncorrelated), linear combinations of the original variables within the dataset. The obtained PCs can be used for multiple purposes including: 1. to construct a new set of variables that are linear combinations of the original variables and that contain exactly the same information as the original variables; 2. to identify patterns of multicollinearity in a data set and use the results to address the collinearity problem in multiple linear regression; 3. to identify variables or factors, underlying the original variables, which are responsible for the variation in the data; 4. to find out the effective number of dimensions over which the data set exhibits variation, with the purpose of reducing the number of dimensions of the problem and; 5. to create a few orthogonal variables that contain most of the information in the data and that simplify the identification of groupings in the observations Johnson and Wichern (1992). PCA

can be considered as a data pre-processing methodology that determines an optimal rotational transformation of the dataset, $\boldsymbol{X}$, and maximizes the amount of variance of the output γ that is explained by the PCs Fodor (2002).

Considering the given dataset $\boldsymbol{X}$, PCA is performed in the following steps:

1. Standardizing $\mathbf{X}$ by transforming it to $\mathbf{Z}$ using the following equations:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{\mathrm{np}} \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 & \ldots & \mathbf{z}_p \end{bmatrix} = \begin{bmatrix} \frac{x_{11}-\overline{x}_1}{s_1} & \frac{x_{12}-\overline{x}_2}{s_1} & \cdots & \frac{x_{1p}-\overline{x}_p}{s_p} \\ \frac{x_{21}-\overline{x}_1}{s_1} & \frac{x_{22}-\overline{x}_2}{s_2} & \cdots & \frac{x_{2p}-\overline{x}_p}{s_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x_{n1}-\overline{x}_1}{s_1} & \frac{x_{n2}-\overline{x}_2}{s_2} & \cdots & \frac{x_{np}-\overline{x}_p}{s_p} \end{bmatrix} \tag{2.1}$$

   where, for $i = 1\ to\ n$ and $j = 1\ to\ p$, x_{ij} is the i^{th} measurement for the j^{th} variable, $\overline{x}_i$ is sample mean for the ith variable, and s_i is sample standard deviation for the i^{th} variable.

2. Determine the unit eigenvectors $e_1, \ldots, e_p$ of $\mathbf{Z}$.

3. Determine the corresponding eigenvalues $\lambda_1, \ldots, \lambda_p$.

4. Rank the eigenvectors according to their eigenvalues.

5. Select the d PCs according to their eigenvalues (or the scree plot).

Selection of the PCs is based on examining the eigenvalues of each PC, which correspond to the amount of variance explained by each PC, and thereby including only the significant PCs as input features. A common selection method is to rank the PCs and select all PCs whose eigenvalues exceed some threshold, k, to ensure that selected components explain the desired amount of variance of γ. Another selection method is to generate and use a scree plot of the percentage contribution of each k $^{\text{th}}$ PC and to visually identify an optimal value of k. Therefore, the first PC contains the most variance possible to be captured in a single axis. The second PC is orthogonal to the first one (their correlation is zero) and contains as much of the remaining variance as possible. The third PC is orthogonal to all previous ones and also contains the most variance possible, etc. Rollins et al. (2006).

Using the JMP statistical software package, the eigenvectors and eigenvalues of the correlation matrix are calculated and presented in Table 2.3 and Table 2.4, respectively. According to the eigenvalues, 89.72% of the variation in the original data is explained by the first 5 PCs. The scree plot, which is a graph of the eigenvalues versus their order, can also be used as a visual inspector of identifying critical PCs. The scree plot is presented in Figure 2.3. The "elbow" point, and the location of this breaking point, indicates the number of critical PCs to be selected. The presented graph illustrates that there are 4 critical PCs.

Based on the PCA results, the first 5 PCs were selected to create pseudo input variables. As mentioned previously, these PCs are a linear combination of the original input variables described by Eq. 2.2

$$pc_i = \sum_{j=1}^{17} \alpha_{ij} x_j + \beta_i \tag{2.2}$$

where, α is the corresponding coefficients, β is constant, and X_j's are the original input variables. Eq. 2.2 is presented in matrix notation by Eq. 2.3

$$PC = AZ^T + B \tag{2.3}$$

where,

Table 2.3: Eigenvectors of the **Z** matrix.

	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}	PC_{11}	PC_{12}	PC_{13}	PC_{14}	PC_{15}	PC_{16}	PC_{17}
z_1	0.26	-0.16	0.23	-0.15	0.29	-0.28	0.14	0.27	0.48	-0.49	0.29	-0.13	-0.04	0.03	0.01	0.05	0.01
z_2	-0.03	-0.31	0.35	0.02	0.67	0.26	-0.15	0.08	-0.26	-0.03	-0.40	0.07	-0.04	0.01	0.02	-0.03	0.03
z_3	-0.25	0.12	-0.17	0.35	-0.02	0.21	-0.39	0.70	0.13	-0.07	0.10	0.17	0.10	0.10	-0.01	0.02	0.04
z_4	0.26	-0.06	0.08	-0.38	-0.03	0.28	-0.57	-0.04	0.18	0.36	0.24	-0.31	-0.20	0.10	-0.03	-0.02	0.03
z_5	0.32	-0.04	0.08	-0.24	-0.14	0.21	-0.12	0.00	-0.05	-0.12	0.06	0.48	0.42	-0.25	0.01	0.13	-0.49
z_6	0.33	0.11	0.02	-0.18	-0.17	0.17	0.05	0.07	-0.01	-0.12	-0.18	0.18	0.24	-0.02	0.03	-0.10	0.79
z_7	0.29	0.28	-0.04	-0.10	-0.10	0.17	0.23	0.20	0.01	-0.06	-0.41	-0.15	-0.08	0.62	-0.09	0.08	-0.29
z_8	0.29	0.30	-0.03	0.04	0.01	0.08	0.09	0.25	-0.22	-0.07	0.01	-0.04	-0.46	-0.44	0.15	-0.49	-0.14
z_9	0.29	0.28	0.04	0.20	0.15	-0.04	0.01	0.03	-0.31	0.05	0.20	-0.02	-0.18	-0.13	-0.09	0.75	0.13
z_{10}	0.26	0.24	0.11	0.35	0.22	-0.11	-0.07	-0.11	-0.20	0.09	0.28	-0.28	0.56	0.19	0.01	-0.33	-0.04
z_{11}	0.22	0.11	0.23	0.49	-0.02	-0.01	-0.15	-0.35	0.43	0.08	-0.09	0.45	-0.28	0.14	-0.01	-0.10	0.01
z_{12}	0.15	-0.29	0.34	0.34	-0.39	0.03	0.06	0.19	0.15	0.16	-0.31	-0.42	0.15	-0.33	0.04	0.13	-0.05
z_{13}	0.10	-0.46	0.26	0.08	-0.35	-0.11	0.04	0.15	-0.47	-0.04	0.32	0.19	-0.21	0.36	-0.04	-0.12	0.04
z_{14}	0.25	-0.22	-0.29	-0.03	0.22	-0.24	0.26	0.28	0.12	0.61	-0.02	0.22	0.05	-0.05	-0.33	-0.08	0.01
z_{15}	-0.16	0.29	0.41	-0.25	-0.01	-0.41	-0.03	0.21	0.00	0.35	-0.08	0.16	0.05	0.08	0.53	0.06	0.00
z_{16}	0.20	-0.30	-0.41	0.15	0.10	0.20	0.14	-0.06	0.05	0.06	0.08	-0.01	-0.01	0.11	0.75	0.11	0.00
z_{17}	-0.23	0.12	0.33	-0.01	0.05	0.58	0.53	0.04	0.15	0.22	0.37	0.04	0.02	0.00	-0.02	-0.01	0.02

Table 2.4: Eigenvalues from the Z Matrix

Number	Eigenvalue	Percent Variance	Cumulative Percent
1	7.979	46.935	46.935
2	2.876	16.920	63.856
3	1.947	11.456	75.311
4	1.539	9.052	84.363
5	0.912	5.363	89.726
6	0.581	3.419	93.145
7	0.421	2.475	95.620
8	0.260	1.529	97.149
9	0.198	1.165	98.314
10	0.152	0.892	99.206
11	0.063	0.371	99.577
12	0.038	0.224	99.801
13	0.016	0.095	99.896
14	0.008	0.049	99.945
15	0.005	0.028	99.973
16	0.003	0.016	99.989
17	0.002	0.011	100.000

$$A^T = \begin{bmatrix} 3.41 \times 10^{-1} & -2.04 \times 10^{-1} & 2.89 \times 10^{-1} & -1.94 \times 10^{-1} & 3.71 \times 10^{-1} \\ -7.6 \times 10^{-8} & -1.17 \times 10^{-6} & 1.31 \times 10^{-6} & 1.05 \times 10^{-7} & 2.47 \times 10^{-6} \\ -6.59 \times 10^{-2} & 3.22 \times 10^{-2} & -4.53 \times 10^{-2} & 9.1 \times 10^{-2} & -7.97 \times 10^{-3} \\ 6.34 \times 10^{-2} & -1.31 \times 10^{-2} & 1.86 \times 10^{-2} & -9.38 \times 10^{-2} & -7.13 \times 10^{-3} \\ 2.11 \times 10^{-2} & -2.33 \times 10^{-3} & 5.26 \times 10^{-3} & -1.63 \times 10^{-2} & -9.03 \times 10^{-3} \\ 2.17 \times 10^{-2} & 7.23 \times 10^{-3} & 1.55 \times 10^{-3} & -1.23 \times 10^{-2} & -1.13 \times 10^{-2} \\ 2.13 \times 10^{-2} & 2.01 \times 10^{-2} & -2.82 \times 10^{-3} & -7.27 \times 10^{-3} & -7.11 \times 10^{-3} \\ 2.79 \times 10^{-2} & 2.89 \times 10^{-2} & -2.98 \times 10^{-3} & 3.79 \times 10^{-3} & 1.41 \times 10^{-3} \\ 3.28 \times 10^{-2} & 3.22 \times 10^{-2} & 4.43 \times 10^{-3} & 2.35 \times 10^{-2} & 1.71 \times 10^{-2} \\ 3.73 \times 10^{-2} & 3.39 \times 10^{-2} & 1.53 \times 10^{-2} & 4.98 \times 10^{-2} & 3.14 \times 10^{-2} \\ 7.08 \times 10^{-2} & 3.3 \times 10^{-2} & 7.38 \times 10^{-2} & 1.54 \times 10^{-1} & -9.05 \times 10^{-3} \\ 1.08 \times 10^{-1} & -2.03 \times 10^{-1} & 2.33 \times 10^{-1} & 2.3 \times 10^{-1} & -2.72 \times 10^{-1} \\ 9.18 \times 10^{-2} & -4.06 \times 10^{-1} & 2.29 \times 10^{-1} & 6.69 \times 10^{-2} & -3.18 \times 10^{-1} \\ 9.98 \times 10^{-2} & -8.77 \times 10^{-2} & -1.12 \times 10^{-1} & -1.14 \times 10^{-2} & 8.85 \times 10^{-2} \\ -1.73 \times 10^{-2} & 3.21 \times 10^{-2} & 4.57 \times 10^{-2} & -2.69E \times 10^{-2} & -9.43 \times 10^{-4} \\ 9.75 \times 10^{-2} & -1.42 \times 10^{-1} & -1.95 \times 10^{-1} & 7.31 \times 10^{-2} & 4.70 \times 10^{-2} \\ -5.50 \times 10^{-4} & 2.73 \times 10^{-4} & 8.2 \times 10^{-4} & -8.92 \times 10^{-6} & 1.36 \times 10^{-4} \end{bmatrix}$$

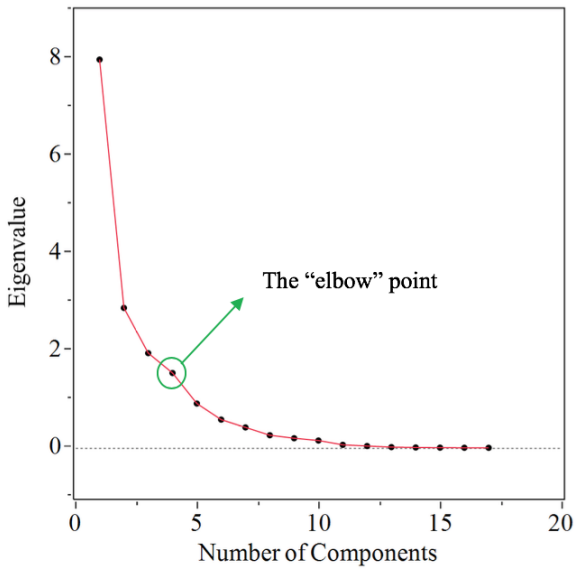

Figure 2.3: Scree plot.

$$B = \begin{bmatrix} -1.72E+01 \\ 5.87E-01 \\ -8.21E+00 \\ 7.36E+00 \\ -4.85E-01 \end{bmatrix}$$

The obtained PCs will be used as new inputs for further modeling efforts.

2.6 Proposed modeling approaches

2.6.1 *K*-fold cross validation

In order to recognize how the results of the statistical analysis will generalize to an independent data set, and to prevent overfitting, a k-fold cross-validation technique is used. Cross validation is a procedure to guard against overfitting models by checking the fitted model against a set of data that was not used in fitting the model Refaeilzadeh et al. (2009). In k-fold cross-validation, the dataset is randomly partitioned into k equal-sized subsamples. When using only a training

set and test set of the k subsamples, a single subsample is retained and the set for testing the model, and the remaining k-1 subsamples are used as training data, e.g., to fit the model. The cross-validation process is then repeated k times (the folds), with each of the k subsamples used exactly once as the test data. The advantage of this method over repeated random subsampling is that all observations are used for both training and testing, and each observation is used for testing exactly once. Based on the size of the data set (83), 3-folds with 24 sample vectors in each of them are randomly selected, and 3-fold cross-validation is done.

2.7 Principal Component Regression (PCR)

Using the first five PCs as input variables and the accumulated strain as the response variable, multiple regression modeling consisting of the second-order quadratic and interaction terms was fit and retrained. Least squares criterion of minimizing the sum of squared residuals (SSE) is used for both linear regression and ANN modeling. For the training fold, it minimizes the sum of squared residuals and develops a closed-form expression for the estimated value of the unknown parameter. Full third-order models and reduced third order models were also developed and fitted. However, these models were rejected because of they did not significantly improve fit to the training data and gave worse fit to the test data.

2.8 Principal Component Neural Network (PCNN)

The proposed Principal Component Neural Network (PCNN) approach was used to develop predictive models of the response that included the pseudo variables as inputs. More specifically, a three-layer feedforward neural networks consisting of an input layer of 5 neurons, a hidden layer of 10 neurons, and an output layer of one neuron was developed using the MATLAB program. Selection of the number of neurons in the hidden layer is based on a trial-and-error procedure between optimization of the cost function and computational time. A four-layer network with two hidden layers was also developed but due to the principal of parsimony, the simplest and most economical way (in terms of computational time) has been selected.

Table 2.5: Linear Regression Coefficients for Equation 2.4

Coefficient	Value
c_0	1.64×10^4
c_1	-8.89×10^2
c_2	-1.24×10^3
c_3	-1.24×10^3
c_4	-92.41
c_5	6.55×10^2
c_6	1.58×10^2
c_7	-4.35×10^2
c_8	2.35×10^2
c_9	3.74×10^2
c_{10}	-8.28×10^2

For each fold, the training process starts with adjusting the initial values of the network's weights and biases in order to obtain a reasonable output and continues modifying the network by minimizing SSE. The iteration continues until the convergence criterion is met (Note that this is the reason a validation set is not used. That is, just as for the linear regression model, training stops when convergence is obtained). The Bayesian Regularization algorithm is implemented for the training efficiency of the network.

2.9 Results and discussion

2.9.1 Results

Results of the proposed methods are presented in this section and their accuracy in predicting the rutting behavior of asphalt mixtures is evaluated and compared. The second-order quadratic linear regression model fit the measured response, y, best and is given by Eq. 2.4 as $\hat{y}$

$$\begin{aligned}\hat{y} = c_0 + c_1 \times pc_1 + c_2 \times pc_2 + c_3 \times pc_3 + c_4 \times pc_4 + c_5 \times pc_5 + c_6 \times pc_1 \times pc_2 \\ + c_7 \times pc_2 \times pc_4 + c_8 \times pc_1 \times pc_3 + c_9 \times pc_2 \times pc_3 + c_{10} \times pc_3 \times pc_5\end{aligned} \tag{2.4}$$

where, c_i's for $i = 0, \ldots, 10$ are presented in Table 2.5.

The successfully trained ANN models can be presented by Eq. 2.5 for ease of use and wider reproduction. Each ANN is presented by the connection weights and biases in a three-layer topology Juang and Chen (1999a).

$$\hat{y} = f_2 \left\{ B_0 + \sum_{j=1}^{n} \left[W_j \cdot f_1 \left(B_{Hj} + \sum_{i=1}^{m} W_{ij} P_i \right) \right] \right\} \tag{2.5}$$

where, B_0 = bias at output layer (just one neuron at this layer); W_j = weight of connection between neuron j of the hidden layer and output layer neuron; B_{Hj} = bias at neuron j of the hidden layer (for $j = 1$ to 10); W_{ij} = weight of connection between input variable i (for $i = 1$ to 5) and neuron j of the hidden layer; P_i = input parameter i; $f_1(t)$ = transfer function of the hidden layer, and $f_2(t)$ = transfer function of the output layer.

Both transfer functions $f_1(t)$ and $f_2(t)$ are sigmoid functions defined in Eq. 2.6 Juang and Chen (1999a).

$$f_k(t) = \frac{1}{1 + e^{-t}} \quad for \quad k = 1, 2. \tag{2.6}$$

The connection weights and biases of the ANN are presented by the following matrixes.

$$W_{\mathrm{ij}} = \begin{bmatrix} -0.447 & 1.702 & -0.811 & -0.854 & -1.295 \\ 0.010 & -1.462 & 0.280 & 0.114 & -1.926 \\ 1.315 & -1.391 & -0.028 & -0.218 & 0.895 \\ -0.089 & 0.182 & -0.135 & -0.323 & -0.150 \\ 0.089 & -0.182 & 0.135 & 0.324 & 0.150 \\ -0.613 & 0.255 & -1.190 & -0.808 & 0.120 \\ 0.089 & -0.182 & 0.135 & 0.324 & 0.150 \\ 0.443 & -1.496 & -0.414 & 1.051 & -0.641 \\ 0.280 & -0.625 & 0.831 & 0.736 & 1.158 \\ -0.007 & 0.168 & 0.003 & -1.380 & -0.292 \end{bmatrix}$$

$$W_j = \begin{bmatrix} 0.673 \\ 0.348 \\ -1.250 \\ -0.294 \\ 0.294 \\ -1.713 \\ 0.294 \\ -0.262 \\ 0.974 \\ 0.986 \end{bmatrix}, \; B_{\mathrm{Hj}} = \begin{bmatrix} -0.084 \\ -1.208 \\ -1.742 \\ -1.528 \\ 0.567 \\ -0.567 \\ -1.320 \\ -0.567 \\ 1.464 \\ -1.750 \end{bmatrix}, \; B_{\mathrm{Hj}} = [1.334]$$

Performance results of the PCR and PCNN models are given based on the following statistics, and presented in Table 2.6. The first statistic is the "average error (AE)" defined as

$$AE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i). \tag{2.7}$$

AE is an estimate of systematic model bias, n is the number of input vectors, y_i is the ith measured response value, and $\widehat{y_i}$ is the ith fitted response value. The second statistic is the "average absolute error (AAE)" and defined as

$$AAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{2.8}$$

This statistic gives the average closeness of the fitted value to the measured response value. The third statistic, r_{fit}, is the correlation of y_i and $\widehat{y_i}$ and defined as

$$r_{fit} = \frac{n\sum_{i=1}^{n} y_i\hat{y}_i - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} \hat{y}_i)}{\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}\sqrt{n\sum_{i=1}^{n} \hat{y}_i^2 - (\sum_{i=1}^{n} \hat{y}_i)^2}} \tag{2.9}$$

The better the fit, the higher r_{fit}will be with a maximum possible value of 1. The last statistic is R^2 or the coefficient of determination. In linear regression, for training, R^2 is the portion of the variation explained by the fitted model. It is only applicable to the PCR since it is linear in parameters but not to PCNN it is non-linear in parameters.

Table 2.6: Statistical Analysis of PCR and PCNN Modeling (na*: not applicable)

		PCR			PCNN		
	Statistics	Fold 1	Fold 2	Fold 3	Fold 1	Fold 2	Fold 3
Training	AD	0	0	0	34.99	-242.99	46.19
	AAD	1497.02	1705.55	1514.59	729.41	1350.87	944.83
	r_{fit}	0.83	0.82	0.85	0.96	0.87	0.94
	R^2	0.69	0.68	0.72	na*	na*	na*
Testing	AD	626.73	-129.91	-226.1	-98.24	149.2	-169.6
	AAD	2007.47	1515.74	2110.64	694.53	719.9	1037.28
	r_{fit}	0.79	0.8	0.73	0.97	0.95	0.92
	R^2	na*	na*	na*	na*	na*	na*

According to the values of r_{fit} presented in Table 2.6, the predicted values of accumulated strain by PCR and PCNN models have a high correlations with the measured ones which means both PCR and PCNN modeled the response well. The second fold of PCR has the highest PCR r_{fit} which is 0.8. The first fold of the PCNN has the highest PCNN r_{fit} of 0.97 which is the highest compare to other folds in both methods.

A phenomenological model is given by Eq. 2.10 below:

$$\frac{\epsilon_p}{\epsilon_r} = aT^bN^c \tag{2.10}$$

where, ϵ_p = accumulated plastic strain at N repetition of load; ϵ_r = resilient strain of the asphalt material as a function of mix properties, temperature, and time rate of loading; N = number of load repetitions; T = pavement temperature, and a, b, and c are unknown model coefficients. Although many researchers including Leahy and Ayres tried to obtain suitable estimates for the unknown parameters by performing repeated load permanent deformation tests, their models were able to the accumulated strain with R^2 of not higher than 0.76 with temperature being the most important variable and loading conditions, material type, and mix parameters being less important ones

Witczak (2007). In comparison with the literature, the overall performance of both the PCR and PCNN models which is expressed in terms of the three test statistics used in the work is significantly higher than previous prediction models used in the AASHTO design procedure Witczak (2007).

Comparing the two best folds for training and testing stages indicates that although the PCR modeling works well in predicting the response variable, PCNN has the best results in both training and testing.

2.9.2 Model Validation

Equation 2.11 is a general regression model for this study:

$$y_i = f(\mathbf{Z}_i, \theta) + \epsilon_i \tag{2.11}$$

where$f$ is the expectation function, θ is the vector of parameters and ϵ_i is random error term assumed to be normally distributed with mean zero and unknown variance σ^2 for $i = 1, \ldots, n$, where n is the number of input vectors. Violation of these assumptions and model adequacy can be investigated by the examination of residuals, defined by Eq. 2.12

$$e_i = y_i - \hat{y}_i \tag{2.12}$$

Through the analysis of residuals, many types of model inadequacies and violations of the underlying assumptions can be assessed. If the model is adequate, the residuals should contain no obvious pattern. Checking the normality assumption can be done by constructing a normal probability plot of the residuals. If the underlying error distribution is normal, this plot will resemble a straight line Devore (2011). These assumptions were checked for the best fold in each method. The plot of residuals for the best folds are presented in Figure 2.4 and 2.5. Since there is no obvious pattern in the residual plot, the assumption of equal variances seems acceptable. The normal probability plots of the residuals for both models are presented in Figure 2.6 and 2.7. The data points are not too far away from a straight line. Therefore, the normality assumption does not appear to be violated.

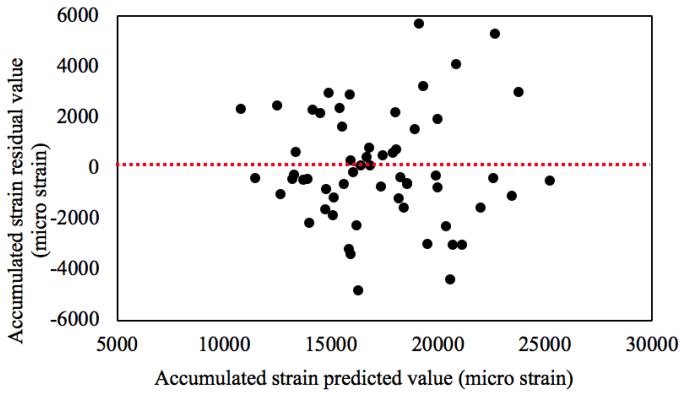

Figure 2.4: Plot of the residuals for best fold of PCR model.

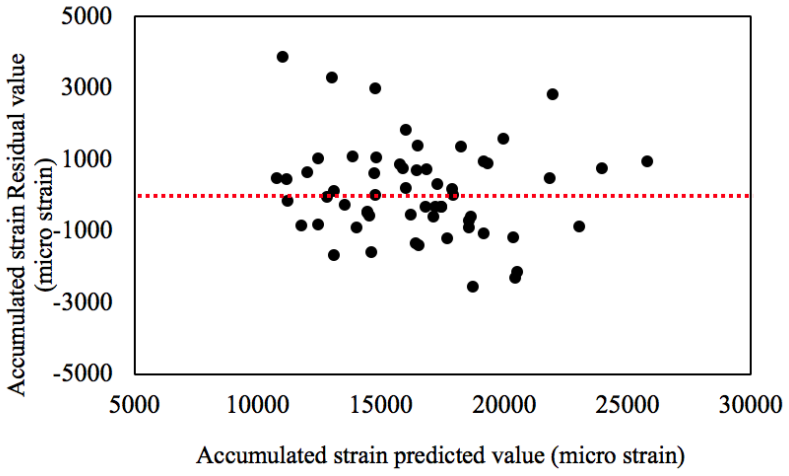

Figure 2.5: Plot of the residuals for best fold of PCNN model.

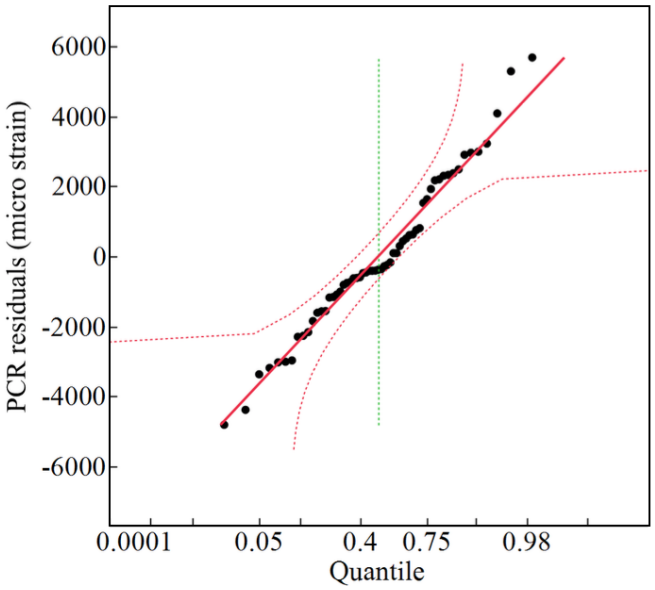

Figure 2.6: Normal probability plot of the residuals for best fold of PCR model.

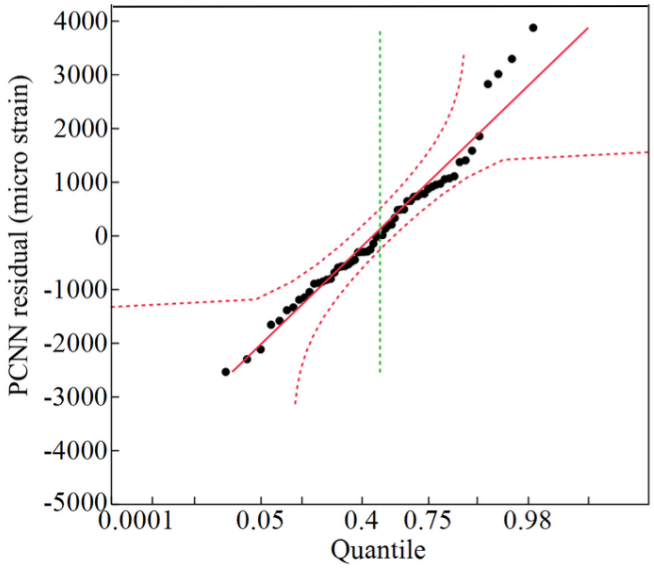

Figure 2.7: Normal probability plot of the residuals for best fold of PCNN model.

2.10 Conclusions and recommendations

This study used the experimental data of the permanent deformation of asphalt mixture and focused on the evaluation and qualifying the input variables to be used in further modeling. A total number of 17 input variables from three categories of material properties including binder, aggregate and mixture properties were selected as the effective parameters in rutting behavior. Cross -correlated input variables were identified by correlation analysis and substituted by the orthogonal pseudo-inputs (PCs) using a dimensionality reduction technique called PCA. This work compared multiple regression and ANN modeling of a small set of pseudo-variables determined from PCA (PCR and PCNN, respectively). Both proposed approaches modeled the amount of permanent deformation well with PCNN fitting the test data significantly better. Nonetheless, both approaches showed better performance as a modeling tool than other regression-based approaches that use standard variables in the AASHTO design procedure. Thus, these PCA approaches are strongly recommended as sound modeling approaches in this application. Moreover, theses methodologies appear to also have much promise in modeling other material properties at every effective temperature and this investigation is recommended in future studies. Another future study to consider is development of an approach to determine the importance of each inputs on the response. Considering the obtained regression model (Eq. 2.4), the linear terms contain orthogonal variables (i.e. PCs) which are in the same normalized scale. Therefore, the linear terms with the largest coefficients (with the absolute value) are the ones with greatest contribution. In order to map this principal component that contributes the most to the original input space and finding the input with the greatest contribution, one can consider the coefficients for each input value in Table 2.4. The one with the largest absolute value is associated variable contributes the most to that pseudo-variable.

2.11 References

AASHTO TP 79-13 (2013). AASHTO TP 79-13 Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Aphalt Mixtures Performance Tester. *Am. Assoc. State Highw. Transp. Off.*, pages 1–19.

Apeagyei, A. K. (2011). Rutting as a Function of Dynamic Modulus and Gradation. *Am. Soc. Civ. Eng.*, 23 (9):1302–1310.

Aslani, M. and Asla, R. (2010). A Novel Hybrid Simplex-Genetic Algorithm For The Optimum Design Of Truss Structures. *Lect. Notes Comput. Sci.*, II.

Bashin, A., Masad, E., Kutay, M. E., Buttlar, W., Kim, Y.-R., Marasteanu, M., Kim, Y. R., Schwartz, C. W., and Carvalho, R. L. (2012). Applications of advanced models to understand behavior and performance of asphalt mixtures. *Transp. Res. E-Circ.*, 1(E-C161).

Birgisson, B., Roque, R., Kim, J., and Pham, L. V. (2004). The Use of Complex Modulus To Characterize Performance of Asphalt Mixtures and Pavements in Florida. Technical Report September 2004, Florida Department of Transportation, Florida Department of Transportation.

Brown, E., Cooley, L. A., Hanson, D., Lynn, C., Powell, B., Prowell, B., and Watson, D. (2002). Ncat test track design, construction, and performance.

Cheng, B. and Titterington, D. M. (1994). Neural networks: A review from a statistical perspective. *Statistical science*, pages 2–30.

Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences*. Cengage learning.

Fodor, I. K. (2002). A Survey of Dimension Reduction Techniques. Technical report, U.S. Department of Energy, U.S. Department of Energy.

Johnson, R. and Wichern, D. (1992). Applied multivariate statistical analysis prentice hall new jersey 641.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.*, 30(3):487.

Juang, C. H. and Chen, C. J. (1999). CPT-Based Liquefaction Evaluation Using Artificial Neural Networks. *Comput. -Aided Civ. Infrastruct. Eng.*, 14(3):221–229.

Kaloush, K. E., Witczak, M., and Sullivan, B. W. (2003). Simple performance test for permanent deformation evaluation of asphalt mixtures. *Sixth Int. RILEM Symp. Perform. Test. Eval. Bitum. Mater.*, pages 498–505.

Kartam, N. (1994). Neural Netwroks in Civil Engineering: Systems and Application. *J. Comput. Civ. Eng.*, 8(2):149–162.

Kartam, Nabil. Flood, I. (1994). Neural Netwroks in Civil Engineering: Principal and Understanding. *J. Comput. Civ. Eng.*, 8(2):131–148.

Kvasnak, A., Robinette, C., and Williams, R. (2007). A Statistical Development of a Flow Number Predictive Equation for the Mechanistic-Empirical Pavement Design Guide.CD-ROM . In Transportation Research Board, Washington, D., editor, *Transportation Research Board, Washington, DC.*, volume 3, pages 1–18, Transportation Research Board, Washington, DC. DC.Transportation Research Board, Washington.

May, R., Dandy, G., and Maier, H. (2011). Review of input variable selection methods for artificial neural networks. In *Artificial neural networks-methodological advances and biomedical applications*. InTech.

Pellinen, T. and Witczak, M. (2002). Use of Stiffness of Hot-Mix Asphalt as a Simple Performance Test. *Transp. Res. Rec. : J. Transp. Res. Board*, 1789(02):80–90.

Rahami, H., Kaveh, A., Aslani, M., and Asl, R. N. (2011). A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization. *Int. J. Optim. Civ. Eng.*, 1(January):29–46.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer.

Rodezno, M., Kaloush, K., and Corrigan, M. (2010). Development of a Flow Number Predictive Model. *Transp. Res. Rec. : J. Transp. Res. Board*, 2181(2181):79–87.

Rollins, D. K., Roggendorf, A., Khor, Y., Mei, Y., Lee, P., and Loveland, S. (2015). Dynamic Modeling With Correlated Inputs: Theory, Method, and Experimental Demonstration. *Industrial & Engineering Chemistry Research*, 54(7):2136–2144.

Rollins, D. K., Zhai, D., Joe, A. L., Guidarelli, J. W., Murarka, A., and Gonzalez, R. (2006). A novel data mining method to identify assay-specific signatures in functional genomic studies. *BMC bioinformatics*, 7:377.

Saltan, M. and Sezgin, H. (2007). Hybrid neural network and finite element modeling of sub-base layer material properties in flexible pavements. *Mater. Des.*, 28(5):1725–1730.

Sousa, J. B., Craus, J., and Monismith, C. L. (1991). Summary report on permanent deformation in asphalt concrete. Technical report.

Timm, D., West, R., Priest, A., Powell, B., Selvaraj, I., Zhang, J., and Brown, R. (2006). Phase ii ncat test track results. *NCAT Report*, 6(05).

Witczak, M. W. (2007). *Specification criteria for simple performance tests for rutting*, volume 1. Transportation Research Board.

Zhang, J., Yang, J., and Kim, Y. R. (2015). Characterization of mechanical behavior of asphalt mixtures under partial triaxial compression test. *Constr. Build. Mater.*, 79:136–144.

CHAPTER 3. PRINCIPAL COMPONENT ANALYSIS-BASED PREDICTIVE MODELING AND OPTIMIZATION OF PERMANENT DEFORMATION IN ASPHALT PAVEMENT: ELIMINATION OF CORRELATED INPUTS AND EXTRAPOLATION IN MODELING

A paper published at the Journal of Structural and Multidisciplinary Optimization, Springer

3.1 Abstract

Permanent deformation in asphalt pavement is a function of material properties, loading, environmental conditions, and structural design (e.g., thickness of pavement layers). Because of the large number of effective variables and their nonlinear interrelationships it is not easy to develop a predictive model for permanent deformation. In this study, a laboratory database containing accumulated strain values (output) and material properties (inputs) from several asphalt pavements has been used to develop a predictive model for permanent deformation. We first show that the inputs are highly correlated, then Principal Component Analysis (PCA) is used to compute a set of orthogonal pseudo-inputs. Two predictive models based on the pseudo-inputs developed using linear regression analysis and Artificial Neural Networks (ANN), and are compared using statistical analysis. Extrapolation using empirical predictive models is highly risky and discouraged by experienced practitioners, so to guard against extrapolation, a method is developed to determine an input hyper-space. The above-developed model, along with an n-dimensional hyper-space, provides sufficient information for supporting an optimization algorithm for finding the minimum accumulated strain. An asphalt pavement design with accumulated strain value of 1772 micro-strain (0.02 inch in a 10-inch-thick asphalt pavement layer) is obtained by solving the optimization problem and the design parameters meet flexible pavement design specifications. The proposed framework

is able to generate accurate predictive models when the original inputs are highly correlated and able to map to an optimal point in the fitted input space.

3.2 Introduction

Design and control of systems require engineers to predict complex behavior, configuration, and environmental loads to which they are exposed. This obviously constitutes a challenging task of determining a mapping from cause to effect, thus engineers model material and system behaviors, from which they make predictions about the expected behavior of specific system configurations. This behavior is often governed by nonlinear multivariate (and sometimes unknown) interrelationships that can vary over time due to degradation of material properties and occur within varying and less controllable physical environments. Another challenge in addition to predicting system behavior is the estimation of unmeasured attributes based on measured attributes.

Flexible pavement is one such complicated system for which engineers try to predict behavior, and permanent deformation (i.e., rutting) is one of the most important flexible pavement design criteria. Rutting usually occurs under the wheel-path and appears as longitudinal depressions with small upheavals to the side, that occur as differential consolidation and cause safety issues for drivers Zhang et al. (2015); Ghasemi et al. (2018b).

Mechanistic-empirical, regression-based, modeling and performance-testing approaches are commonly used in asphalt mixtures rut susceptibility analysis, and viscoelastic, viscoplastic, and viscoelastoplastic continuum damage-based modeling can all be used to explain the rutting behavior of asphalt mixtures Bashin et al. (2012). A common problem with such models is that they are highly dependent on empirical data and require accurate characterization of asphalt binder behavior. Rutting is commonly simulated using a rutting resistance indicator parameter called Flow Number (FN), and defined as the point where the permanent strain rate reaches a minimum value, as shown in Figure 3.1, that can be measured through repeated loading and unloading tests. This parameter has been found indicative of good correlation with field rutting of asphalt mixtures exposed to different levels of traffic AASHTO TP 79-13 (2013); Arabali et al. (2017).

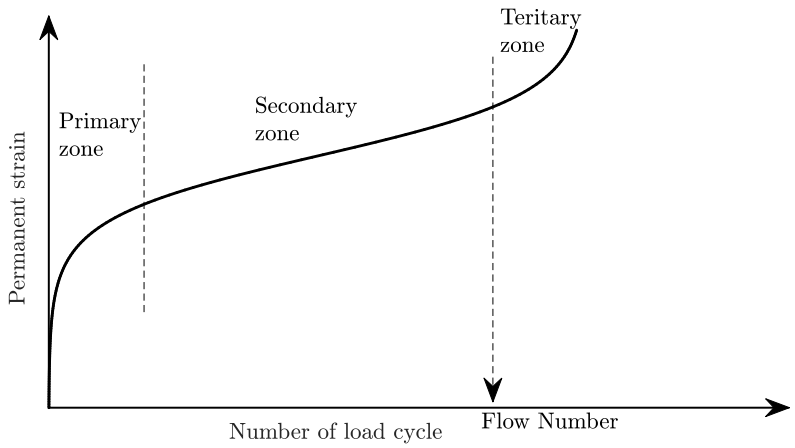

Figure 3.1: Typical repeated load permanent deformation behavior.

During the design procedure, there is generally a limiting criterion of 0.4 inches (10.16 mm) with respect to the total deformation of a pavement structure. According to the existing literature Kaloush et al. (2003); Witzcak (2002), parameters affecting the rutting behavior of asphalt mixtures can be classified into three categories, i.e., asphalt binder properties representing viscoelastic and viscoplastic behavior, aggregate properties representing elastic/plastic behavior, and mixture properties. Many studies have demonstrated that the amount of rutting depends on mixture volumetric properties, binder viscosity, and testing temperature. Kvasnak et al. Kvasnak et al. (2007), demonstrated that binder grade, binder viscosity, asphalt content, testing temperature, nominal maximum aggregate size (NMAS), voids in mineral aggregate (VMA), percentage aggregate passing through sieve sizes No.4, No.16, No.200, and the number of gyrations all can affect the FN of asphalt mixture. Rodezno et al. Rodezno et al. (2010) designated 12 parameters affecting rutting behavior of asphalt mixtures, i.e., testing temperature, maximum shear stress, normal stress, binder viscosity, percentage aggregate passing through sieve sizes 3/4-inch, 3/8-inch, and No.4, percentage air voids, effective binder content, binder content, VMA, and voids filled with asphalt (VFA). Although Apeagyei, et al. Apeagyei (2011), indicated that using dynamic modulus test results at specific temperature and loading frequencies in conjunction with aggregate gradation seems to exhibit good correlation with FN test results, there is some disagreement on the existence of correlation between the results of these two performance tests Birgisson et al. (2004a); Pellinen and

Witczak (2002); Timm et al. (2006), so engineers have tried to predict asphalt pavement rutting behavior based on system properties Kvasnak et al. (2007); Rodezno et al. (2010).

Predictive modeling is a process that can use data mining tools and probability theory techniques to forecast outcomes from a given system, with each model constructed with several predictors likely to influence future results. Once data have been collected for relevant predictors, a statistical model is formulated that may use a simple linear equation, or may rather use be a complex structure such as an Artificial Neural Network (ANN) obtained through sophisticated optimization algorithms Carpenter and Barthelemy (1993); Nicholas et al. (2015). ANN modeling has the same goal as statistical regression modeling because, both attempt to capture the relationship between a set of model inputs and corresponding outputs Cheng and Titterington (1994a); Liew et al. (2004) by estimating a set of coefficients that provide the best fit with the data.

The asphalt pavement properties discussed above are not necessarily accepted as an independent set of variables because they may not be adequate for modeling. Since cross-correlated inputs in the modeling data set can adversely affect accurate estimation of their causative effects on the response variable and this impede the capability of the model to accurately estimate the response variable Rollins et al. (2006); K Rollins Sr (2015), a pre-processing step is often useful to examine the quality and relationships among the input variables, although this step is rarely practiced by design engineers in this particular application.

In the presence of correlated input variables, orthogonal (uncorrelated) variables can be obtained using a pseudo-variable approach called principal component analysis (PCA) Rollins et al. (2006); Jolliffe (2002); Noh et al. (2009); Yi et al. (2017); Lever et al. (2017). PCA is a multivariate statistical procedure that uses an orthogonal transformation to convert a set of correlated variables into a set of uncorrelated (pseudo-) variables called principal components (PCs). The PCs are a set of orthogonal variables that are linear combinations of the original variables in the dataset. Such an approach has been applied to a number of engineering problems as an essential part of the design procedure Xiong et al. (2017); Su and Tong (1997); Happ and Greven (2018); Hargrove et al. (2017); Malik et al. (2018); Xiong et al. (2018); Lim et al. (2018); Granato et al. (2017).

Extrapolation using empirical predictive models is highly risky and therefore often discouraged by experienced practitioners. Given a training data set $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ a common purpose of linear/nonlinear predictive modeling is to find an accurate approximation, e.g., $\hat{f}(x)$, of a function $f(x) = E(y|x)$, that can be used to predict the value of a new y given x. When x lies outside the convex hull of x_i values, the prediction is called an extrapolation and the behavior of the model is not predictable Devore (2011). Extrapolation inflates standard prediction error compared to interpolation, a process in which, x lies inside the convex hull. To avoid using locations outside the convex hull, one might define a hyper-space for the input set. This n-dimensional domain might satisfy certain qualifications, including convexity and minimum volume enclosing data points, to prevent extrapolation.

The focus of the current study is to define a machine-learning based framework to predict rutting behavior of asphalt mixture. The proposed framework is uniquely addressing the correlated inputs in modeling using PCA. Unlike most of the existing pavement performance predictive models, the proposed framework is seeking to define the n-dimensional input space to guard against extrapolation. These issues are commonly ignored in pavement predictive modeling problems. This framework can be used to find an optimal design using an optimization algorithm. The minimum volume enclosing ellipsoid enters as a constraint into the optimization problem to ensure that no extrapolation is occurring. Although the domain is convex, the modeled function is multimodal, making this a nonconvex optimization problem that we tackle in this study with an evolutionary optimization solution approach called Mean-Variance Mapping Optimization (MVMO) Rueda and Erlich (2015); Erlich et al. (2010). The current implementation of MVMO employs adaptive strategies on essential control parameters in order to explore the search domain initially and exploit the domain of interest in the later stages of optimization. A major interest in this algorithm for pavement rutting design optimization (where a moderately accurate solution is assumed to be sufficiently good as long as it meets the design specifications) is the fast convergence rate of MVMO compared to classical evolutionary algorithms. This property of the algorithm has been confirmed through rigorous analytical and engineering-based benchmarking Rueda and Erlich (2015).

This manuscript is organized as follows: the materials and the developed methodologies are presented in Section 3.3 and the results and discussion and an application of the developed framework are presented in Section 3.4, followed by conclusions in Section 3.5.

3.3 Materials and Methodologies

3.3.1 Data collection

Eighty-three specimens from 21 different asphalt mixtures collected from different projects in the State of Wisconsin provided the database for subsequent modeling. Asphalt mixtures required for the study were sampled at the plant site directly from the back of the delivery trucks, and asphalt binders corresponding to each pavement section were sampled during mix plant production. Maximum theoretical values of specific gravity were measured and used to determine other volumetric properties of the asphalt mixtures. Asphalt mixtures were compacted in the laboratory using a Superpave Gyratory Compactor (SGC) into test samples whose dimensions were 150 mm in diameter by 170 mm in height.

Each 100-mm diameter by 150-mm height cylindrical specimen was cored, trimmed and prepared for dynamic modulus testing. The specimens were tested under repeated sinusoidal compressive stress at an effective test temperature of 36.6 $^{\circ}C$ and at the four loading frequencies 25, 10, 1, and 0.1 Hz under unconfined conditions. Flow number tests were performed on the same specimens under repeated haversine compressive stress at a single effective temperature of 36.6 $^{\circ}C$, with the load applied for a duration of 0.1 second and a dwell period of 0.9 second. No confining pressure was used and the axial stress was similar to the deviator stress (600 kPa). To obtain binder shear properties, a complex shear modulus test was conducted at the same effective test temperature and similar frequencies used in the the dynamic modulus tests.

It should be mentioned that the effective test temperature for all of the laboratory tests (36.6° C) was based on climatic condition of the Midwestern area of the United States (central Wisconsin), and was considered the typical temperature at which permanent deformation happens, equivalent to a seasonal correction throughout the year.

3.3.2 Pre-processing step: input variable selection

A fundamental and crucial step in identifying the optimal functional form of statistical models is input variable selection. In general, this requires a set of input variables sufficiently high in information content that maps to the output space. The difficulty in selecting a parsimonious set of input variables can arise from several conditions, including too many available variables, highly-correlated input variables, and input variables weakly-related or unrelated to the response Fodor (2002).

Mathematical modeling is the process of functionally relating measured input variables to output variables. The modeler selects a mathematical structure and a process for estimating unknown model parameters. For a general model structure, let the expectation function of the response be represented as $\eta_i = f(\mathbf{x}_i, \theta)$, where η_i is the expected value of the response (i.e., output) at the ith sampling time, $i = 1, \ldots, \ n$, $\mathbf{x}_i$ is the vector of input values at the ith sampling time, and θ is the vector of unknown model parameters with $\theta = [\theta_1 \ldots \theta_q]^T$. Let the element of the Jacobian matrix, $\mathbf{J}$, in the ith row and jth column be $\frac{\partial \eta_i}{\partial \theta_j}$ i.e., $\mathbf{J} = \left\{ \frac{\partial \eta_i}{\partial \theta_j} \right\}$. Note that the jth column represents θ_j and its column vector represents the change in the response space as θ_j changes for a particular set of experimental conditions. If two columns such as j and k are orthogonal, their correlation coefficient is zero. More specifically, if these two columns are orthogonal, the information used to estimate θ_j is decoupled (i.e., separate or independent) from the information to estimate θ_k, and vice versa. The advantage of the orthogonal attribute is that it strengthens causative relationships of inputs on the output while maximizing parameter accuracy and consequently estimation accuracy for the modeled output. Correlated columns in the Jacobian Matrix arise from pairwise correlation of inputs, so for a given set of experimental data, to minimize standard parameter errors and maximize accuracy mapping of input behavior into the response space, we should seek to minimize pairwise correlation of the inputs used to model the response behavior.

In the present study, the vector of input values ($\mathbf{x}$), represents the asphalt mixture's component properties, recalling from the literature Kaloush et al. (2003); Witzcak (2002); Kvasnak et al. (2007); Rodezno et al. (2010), that the rutting behavior of an asphalt mixture can be legitimately

Table 3.1: Original Input Variables

Variable	Identity	Min.	Max.	Ave.	Std. Ave.
x_1	Binder %	3.400	6.600	5.093	0.773
x_2	G*	210800.543	1163559.917	612179.576	265903.652
x_3	NMAS	12.500	25.000	15.922	3.761
x_4	Passing 3/4”	81.300	100.000	98.554	4.085
x_5	Passing 1/2”	38.300	98.800	87.130	15.199
x_6	Passing 3/8”	34.100	89.900	76.340	15.099
x_7	Passing #4	26.200	72.500	56.248	13.741
x_8	Passing #8	17.500	54.000	42.249	10.512
x_9	Passing #16	14.200	47.400	32.178	8.740
x_{10}	Passing #30	9.600	47.400	32.178	8.740
x_{11}	Passing #50	5.700	18.600	12.022	3.179
x_{12}	Passing #100	3.700	9.800	6.187	1.424
x_{13}	Passing #200	2.800	8.500	4.322	1.115
x_{14}	VMA	10.323	21.000	16.452	2.502
x_{15}	VFA	46.450	91.719	65.189	9.062
x_{16}	Va%	1.019	9.825	5.868	2.088
x_{17}	E*	395.700	2299.400	869.410	411.524

characterized by its component properties. The component properties and their ranges used in the present study are summarized in Table 3.1.

These seventeen properties are used as input variables to predict accumulated strain value at the flow number (FN), and the cross or pairwise correlation matrix of the input variables is given by Table 3.2. Results with absolute values of 0.5 or higher are shown in bold and red text. As can be seen, the absolute value of 273 correlation coefficients are above 0.1, with 41 of them greater than 0.5, indicating that several of the input variables appear to be highly correlated. Therefore, to accurately map the inputs to the response, it would be useful to obtain a smaller set of orthogonal pseudo-variables as linear combinations of the inputs using the PCA method and mapping them to the response is required Fodor (2002).

Table 3.2: Correlation Matrix for the Input Variables

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	1	0.33	-0.73	0.6	0.71	0.63	0.47	0.45	0.46	0.45	0.41	0.43	0.42	0.61	-0.18	0.35	-0.43
x_2	0.33	1	-0.11	0.08	-0.03	-0.23	-0.38	-0.32	-0.18	-0.06	0.01	0.23	0.35	0.02	-0.02	0.02	0.24
x_3	-0.73	-0.11	1	-0.64	-0.76	-0.69	-0.51	-0.41	-0.39	-0.32	-0.25	-0.30	-0.39	-0.53	0.13	-0.30	0.37
x_4	0.6	0.08	-0.64	1	0.88	0.78	0.58	0.51	0.42	0.32	0.25	0.23	0.24	0.46	-0.21	0.32	-0.45
x_5	0.71	-0.03	-0.76	0.88	1	0.95	0.77	0.71	0.61	0.49	0.41	0.39	0.36	0.56	-0.33	0.44	-0.51
x_6	0.63	-0.23	-0.69	0.78	0.95	1	0.92	0.86	0.75	0.61	0.49	0.30	0.16	0.53	-0.27	0.38	-0.49
x_7	0.47	-0.38	-0.51	0.58	0.77	0.92	1	0.95	0.84	0.69	0.49	0.11	-0.14	0.43	-0.16	0.26	-0.37
x_8	0.45	-0.32	-0.41	0.51	0.71	0.86	0.95	1	0.95	0.83	0.59	0.11	-0.15	0.42	-0.16	0.26	-0.41
x_9	0.46	-0.18	-0.39	0.42	0.61	0.75	0.84	0.95	1	0.96	0.74	0.18	-0.11	0.40	-0.16	0.25	-0.41
x_{10}	0.45	-0.06	-0.32	0.32	0.49	0.61	0.69	0.83	0.96	1	0.84	0.28	-0.06	0.35	-0.15	0.22	-0.37
x_{11}	0.41	0.01	-0.25	0.25	0.43	0.49	0.49	0.59	0.74	0.84	1	0.59	0.17	0.21	-0.20	0.20	-0.25
x_{12}	0.43	0.23	-0.30	0.23	0.39	0.30	0.11	0.11	0.18	0.28	0.59	1	0.82	0.24	-0.28	0.28	-0.15
x_{13}	0.42	0.35	-0.39	0.24	0.36	0.16	-0.14	-0.15	-0.11	-0.06	0.17	0.82	1	0.29	-0.30	0.32	-0.22
x_{14}	0.61	0.02	-0.53	0.46	0.56	0.53	0.43	0.42	0.40	0.35	0.21	0.24	0.29	1	-0.62	0.83	-0.71
x_{15}	-0.18	-0.02	0.13	-0.21	-0.33	-0.27	-0.16	-0.16	-0.16	-0.15	-0.20	-0.28	-0.30	-0.62	1	-0.94	0.52
x_{16}	0.35	0.02	-0.30	0.32	0.44	0.38	0.26	0.26	0.25	0.22	0.20	0.28	0.32	0.83	-0.94	1	-0.63
x_{17}	-0.43	0.24	0.37	-0.45	-0.51	-0.49	-0.37	-0.41	-0.41	-0.37	-0.25	-0.15	-0.22	-0.71	0.52	-0.63	1

3.3.3 Principal component analysis (PCA)

PCA is an orthogonal linear transformation of a set of variables to a new coordinate system. Among the transformed variables, the first has the greatest variance or spread along its axis, followed by the second, then the third, and so on. The transformation can be thought of as fitting an n-dimensional ellipsoid to the data, where each axis represents one of the principal components (PCs) such that the variance along the small axis of the ellipsoid is also small, so omitting that axis and its corresponding principal component from the representation of the data set results in losing, only a small portion of the information. Another explanation of the PCA could be in its revealing of the internal structure of the data in a way that best explains the variance within the data. If a multivariate dataset is visualized as a set of coordinates in an n-dimensional data space, with one axis per variable, PCA can provide a lower-dimensional picture, a projection of the object viewed from its most informative viewpoint Jolliffe (2002). This operation is visualized for three

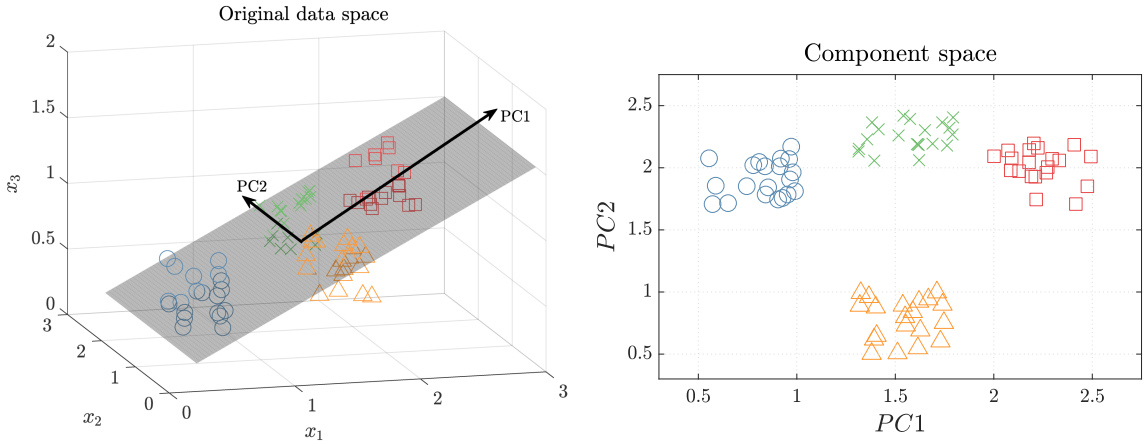

Figure 3.2: Schematic of the PCA transformation.

input variables in Figure 3.2. The dimensionality of the transformed data is reduced by using only the first few principal components representing the highest variations in the original data.

Considering a data set $\mathbf{X}$, PCA reduces the dimension of $\mathbf{X}$ by expressing the p (p =17) original variables $(x_1, \ldots, x_p)$ as d new pseudo-variables (principal components, PCs), where $d < p$. The PCs are a set of orthogonal (i.e., uncorrelated), linear combinations of the original variables within the dataset.

As a data pre-processing methodology, PCA determines an optimal rotational transformation of the dataset $\mathbf{X}$ into an orthogonal set of PCs (pseudo-variables) that maximizes the information content of the original data for a smaller set of these pseudo variables Fodor (2002); Johnson et al. (2014). With respect to dataset, $\mathbf{X}$, PCA is performed in the following steps:

1. Standardizing $\mathbf{X}$ by transforming it to $\mathbf{Z}$ using the following equations:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{\mathrm{np}} \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 & \cdots & \mathbf{z}_p \end{bmatrix} = \begin{bmatrix} \frac{x_{11}-\overline{x}_1}{s_1} & \frac{x_{12}-\overline{x}_2}{s_1} & \cdots & \frac{x_{1p}-\overline{x}_p}{s_p} \\ \frac{x_{21}-\overline{x}_1}{s_1} & \frac{x_{22}-\overline{x}_2}{s_2} & \cdots & \frac{x_{2p}-\overline{x}_p}{s_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x_{n1}-\overline{x}_1}{s_1} & \frac{x_{n2}-\overline{x}_2}{s_2} & \cdots & \frac{x_{np}-\overline{x}_p}{s_p} \end{bmatrix} \tag{3.1}$$

   where, for $k = 1$ *to* n and $j = 1$ *to* p, x_{kj} is the kth measurement for the jth variable, $\overline{x}_k$ is sample mean for the kth variable, and s_k is sample standard deviation for the kth variable.

2. Determine the unit eigenvectors $e_1, \ldots, e_p$ of $\mathbf{Z}$.

3. Determine the corresponding eigenvalues $\lambda_1, \ldots, \lambda_p$.

4. Rank the eigenvectors according to their eigenvalues.

5. Select d PCs according to their eigenvalues (or the scree plot).

Examining the eigenvalues of each PC, corresponding to the amount of variance explained by each, is the basis for selecting the PCs. A common method is to rank the PCs and select those with eigenvalues exceeding a significant threshold. The first PC contains the most variance captured in a single axis, the second PC is orthogonal to the first (their correlation is zero) and contains the second greatest variance, the third PC is orthogonal to all previous ones and also contains the third greatest variance, etc. Rollins et al. (2006). Generating a plot of the percentage contributions of each PC indicating eigenvalues versus their order, called a scree plot, visualizes the PC selection strategy. The "elbow point" in this plot indicates the number of significant PCs needed to properly represent the data.

The eigenvectors and eigenvalues of the correlation matrix are calculated and presented in Tables 3.3 and 3.4, respectively. Table 3.4 shows the fraction of total variation in the data as represented by each eigenvalue. The eigenvalues are ordered by decreasing in contribution to the total variation. According to the eigenvalues, 89.72% of the variation in the original data is represented by the first five PCs. The scree plot presented in Figure 3.3 indicates that the elbow point is located at the fourth PC, meaning that first four PCs appear to be significant. We conservatively selected the first five PCs as our pseudo-input variables.

Forming a linear combination of the original input variables, the PCs are obtained using Eq. 3.2:

$$pc_i = \sum_{j=1}^{17} \alpha_{ij} x_j + \beta_i \tag{3.2}$$

where $i = 1\ to\ 5$, the α_{ij} is the corresponding coefficients, the β_i are constants, and the x_js are the original input variables. Eq. 3.2 can be presented in matrix notation as in Eq. 3.3:

$$\mathbf{p} = \mathbf{Mz} + \mathbf{n} \tag{3.3}$$

where

$$\mathbf{p} = \begin{bmatrix} pc_1 \\ pc_2 \\ pc_3 \\ pc_4 \\ pc_5 \end{bmatrix}$$

$$\mathbf{M}^T = \begin{bmatrix} 3.41 \times 10^{-1} & -2.04 \times 10^{-1} & 2.89 \times 10^{-1} & -1.94 \times 10^{-1} & 3.71 \times 10^{-1} \\ -7.60 \times 10^{-8} & -1.17 \times 10^{-6} & 1.31 \times 10^{-6} & 1.05 \times 10^{-7} & 2.47 \times 10^{-6} \\ -6.59 \times 10^{-2} & 3.22 \times 10^{-2} & -4.53 \times 10^{-2} & 9.10 \times 10^{-2} & -7.97 \times 10^{-3} \\ 6.34 \times 10^{-2} & -1.31 \times 10^{-2} & 1.86 \times 10^{-2} & -9.38 \times 10^{-2} & -7.13 \times 10^{-3} \\ 2.11 \times 10^{-2} & -2.33 \times 10^{-3} & 5.26 \times 10^{-3} & -1.63 \times 10^{-2} & -9.03 \times 10^{-3} \\ 2.17 \times 10^{-2} & 7.23 \times 10^{-3} & 1.55 \times 10^{-3} & -1.23 \times 10^{-2} & -1.13 \times 10^{-2} \\ 2.13 \times 10^{-2} & 2.01 \times 10^{-2} & -2.82 \times 10^{-3} & -7.27 \times 10^{-3} & -7.11 \times 10^{-3} \\ 2.79 \times 10^{-2} & 2.89 \times 10^{-2} & -2.98 \times 10^{-3} & 3.79 \times 10^{-3} & 1.41 \times 10^{-3} \\ 3.28 \times 10^{-2} & 3.22 \times 10^{-2} & 4.43 \times 10^{-3} & 2.35 \times 10^{-2} & 1.71 \times 10^{-2} \\ 3.73 \times 10^{-2} & 3.39 \times 10^{-2} & 1.53 \times 10^{-2} & 4.98 \times 10^{-2} & 3.14 \times 10^{-2} \\ 7.08 \times 10^{-2} & 3.30 \times 10^{-2} & 7.38 \times 10^{-2} & 1.54 \times 10^{-1} & -9.05 \times 10^{-3} \\ 1.08 \times 10^{-1} & -2.03 \times 10^{-1} & 2.33 \times 10^{-1} & 2.30 \times 10^{-1} & -2.72 \times 10^{-1} \\ 9.18 \times 10^{-2} & -4.06 \times 10^{-1} & 2.29 \times 10^{-1} & 6.69 \times 10^{-2} & -3.18 \times 10^{-1} \\ 9.98 \times 10^{-2} & -8.77 \times 10^{-2} & -1.12 \times 10^{-1} & -1.14 \times 10^{-2} & 8.85 \times 10^{-2} \\ -1.73 \times 10^{-2} & 3.21 \times 10^{-2} & 4.57 \times 10^{-2} & -2.69 \times 10^{-2} & -9.43 \times 10^{-4} \\ 9.75 \times 10^{-2} & -1.42 \times 10^{-1} & -1.95 \times 10^{-1} & 7.31 \times 10^{-2} & 4.70 \times 10^{-2} \\ -5.50 \times 10^{-4} & 2.73 \times 10^{-4} & 8.20 \times 10^{-4} & -8.92 \times 10^{-6} & 1.36 \times 10^{-4} \end{bmatrix}$$

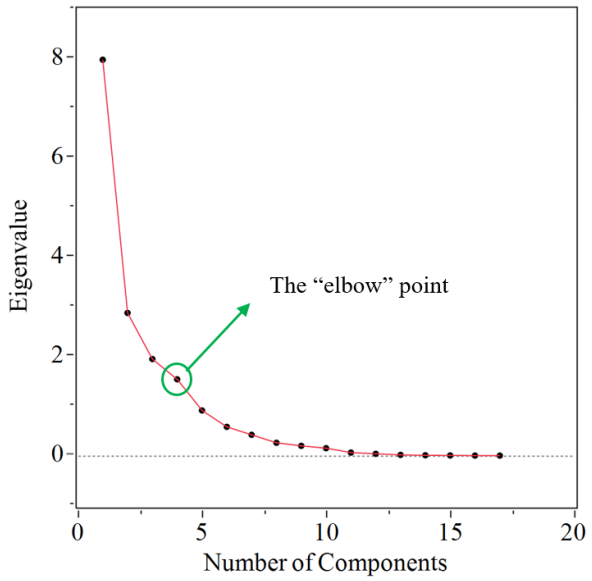

Figure 3.3: Scree plot of components' contribution.

$$\mathbf{n} = \begin{bmatrix} -1.72 \times 10 \\ 5.87 \times 10^{-1} \\ -8.21 \\ 7.36 \\ -4.85 \times 10^{-1} \end{bmatrix}$$

The PCs obtained in this way will be used in further modeling.

3.3.4 K-fold cross validation

To recognize how the results of the statistical analysis can generalize to an independent dataset, and to prevent overfitting, a k-fold cross validation technique, a procedure to avoid overfitting models by checking the fitted model against a set of data not used in fitting the model Refaeilzadeh et al. (2009), is used. During this procedure, the given dataset is randomly partitioned into k equal-sized subsamples from which k-1 subsamples are used as training data, and the remaining single subsample is used for testing the developed model. This procedure is then repeated k times (number of the folds), with each of the k subsamples used exactly once as test data. In this way, all

Table 3.3: Eigenvectors of the **Z** matrix

	pc_1	pc_2	pc_3	pc_4	pc_5	pc_6	pc_7	pc_8	pc_9	pc_{10}	pc_{11}	pc_{12}	pc_{13}	pc_{14}	pc_{15}	pc_{16}	pc_{17}
z_1	0.26	-0.16	0.23	-0.15	0.29	-0.28	0.14	0.27	0.48	-0.49	0.29	-0.13	-0.04	0.03	0.01	0.05	0.01
z_2	-0.03	-0.31	0.35	0.02	0.67	0.26	-0.15	0.08	-0.26	-0.03	-0.4	0.07	-0.04	0.01	0.02	-0.03	0.03
z_3	-0.25	0.12	-0.17	0.35	-0.02	0.21	-0.39	0.7	0.13	-0.07	0.1	0.17	0.1	0.1	-0.01	0.02	0.04
z_4	0.26	-0.06	0.08	-0.38	-0.03	0.28	-0.57	-0.04	0.18	0.36	0.24	-0.31	-0.2	0.1	-0.03	-0.02	0.03
z_5	0.32	-0.04	0.08	-0.24	-0.14	0.21	-0.12	0	-0.05	-0.12	0.06	0.48	0.42	-0.25	0.01	0.13	-0.49
z_6	0.33	0.11	0.02	-0.18	-0.17	0.17	0.05	0.07	-0.01	-0.12	-0.18	0.18	0.24	-0.02	0.03	-0.1	0.79
z_7	0.29	0.28	-0.04	-0.1	-0.1	0.17	0.23	0.2	0.01	-0.06	-0.41	-0.15	-0.08	0.62	-0.09	0.08	-0.29
z_8	0.29	0.3	-0.03	0.04	0.01	0.08	0.09	0.25	-0.22	-0.07	0.01	-0.04	-0.46	-0.44	0.15	-0.49	-0.14
z_9	0.29	0.28	0.04	0.2	0.15	-0.04	0.01	0.03	-0.31	0.05	0.2	-0.02	-0.18	-0.13	-0.09	0.75	0.13
z_{10}	0.26	0.24	0.11	0.35	0.22	-0.11	-0.07	-0.11	-0.2	0.09	0.28	-0.28	0.56	0.19	0.01	-0.33	-0.04
z_{11}	0.22	0.11	0.23	0.49	-0.02	-0.01	-0.15	-0.35	0.43	0.08	-0.09	0.45	-0.28	0.14	-0.01	-0.1	0.01
z_{12}	0.15	-0.29	0.34	0.34	-0.39	0.03	0.06	0.19	0.15	0.16	-0.31	-0.42	0.15	-0.33	0.04	0.13	-0.05
z_{13}	0.1	-0.46	0.26	0.08	-0.35	-0.11	0.04	0.15	-0.47	-0.04	0.32	0.19	-0.21	0.36	-0.04	-0.12	0.04
z_{14}	0.25	-0.22	-0.29	-0.03	0.22	-0.24	0.26	0.28	0.12	0.61	-0.02	0.22	0.05	-0.05	-0.33	-0.08	0.01
z_{15}	-0.16	0.29	0.41	-0.25	-0.01	-0.41	-0.03	0.21	0	0.35	-0.08	0.16	0.05	0.08	0.53	0.06	0
z_{16}	0.2	-0.3	-0.41	0.15	0.1	0.2	0.14	-0.06	0.05	0.06	0.08	-0.01	-0.01	0.11	0.75	0.11	0
z_{17}	-0.23	0.12	0.33	-0.01	0.05	0.58	0.53	0.04	0.15	0.22	0.37	0.04	0.02	0	-0.02	-0.01	0.02

Table 3.4: Eigenvalues of the **Z** Matrix and the Corresponding Percent Variance

Number	Eigenvalue	Percent Variance	Cumulative Percent
1	7.979	46.935	46.935
2	2.876	16.920	63.856
3	1.947	11.456	75.311
4	1.539	9.052	84.363
5	0.912	5.363	89.726
6	0.581	3.419	93.145
7	0.421	2.475	95.620
8	0.260	1.529	97.149
9	0.198	1.165	98.314
10	0.152	0.892	99.206
11	0.063	0.371	99.577
12	0.038	0.224	99.801
13	0.016	0.095	99.896
14	0.008	0.049	99.945
15	0.005	0.028	99.973
16	0.003	0.016	99.989
17	0.002	0.011	100.000

observations are used for both training and testing exactly once, constituting which is the advantage of k-fold cross-validation technique over repeated random sub-sampling.

In the present study, based on the size of the data set (83 values), 3-folds with 24 sample vectors in each have been randomly selected, and 3-fold cross validation was performed.

3.3.5 Principal Component Regression (PCR)

In statistics, ordinary least squares (OLS) is a method for estimating the unknown parameters in a linear regression model, with the goal of minimizing the sum of the squares of the differences between the observed responses in the given dataset and those fitted by a linear function of parameters using a set of explanatory variables. This can be visually seen as the sum of the squared vertical distances between each data point in the set and the corresponding point on the regression line (the smaller the differences, the better the model fits the data).

Several multiple regression structures including those with second-order quadratic and interaction terms considered to represent the relationship between the dependent variable (accumulated strain) and explanatory variables (PCs). The least-squares criterion of minimizing the sum of squared residuals (SSE) is used for estimating the values of the unknown model coefficients. For the training fold, this minimizes the sum of squared residuals and develops a closed-form expression for the estimated value of the unknown parameter. While full third-order models and reduced third-order models were also developed and fitted, they were rejected due to their inability to improve the fit to the training data and also because they were deficient in fitting to the test data.

3.3.6 Principal Component Neural Network (PCNN)

Using artificial neural networks (ANN), a predictive model of the response that we call, "Principal Component Neural Network (PCNN)" was developed. ANNs are computational tools composed of highly-interconnected networks of many simple processors or nodes arranged in different layers Kartam (1994); García-Segura et al. (2017); May et al. (2011). Our three-layer feed-forward neural network, consisting of an input layer of 5 neurons, a hidden layer of 10 neurons, and an output

layer of one neuron was developed using the MATLAB MATLAB (2012) program. Selection of the number of hidden neurons is based on a trial- and- error optimization procedure balancing between the cost function and computational time.

The input of each processing element, pc_i, is multiplied by an adjustable connection weight w_{ij}. At each processing element, the weighted input signals are summed and a threshold value, b_{Hj}, is added. This combined input is then passed through a non-linear transfer function, $f(.)$, to produce the output of the first layer, ν_j, forming the input to the next layer. The network adjusts its weights on the presentation of a training dataset and uses a learning rule to find a set of weights that will produce the input/output mapping that results in the smallest possible error. This process is called "learning" or "training". Once the training phase of the model has been successfully accomplished, the performance of the trained model must be validated using an independent testing set Cheng and Titterington (1994a); Kartam (1994); Juang and Chen (1999b). This process is visually represented in Figure 3.4.

The output ν_j from the j^{th} hidden nodes is given by

$$\nu_j = f_1(pc_i, w_{ij}), \quad i = 1, ..., 5 \quad and \quad j = 1, ..., 10 \tag{3.4}$$

and the single output $\hat{y}$ is:

$$\hat{y} = f_2\left(f_1(pc_i, w_{ij})\right). \tag{3.5}$$

Then the expression of $\hat{y}$ as a function of pc becomes a complicated nonlinear regression function with the j sets of weights, as parameters. As a special case it is assumed that

$$f_2(\nu_j, w_{Hj}) = b_0 + \sum_j \nu_j w_{Hj}, \tag{3.6}$$

and for each j,

$$f_1(pc_i, w_{ij}) = b_{Hj} + \sum_i pc_i w_{ij} \tag{3.7}$$

so a general form of the feed forward neural network is described in Eq. 3.8:

$$\hat{y} = f_2\left\{b_0 + \sum_{j=1}^{n}\left[w_{Hj} \cdot f_1\left(b_{H_j} + \sum_{i=1}^{m} pc_i w_{ij}\right)\right]\right\} \tag{3.8}$$

where, b_0 is bias at output layer (just one neuron at this layer); w_{Hj} is weight of connection between neuron j of the hidden layer and output layer neuron; b_{H_j} is bias at neuron j of the hidden layer (for $j = 1$ to 10); w_{ij} is weight of connection between input variable i (for $i = 1$ to 5) and neuron j of the hidden layer; pc_i is pseudo input parameter i; $f_1(t)$ is transfer function of the hidden layer, and $f_2(t)$ is transfer function of the output layer.

Both transfer functions $f_1(t)$ and $f_2(t)$ used in this paper are sigmoid functions which, for an arbitrary variable t, are defined in Eq. 3.9,

$$f_k(t) = \frac{1}{1 + e^{-t}} \quad for \quad k = 1, 2. \tag{3.9}$$

For each fold, the training procedure starts with adjusting the initial values of the network's weights and biases to obtain a reasonable output and continue to modify the network by minimizing the value of the sum of squared errors (SSE). The iteration continues until the convergence criterion is met. It is worth mentioning that this is the reason that a validation set is not used, and just as for the linear regression model, training finishes when the convergence criterion in satisfied. The Bayesian Regularization algorithm is implemented for network training efficiency.

3.3.7 Effective variable space

As discussed above, these empirical models are prone to high uncertainty when used in extrapolation. To avoid extrapolation, one should find the n-dimensional hyper-space that is representative of the space of the n inputs used to train the model. As discussed earlier, to achieve this goal, a symmetrical convex space containing all the input data used in the development of the model should be determined.

It is assumed that each input variable (x_i) has a normal distribution, so their joint distribution is bi-variate normal, and such distributions are frequently portrayed in terms of a contour diagram. A contour curve on such a diagram is composed of all the points on the surface equidistant from the $x_i x_j$ plane. Expressed in a different way, a contour curve is composed of all (x_i, x_j) outcomes that have a constant density function Kutner et al. (2004), so one can picture a contour as the

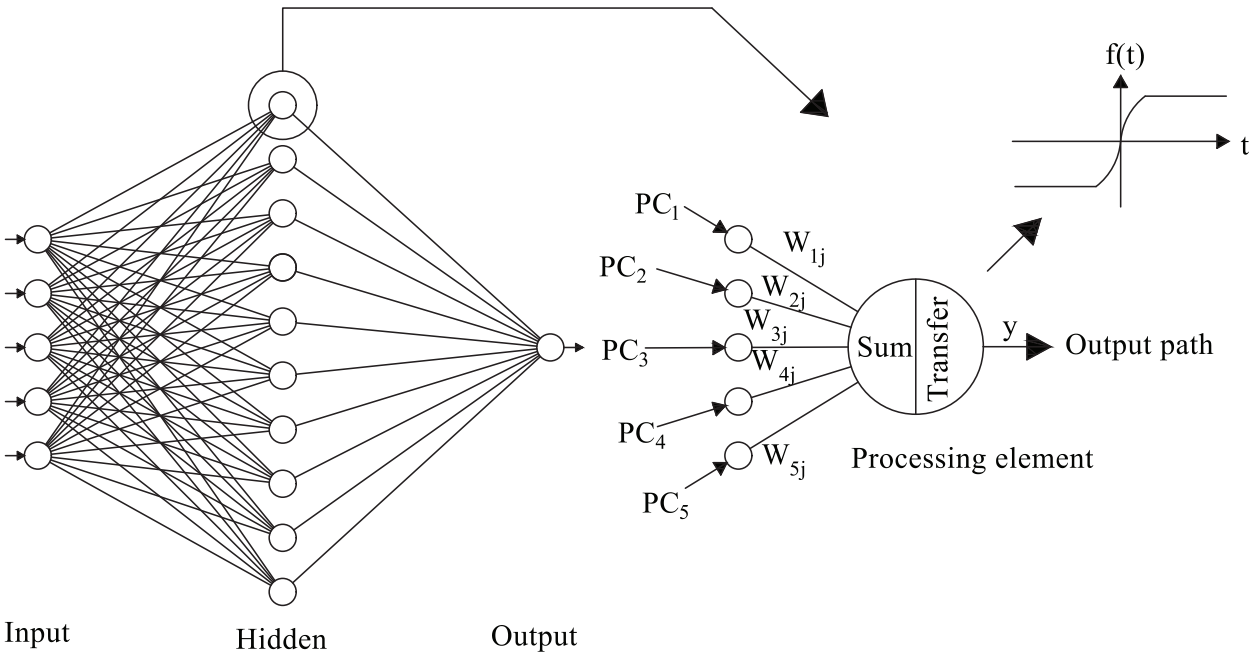

Figure 3.4: Schematic of the training process in a processing element

cross section obtained by slicing a bivariate normal surface at a fixed distance from the $x_i x_j$ plane. A schematic view of this concept is presented in Figure 3.5.

Therefore, the minimum volume-enclosing ellipsoid should be found and added as a constraint to the developed models. Minimum volume enclosing ellipsoids play an important role in several diverse applications such as optimal design, computational geometry, convex optimization, pattern recognition, and statistics.

The minimum enclosing ellipsoid will provide the n-dimensional symmetrical convex space that includes all the data points. The ellipsoid method generates a sequence of ellipsoids whose volume uniformly decreases at each step, thus enclosing a minimizer of a convex function. As an algorithm for solving linear programming problems with rational data, the ellipsoid algorithm was studied by Leonid Khachiyan Todd and Yıldırım (2007). The idea is to formulate an optimization problem that approximately finds a minimum volume enclosing ellipsoid, Vol($\mathcal{E}$), given p data points in n-dimensions.

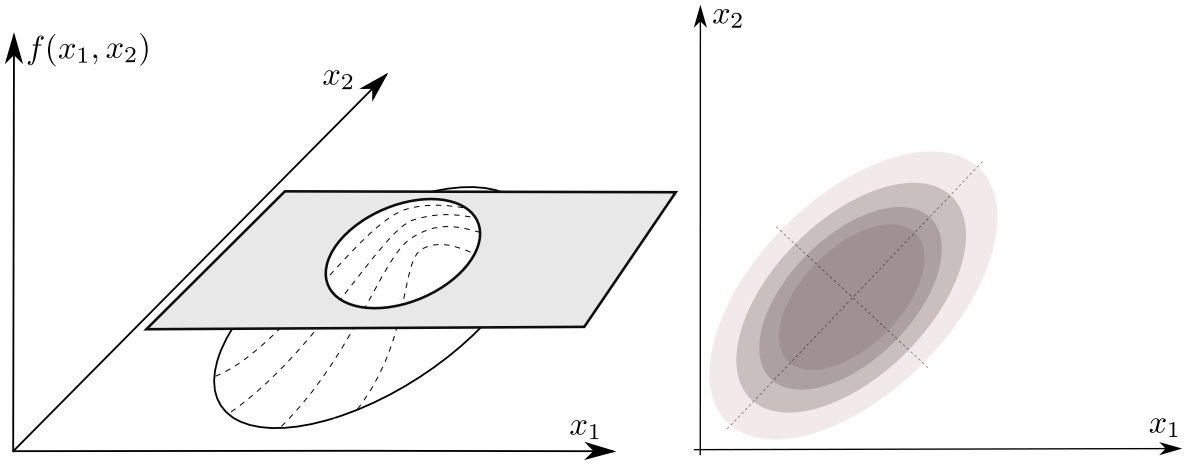

Figure 3.5: Graphic representation of a bivariate normal distribution

An arbitrary oriented n-dimensional ellipsoid can be defined in the form of

$$(\mathbf{x} - \mathbf{v})^T \mathbf{A} (\mathbf{x} - \mathbf{v}) \leq \mathbf{1}, \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{3.10}$$

where the ellipsoid is centered at $\mathbf{v} \in \mathbb{R}^n$ and $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{n \times n}$ is a positive-definite matrix. The eigenvectors of $\mathbf{A}$ thus define the principal semi-axes of the ellipsoid and the eigenvalues of $\mathbf{A}(\lambda_i)$ are the reciprocals of the squares of the principal semi-axes, i.e., $s_i = \lambda_i^{-\frac{1}{2}}$.

The following minimization problem for the volume of the ellipsoid (Vol($\mathcal{E}$)) must be solved given a matrix containing p points in n dimensions ($\mathbf{M} \subset \mathbb{R}^{n \times p}$ where the i^{th} point is $\mathbf{M}_i$) and a pre-specified threshold Δ, find the approximate values (vector $\mathbf{v}$ and matrix $\mathbf{A}$) for the optimal ellipsoid as follows:

$$\begin{aligned} \text{minimize} \quad & Vol(\mathcal{E}) \\ \text{with respect to} \quad & \mathbf{v} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{n \times n} \\ \text{subject to} \quad & (\mathbf{M}_i - \mathbf{v})^T \mathbf{A} (\mathbf{M}_i - \mathbf{v}) \leq \mathbf{1} \quad for \quad i = 1, 2, ..., p \end{aligned} \tag{3.11}$$

The Vol($\mathcal{E}$) can be replaced by $k_1 \times \det\left[\mathbf{A}^{-\frac{1}{2}}\right]$ where k_1 is a positive constant. In the Khachiyan approach, the objective is replaced by $\log(\text{Vol}(\mathcal{E})) = -k_2 \times log(\det[\mathbf{A}])$ for convexity purposes. To produce a centrally symmetric ellipsoid, the points will be lifted into a centrally symmetric (n + 1)-dimensional space with coordinates located at the center of the ellipsoid. This lifting requires

creating a new set of points in $\mathbb{R}^{n+1}$ using:

$$\begin{aligned} \tilde{\mathbf{M}}_{2p\times(n+1)} &= \left\{\pm\tilde{\mathbf{M}}_1, \pm\tilde{\mathbf{M}}_2, ..., \pm\tilde{\mathbf{M}}_p\right\} \\ \tilde{\mathbf{M}}_i &= [\mathbf{M}_i \quad 1]^T \quad for \quad i = 1, 2, ..., p. \end{aligned} \tag{3.12}$$

We can then write the Lagrange dual problem of the above minimization problem in the following format

$$\begin{aligned} \text{maximize} \quad & log\left(det\left[\sum_{i=1}^{p} q_i \tilde{\mathbf{M}}_i (\tilde{\mathbf{M}_i})^T\right]\right) \\ \text{with respect to} \quad & \mathbf{q} \in \mathbb{R}^p \\ \text{subject to} \quad & \mathbf{e^T q} = 1, \quad \mathbf{q} \geq 0, \end{aligned} \tag{3.13}$$

where $\mathbf{e}$ is the unit vector and $\mathbf{q}$ is the unknown variable. With the optimality condition (KKT condition) simplified to solve the following complementary slackness condition:

$$\omega_i(\mathbf{q}^*) = n \quad for \quad i = 1, 2, ..., p \tag{3.14}$$

where, $\omega_i(\mathbf{q}) = (\tilde{\mathbf{M}}_i)^T \left(\sum_{i=1}^{p} q_i \tilde{\mathbf{M}}_i (\tilde{\mathbf{M}}_i)^T\right)^{-1} \tilde{\mathbf{M}}_i$.

The above algorithm satisfies the so-called Δ-relaxed optimality conditions (approximate $\mathbf{q}^*$ by $\tilde{\mathbf{q}}$) defined by:

$$(1 - \Delta)n \leq \omega_i(\tilde{\mathbf{q}}) \leq (1 + \Delta)n \tag{3.15}$$

Therefore, such a solution satisfies a very weak approximation form of the complementary slackness conditions. Finally, the algorithm starts with a feasible solution ($\tilde{\mathbf{q}} > 0$) and improves upon the objective function value by increasing only one component of $\tilde{\mathbf{q}}$ at each iteration, followed by rescaling to regain feasibility.

Once the approximate solution ($\mathbf{q}^*$) is found the solution to the primal problem can be found using:

$$\mathbf{A}^* = \frac{1}{n}\left(\tilde{\mathbf{M}}\mathbf{Q}^*\tilde{\mathbf{M}}^T - (\tilde{\mathbf{M}}\mathbf{q}^*)(\tilde{\mathbf{M}}\mathbf{q}^*)^T\right)^{-1} \tag{3.16}$$

where, $\mathbf{Q} = diag(\mathbf{q})$. The center of the ellipsoid is then:

$$\mathbf{v}^* = \mathbf{A}\mathbf{q}^* \tag{3.17}$$

Algorithm 1 Minimum Volume Elliposoid

procedure CONSTRUCTS A SEQUENCE OF ELLIPSOIDS
input set of p points $\mathbf{M} \subset \mathbb{R}^n$ where the *ith* point is $\mathbf{M}_i$ and threshold $\Delta > 0$
set $k \leftarrow 0$, $\mathbf{q}^0 \leftarrow (1/p)e$, and $\tilde{\mathbf{M}}_i \leftarrow (\mathbf{M}_i, 1)^T \quad$ for $\quad i = 1, 2, .., p$
while $\mathbf{q}^{(k)}$ does not satisfy Eq. 3.15, do
loop
$j \leftarrow argmax_{(i=1,...,p)} (\mathbf{M}_i)^T \left(\sum_{i=1}^{p} q_i^k \tilde{\mathbf{M}}_i (\tilde{\mathbf{M}}_i)^T \right)^{-1} (\mathbf{M}_i)$
$\kappa \leftarrow (\mathbf{M}_j)^T \left(\sum_{i=1}^{p} q_i^k \tilde{\mathbf{M}}_i (\tilde{\mathbf{M}}_i)^T \right)^{-1} (\mathbf{M}_j)$
$\beta \leftarrow \frac{\kappa-(n+1)}{(n+1)(\kappa-1)}$
$\mathbf{q}^{k+1} \leftarrow (1-\beta)\mathbf{q}^k + \beta e^j, k \leftarrow k+1$
end loop

The pseudo-code of the above algorithm that satisfies Eq. 3.15 can be summarized as follows

This algorithm was used to find the 17-dimensional and 5-dimensional enclosing ellipsoids in the primary space and pseudo space of the data set, respectively.

Figure 3.6 indicates a summary of the methodology for producing the proposed framework. The flowchart starts with creating a dataset using laboratory data and continues with the pre-processing step of input variable qualification followed by model development over the defined n-dimensional input space. The obtained models will be used in further performance prediction, optimization problems, etc.

3.4 Results and discussion

3.4.1 Performance measurement

Results of the proposed methods are presented in this section, and their capability and accuracy in predicting rutting behavior of asphalt mixtures over the presented effective variable space are evaluated and compared. The problem of determining optimal design parameters that minimize the amount of accumulated strain over the effective variable space is defined and solved. The obtained optimal design parameters can be used as a first step in the design procedure.

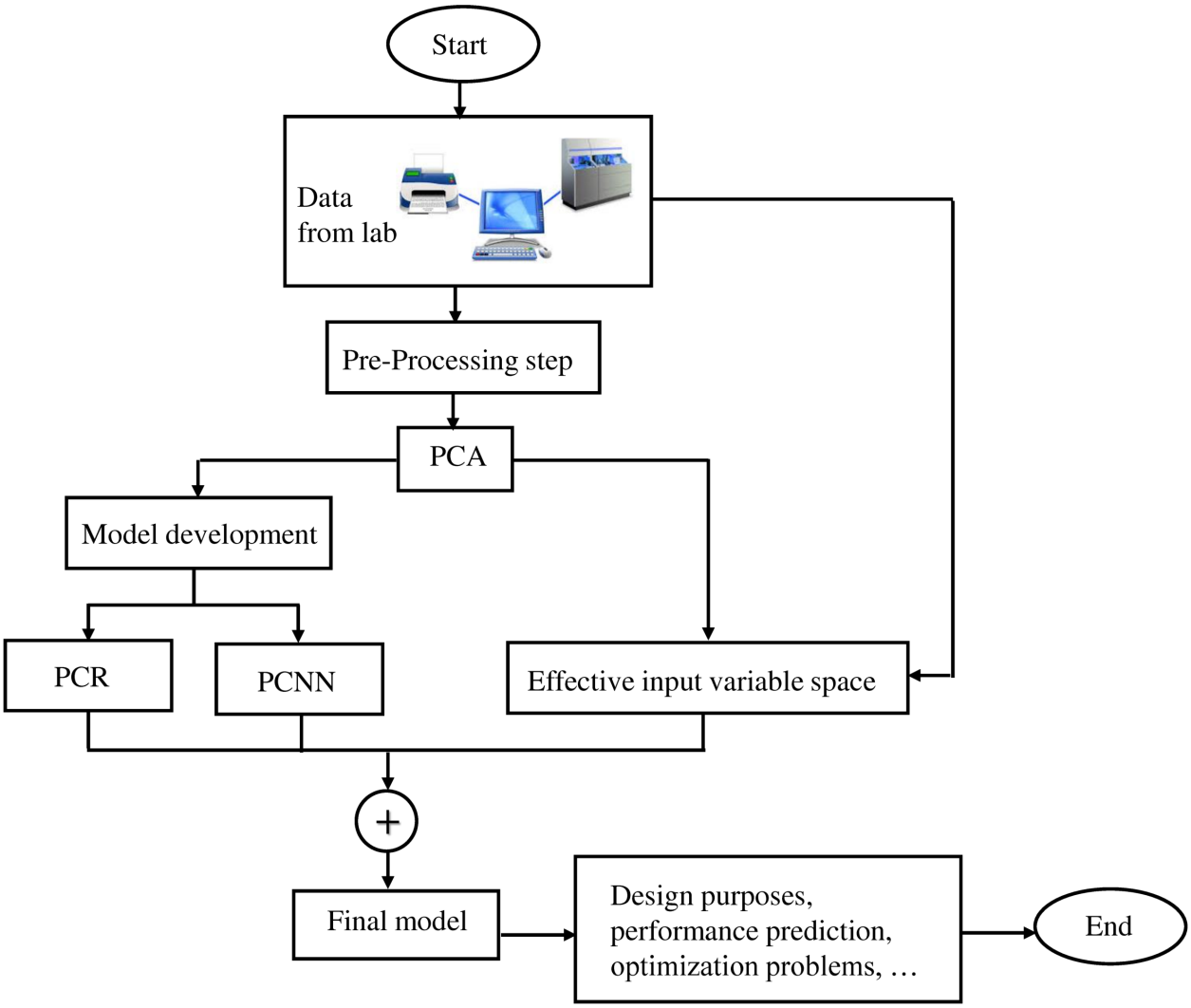

Figure 3.6: A summary of the proposed machine learning-based framework

Table 3.5: Linear Regression Coefficients for Equation 3.18

Coefficient	Value
c_0	1.64×10^4
c_1	-8.89×10^2
c_2	-1.24×10^3
c_3	-1.24×10^3
c_4	-92.41
c_5	6.55×10^2
c_6	1.58×10^2
c_7	-4.35×10^2
c_8	2.35×10^2
c_9	3.74×10^2
c_{10}	-8.28×10^2

The best second-order quadratic linear regression model fitting the measured response, y, is presented in Eq. 3.18:

$$\begin{aligned}\hat{y} = c_0 + c_1 \times pc_1 + c_2 \times pc_2 + c_3 \times pc_3 + c_4 \times pc_4 + c_5 \times pc_5 + c_6 \times pc_1 \times pc_2 \\ + c_7 \times pc_2 \times pc_4 + c_8 \times pc_1 \times pc_3 + c_9 \times pc_2 \times pc_3 + c_{10} \times pc_3 \times pc_5\end{aligned} \tag{3.18}$$

where, c_i's for $i = 0, \ldots, 10$ are presented in Table 3.5.

As mentioned previously, the successfully trained three-layer ANN can be presented as in Eq. 3.8. For ease of use and wider reproduction, the connection weights and biases are presented using the following matrices:

$$W^T = \begin{bmatrix} -0.447 & 1.702 & -0.811 & -0.854 & -1.295 \\ 0.010 & -1.462 & 0.280 & 0.114 & -1.926 \\ 1.315 & -1.391 & -0.028 & -0.218 & 0.895 \\ -0.089 & 0.182 & -0.135 & -0.323 & -0.150 \\ 0.089 & -0.182 & 0.135 & 0.324 & 0.150 \\ -0.613 & 0.255 & -1.190 & -0.808 & 0.120 \\ 0.089 & -0.182 & 0.135 & 0.324 & 0.150 \\ 0.443 & -1.496 & -0.414 & 1.051 & -0.641 \\ 0.280 & -0.625 & 0.831 & 0.736 & 1.158 \\ -0.007 & 0.168 & 0.003 & -1.380 & -0.292 \end{bmatrix}$$

$$W_H = \begin{bmatrix} 0.673 \\ 0.348 \\ -1.250 \\ -0.294 \\ 0.294 \\ -1.713 \\ 0.294 \\ -0.262 \\ 0.974 \\ 0.986 \end{bmatrix}, \quad B_{\mathrm{H}} = \begin{bmatrix} -0.084 \\ -1.208 \\ -1.742 \\ -1.528 \\ 0.567 \\ -0.567 \\ -1.320 \\ -0.567 \\ 1.464 \\ -1.750 \end{bmatrix}, \quad B_0 = [1.334]$$

Performance results of the PCR and PCNN models are given based on the following statistics, and presented in Table 3.6. The first statistic is the "average error (AE)" defined as

$$AE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i). \tag{3.19}$$

AE is an estimate of systematic model bias, n is the number of input vectors, y_i is the i^{th} measured response value, and $\hat{y}_i$ is the i^{th} fitted response value.

The second statistical component is the "average absolute error (AAE)" defined as

$$AAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|. \tag{3.20}$$

This statistic gives the average closeness of the fitted value to the measured response value.

The third statistical component, r_{fit}, is the correlation of y_i and $\hat{y}_i$ defined as

$$r_{fit} = \frac{n\sum_{i=1}^{n} y_i\hat{y}_i - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} \hat{y}_i)}{\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}\sqrt{n\sum_{i=1}^{n} \hat{y}_i^2 - (\sum_{i=1}^{n} \hat{y}_i)^2}} \tag{3.21}$$

The better the fit, the higher r_{fit} will be, with a maximum possible value of 1. The last statistical component is R^2 or the coefficient of determination. In linear regression, for training, R^2 is the portion of the variation explained by the fitted model. It is only applicable to the PCR since it is linear in its parameters and not to PCNN that is non-linear in its parameters.

According to the values of r_{fit} presented in Table 3.6, the predicted values of accumulated strain given by PCR and PCNN models have a high correlation with measured values, meaning that both PCR and PCNN modeled the response well. The second fold of PCR has the highest PCR r_{fit}, 0.80, a very good fit for modeling this type of real data in this application. The first fold of the PCNN has the highest PCNN r_{fit} 0.97 is the highest compared to other folds in both methods.

The prediction models used in the flexible pavement design guide can be expressed in the general form of Eq. 3.22 Bashin et al. (2012); ARA (2004)

$$\frac{\epsilon_p}{\epsilon_r} = aT^bN^c \tag{3.22}$$

where ϵ_p is accumulated plastic strain at N repetition of load, ϵ_r is resilient strain of the asphalt material as a function of mix properties, temperature, and time rate of loading, N is number of load repetitions, $T$is pavement temperature, and a, b, and c are unknown model coefficients. Many researchers (see Ref. ARA (2004)) have tried to obtain suitable estimates for the unknown parameters by performing repeated load permanent deformation tests. Leahy's model was able to estimate the accumulated strain with an R^2 of 0.76, with temperature being the most important variable while loading conditions, material type, and mix parameters were less important. Ayres re-analyzed the original Leahy data and added additional laboratory data to develop a predictive

Table 3.6: Statistical Analysis of PCR and PCNN Modeling (na*: not applicable)

		PCR			PCNN		
	Statistics	Fold 1	Fold 2	Fold 3	Fold 1	Fold 2	Fold 3
Training	AE	0	0	0	34.99	-242.99	46.19
	AAE	1497.02	1705.55	1514.59	729.41	1350.87	944.83
	r_{fit}	0.83	0.82	0.85	0.96	0.87	0.94
	R^2	0.69	0.68	0.72	na*	na*	na*
Testing	AE	626.73	-129.91	-226.1	-98.24	149.2	-169.6
	AAE	2007.47	1515.74	2110.64	694.53	719.9	1037.28
	r_{fit}	0.79	0.8	0.73	0.97	0.95	0.92
	R^2	na*	na*	na*	na*	na*	na*

model with fewer input variables and a slightly lower R^2 of 0.725 Bashin et al. (2012); ARA (2004). Although the reported R^2 values for both models seem reasonable, they did not use separate data sets for training and testing possibly resulting in overfitted models. Also, as mentioned previously, the ability of the model in fitting to the empirical data should be expressed in the r_{fit} value and not the R^2 value.

In comparison with results from the literature, the overall performance of both the PCR and PCNN models, as expressed in terms of the three test statistics used in the work, is significantly higher than that of the previous prediction models used in the American Association of State Highway and Transportation Officials (AASHTO) design procedure.

Comparing the two best folds for training and testing stages indicates that although the PCR modeling works well in predicting the response variable, PCNN produced the best results in both training and testing. Since ANNs learn from data examples presented to them and use these data to adjust their weights in an attempt to capture the relationship between the model input variables and the corresponding outputs, ANNs require no prior knowledge about the nature of the relationship between the input/output variables, this is one of the advantages of ANNs compared with most empirical and statistical methods. Although PCNN indicates higher performance in

terms of predicting the response variable, linear regression models are always easier to use so are often preferred by designers.

3.4.2 Model validation

A general regression model for this study can be represented as

$$y_i = f(\mathbf{Z_i}, \theta) + e_i \tag{3.23}$$

where f is the expectation function; θ is the vector of parameters, and e_i is a random error term assumed to be normally distributed with a mean of zero and unknown variance σ^2 for $i = 1, \ldots, n$, where n is the number of input vectors. Violation of these assumptions and model adequacy can be investigated by the examination of residuals as defined by

$$e_i = y_i - \hat{y}_i. \tag{3.24}$$

Through analysis of residuals, many types of model inadequacies and violations of the underlying assumptions can be assessed. If the model is adequate, the residuals should contain no obvious pattern. Checking the normality assumption can be done by constructing a normal probability plot of the residuals. If the underlying error distribution is normal, this plot will resemble a straight line Devore (2011). These assumptions were checked for the best fold in each method, and the plot of residuals for the best folds are presented in Figure 3.7. Since there is no obvious pattern in the residual plot, the assumption of equal variances seems acceptable. The normal probability plots of the residuals for both models are presented in Figure 3.8, and the data points are not too far away from a straight line, so the normality assumption does not appear to be violated. Although the criteria based on squared residuals as SSE, can give rise to misleading results when the initial assumptions about normality or data being independent are violated Marti-Vargas et al. (2013), the residual plot and normal probability plot of the residuals indicate that these assumptions in this modeling problem hold well without concern.

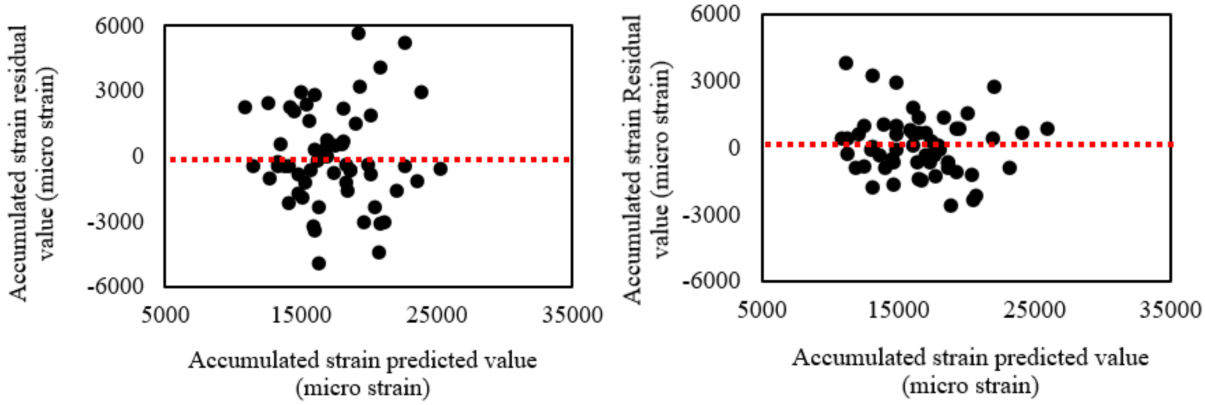

Figure 3.7: Plot of the residuals for the best fold of (a) PCR and (b) PCNN models

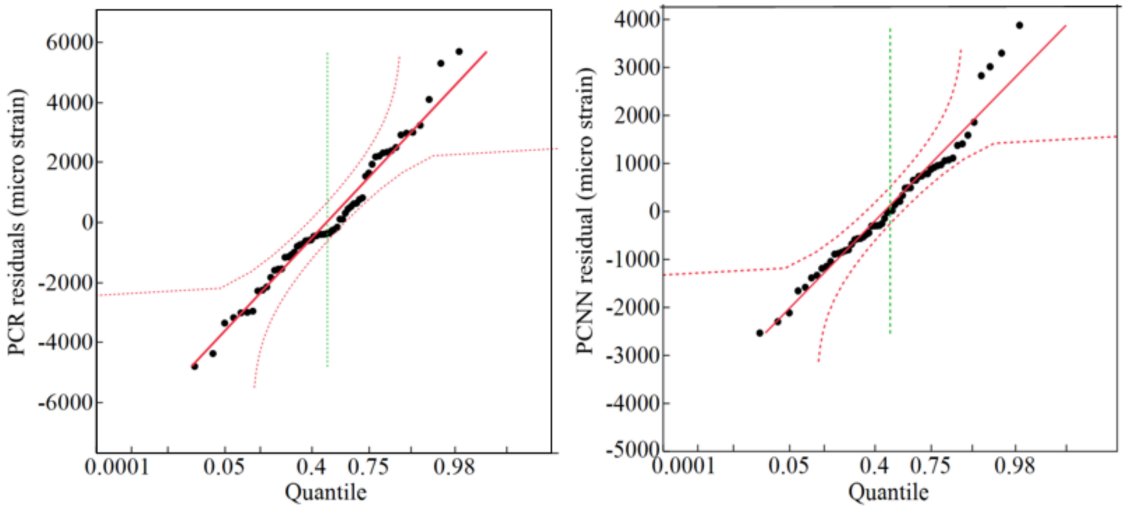

Figure 3.8: Normal probability plot.

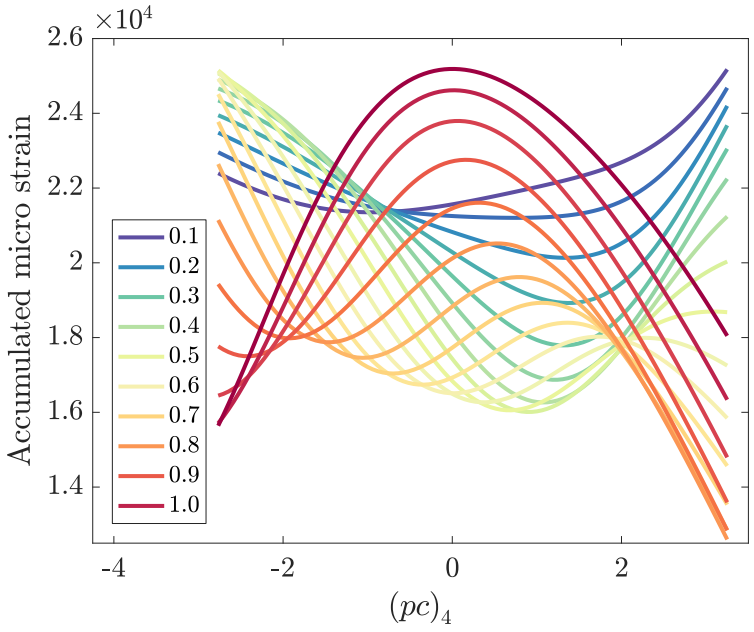

Figure 3.9: Non-linear nature of PCNN function.

3.4.3 Application of the framework: Optimal design

3.4.3.1 Description of the problem and optimization algorithm

One application of the above framework is to find the optimal design parameters that minimize the predicted value of accumulated strain. An appropriate optimization algorithm is required to solve this problem reliably, and the inter-connected functions in ANN create a nonlinear objective function that resemble the properties of a non-convex function leading to multiple local minima. Figure 3.9 depicts the non-linear nature of the function (PCNN). In this figure accumulated strain is plotted against pc_4 while, other PCs are changing over their ranges. On the other hand, the effective variable space is restricted by the bounds defined using high-dimensional ellipsoids (for original variables and principal component space variables) and this will add to the complexity of the problem by introducing a number of constraints to the optimization problem.

Given the various aspects of the above problem, gradient-based optimization algorithms are not useful, while a population-based evolutionary search algorithm represents a potentially active area of the research. A number of novel techniques have been employed in recent years to solve

complex real-world problems Talatahari et al. (2015); Rahami et al. (2011b); Aslani et al. (2010); Ebrahimi et al. (2017). Elrich, et al. Rueda and Erlich (2015); Erlich et al. (2010), developed a new optimization algorithm, called Mean-Variance Mapping Optimization (MVMO) that preserves an archive of the best points that implicitly resemble elitism. This archive of best solutions provides guidelines for a mapping algorithm based on the mean and variance of the archive on which the mapping function depends. MVMO has the ability to carry out a global search using the best-so-far solutions. It is initially designed to work with a single-solution on a normalized domain. While there exist a vast number of optimization algorithms in the literature, in this study a constrained version of MVMO algorithm is implemented to minimize the amount of accumulated strain in asphalt pavements. Although, MVMO is an algorithm for unconstrained problems, Aslani et al. Aslani et al. (2018) showed that constrained MVMO is an attractive tool for engineering-based optimization problems with an improved convergence rate. The primary interest in the MVMO algorithm, for this particular problem, is based on the rigorous numerical comparisons with some basic and enhanced evolutionary algorithms which have shown that MVMO exhibits a better performance, especially in terms of convergence speed (see Refs.Rueda and Erlich (2015); Erlich et al. (2010); Aslani et al. (2018).)

In the present study, the constraints are the effective space for input and PCA variables. If a solution point (an extrapolated point outside the enclosing ellipsoid) violates the constraint the penalty is applied through an adaptive quadratic function. The design optimization problem that minimizes the accumulated strain in asphalt pavements is introduced as follows:

$$\begin{aligned} \text{minimize} \quad & \epsilon = F_{ANN}(\mathbf{x}) \\ \text{with respect to} \quad & \mathbf{x} = (x_1, ..., x_{17}) \\ \text{subject to} \quad & (\mathbf{x} - \mathbf{v})^T \mathbf{A} (\mathbf{x} - \mathbf{v}) \leq \mathbf{1}, \\ & (\mathbf{x}_{pca} - \mathbf{v}')^T \mathbf{A}' (\mathbf{x}_{pca} - \mathbf{v}') \leq \mathbf{1}, \end{aligned} \tag{3.25}$$

where ϵ is the accumulated strain; $\mathbf{x}$ is the vector of variables, and $(\mathbf{x} - \mathbf{v})^T \mathbf{A} (\mathbf{x} - \mathbf{v}) \leq \mathbf{1}$ is the enclosing ellipsoid constraint equation for the original and mapped variables. A widely-practiced approach is to use a penalty function that penalizes (increases, in the case of the minimization

Table 3.7: User-defined parameters in MVMO single-solution

Parameter and Description	Value
Archive Size ($n_{archive}$)	6
Factor (f_s)	(Adaptive, $f_s^{start} = 0.8$, $f_s^{final} = 2$)
Increment (Δd_i)	(Adaptive, $\Delta d_i^{start} = 0.01$, $\Delta d_i^{final} = 0.4$)
Number of mapped variables (m_{mapped})	6
No. of Iterations (Convergence criterion)	30,000

problem) the objective value for each constraint in proportion to its degree of closeness/violation of the corresponding constraint. Similar to other algorithms, MVMO contains a number of parameters which could potentially alter its convergence behavior. In this study, adaptive strategies are used for control parameters including shape factor (f_s) and the local increment size (Δd_i) in order to mimic exploration to exploitation behavior automatically. A number of initial tests are performed to select the initial and final values of these parameters. Table 3.7 contains the parameters used in this study for a single-solution MVMO algorithm. A larger archive size and number of mapped variables will make the search more intensive, however, it will inversely affect the performance of the algorithm in terms of computational resources. Initial tests on this problem revealed that once the number of archive size and mapped variables are above 6, then the final results converge and the problem is independent of these control parameters.

It should be noted that given the uncertainty in the parameters of the ellipsoid, the above penalty function is also prone to an uncertainty. If the threshold in Eq. 3.15 is Δ, then it can be shown that for small values of Δ, the penalty function can be modified as follows:

$$(\mathbf{x} - \mathbf{v})^T \mathbf{A} (\mathbf{x} - \mathbf{v}) \leq \mathbf{1} - |\Delta| max(s_i) \tag{3.26}$$

where $s_i = \lambda_i^{-\frac{1}{2}}$, is the principal semi-axes of the ellipsoid which will be different for original and pseudo inputs.

Figure 3.10 shows the convergence history for the above problem solved using a constrained MVMO algorithm. The algorithm starts with a random initial point and the objective function decreases as the iteration evolves. MVMO uses adaptive strategies to mimic exploration and ex-

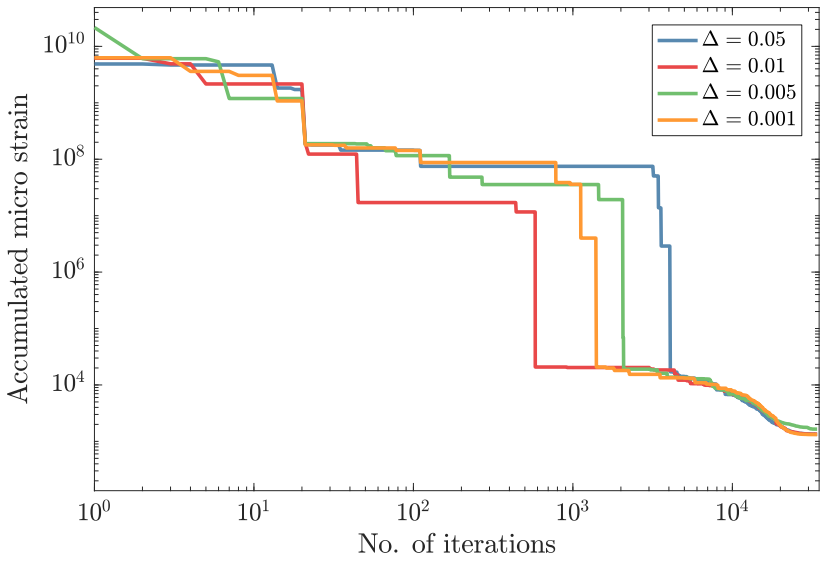

Figure 3.10: Convergence study

ploitation strategies during optimization. MVMO is initiated with a random point and it is terminated after executing 30,000 function evaluations. Multiple experiments have shown that the above stopping criterion is sufficient for the algorithm to converge (when the change in the solution is not significant). The convergence plot shows that the algorithm converges to virtually similar results when Δ varies from 0.001 to 0.05. It should be noted that using relatively small values of Δ could make the matrices in Algorithm 1 ill-conditioned.

Solving the optimization problem using PCNN predictive model over the available effective variables space produces a minimum value of accumulated strain of 1772 micro strain with corresponding material properties presented in Table 3.8. The constrained MVMO was used once for PCR and a minimum value of accumulated strain of 4181 micro strain was obtained. Corresponding design variables were also obtained but, since the PCNN produced better accuracy, those results are not reported in this section. The constrained single-solution MVMO is implemented in MATLAB and a serial simulation (calculation on a single core Intel processor with 2.60 GHz Clock Speed) with 30,000 function evaluations takes around 400 seconds on and takes roughly 1.1 GB of memory.

3.4.3.2 Statistical description of solution and global optimum estimation

In order to verify the quality of the solution obtained using the optimization algorithm, 1000 MVMO simulations (each one initiated with a random point) are conducted. These tests are performed following the parameters described for MVMO (see Section 3.4.3) and the threshold parameter Δ is 0.005. Figure 3.11 (left) shows the accumulated strain (in micro strain unit) obtained from each run. The statistical description of the 1000 runs is the following: the maximum and minimum values are 1787.1 and 1765, respectively; the sample mean value is 1773; the standard deviation of the sample is 4.19. A statistical approach is employed following Refs. LOMBARDERO et al. (2012); Paya-Zaforteza et al. (2010) where extreme value theory based on Weibul distribution is implemented to estimate the global optimum of the solution obtained by MVMO. The fitted Weibul cumulative distribution function (cdf) parameters are $\eta = 11.1212$ (scale parameter) $\beta = 2.570046$ (shape parameter), and $\gamma = 1763.21$ (location parameters). The fitted cdf is plotted in Figure 3.11 (right) and compared to the experimental data obtained from MVMO. The difference between the estimated global minimum and the minimum obtained from MVMO is only 0.1% which is small enough given the empirical nature of the predictive model and the application of the current framework. Thus, the proposed optimal design (which successfully satisfy the design consideration as discussed in the next section) can be used confidently as a sound starting point in flexible pavement design problem.

3.4.3.3 Design considerations and discussion

Although there is no design specification for the amount of asphalt binder to be used in the mixture, 4% by weight of total mix is reasonable for a nominal maximum aggregate size (NMAS) of 19 mm (3/4"). The percentage of aggregate passed by sieve size and located in the allowable ranges of a typical gradation specification is presented in Figure 3.12 (The obtained aggregate gradation is shown by the red line). The obtained percentage air void is 4%, the target value of the design specification. Voids in mineral aggregate (VMA) for NMAS of 19 mm (3/8") should be greater than 13%, a condensation that was satisfied. The acceptable range for voids filled with asphalt

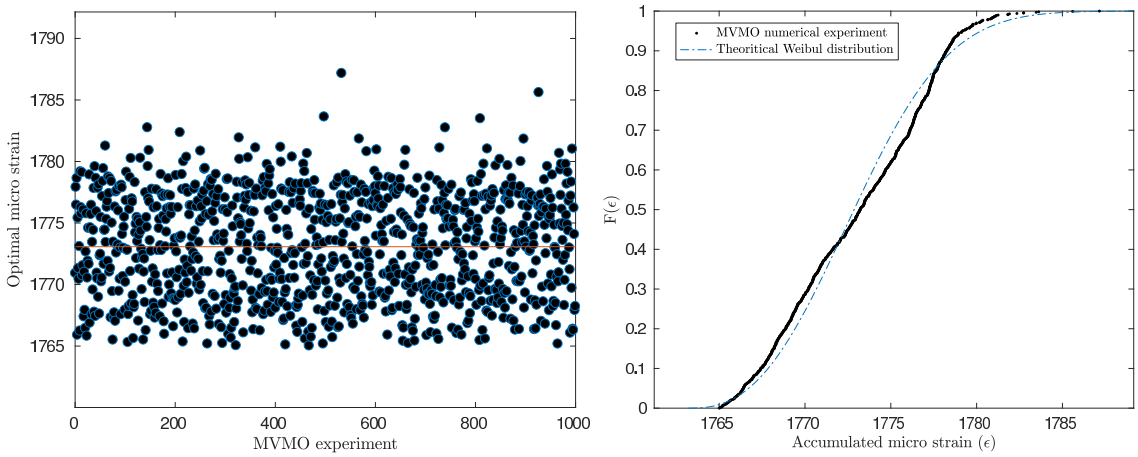

Figure 3.11: The results for 1000 runs

(VFA) depends on the amount of traffic, and for traffic loading less than 0.3 million Equivalent Single Axel Load(ESALs), VFA should be in a range between 70 and 80, for traffic loading between 0.3 and 3 million ESALs, the range for VFA should be 65 to 78, and for traffic loading greater than 3 million of ESALs, the range should be 65 to 75. The obtained VFA value of 76% is satisfied for all the loading categories.

Recalling the PCNN's higher prediction capability, the offered optimum design parameter based on PCNN would be more reliable as a design starting point. It is also worth pointing out that the obtained optimum design parameters and minimal accumulated strain are based on the available empirical database. Eventually, for creating a comprehensive predictive model, a large empirical database is required for obtaining more reliable and applicable results. The purpose of this study was to provide and introduce a framework that could be reproducible and easy to use for every database.

3.5 Conclusions and recommendations

This study used the experimental data of permanent deformation of asphalt mixtures and focused on evaluation and qualification of input variables to be used in further modeling. Cross-correlated input variables are identified using correlation analysis and substituted by orthogonal pseudo-inputs using a dimensionality reduction technique called PCA. This work compared mul-

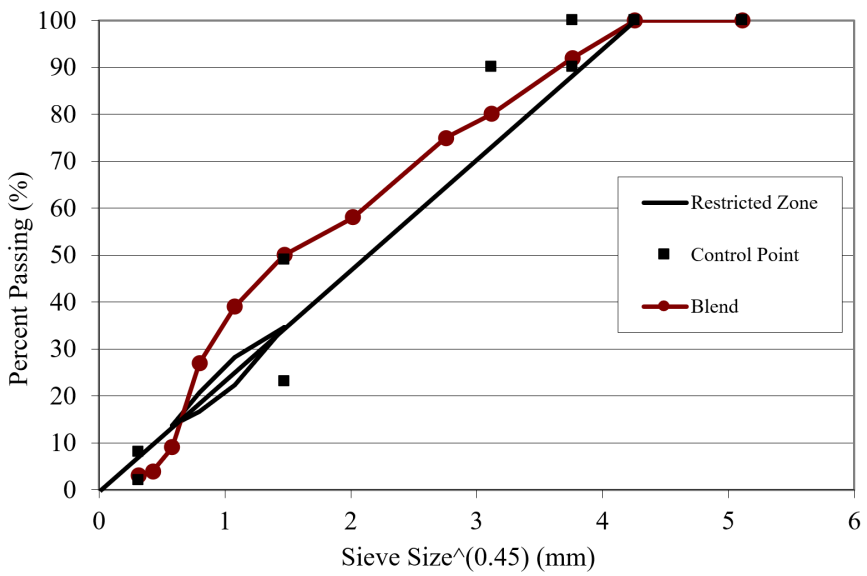

Figure 3.12: Optimal aggregate gradation graph

Table 3.8: Optimal design parameters

Variable	Identity	Optimal values from PCNN	Current design specification			
			Control points		Restricted zone	
			Lower	Upper	Lower	Upper
x_1	Binder %	4	-	-	-	-
x_2	G* (Pa)	270190	-	-	-	-
x_3	NMAS	19	-	-	-	-
x_4	Passing 3/4”	92	90	100	-	-
x_5	Passing 1/2”	66	-	90	-	-
x_6	Passing 3/8”	65	-	-	-	-
x_7	Passing #4	58	-	-	-	-
x_8	Passing #8	50	23	49	34.6	34.6
x_9	Passing #16	39	-	-	22.3	28.3
x_{10}	Passing #30	27	-	-	16.7	20.7
x_{11}	Passing #50	9	-	-	13.7	13.7
x_{12}	Passing #100	4	-	-	-	-
x_{13}	Passing #200	3	2	8	-	-
x_{14}	VMA	16	13	-	-	-
x_{15}	VFA	76	65	80	-	-
x_{16}	Va%	4	4	4	-	
x_{17}	E*(Mpa)	713	-	-	-	-

tivariate regression and ANN modeling of a small set of pseudo-variables determined from PCA (PCR and PCNN, respectively). Both proposed approaches modeled the amount of permanent deformation satisfactorily with PCNN fitting the test data significantly better. A simple method is developed to characterize the effective variable space because empirical-based predictive models can lead to spurious values under extrapolation. The obtained n-dimensional enclosing ellipsoid is defined as the effective variable space in which both predictive models can be used while guard against extrapolation. Nonetheless, both approaches showed better performance as modeling tools than other regression-based approaches that use standard variables in the AASHTO design procedure. These PCA-based approaches are therefore strongly recommended as sound modeling approaches in this application. Moreover, these methodologies appear to also have much promise in modeling other material properties at other effective temperatures and such investigations are recommended as future studies. To present one of this framework's applications, an optimization problem of finding minimum accumulated strain over the effective variable space is defined and solved. The value of 1772 micro-strain is obtained as the minimum accumulated strain and the design parameters corresponding to this minimum value are calculated for PCNN. The results meet the flexible pavement design specifications and could be used as a sound starting point in the design procedure. The proposed approach is applicable for predictive modeling problems with correlated inputs and when seeking to use the model to map back to input space, without extrapolation, to find an input vector for near optimal response. Future research would address the scalability and efficiency of the method in other modeling problems, particularly for larger data sets.

3.6 References

AASHTO TP 79-13 (2013). AASHTO TP 79-13 Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Aphalt Mixtures Performance Tester. *Am. Assoc. State Highw. Transp. Off.*, pages 1–19.

Apeagyei, A. K. (2011). Rutting as a Function of Dynamic Modulus and Gradation. *Am. Soc. Civ. Eng.*, 23 (9):1302–1310.

ARA, Inc., E. C. D. (2004). Guide for mechanistic–empirical design of new and rehabilitated pavement structures. *Final Rep., Nchrp Proj. 1-37a*.

Arabali, P., Sakhaeifar, M. S., Freeman, T. J., Wilson, B. T., and Borowiec, J. D. (2017). Decision-making guideline for preservation of flexible pavements in general aviation airport management. *J. Transp. Eng. Part B: Pavements*, 143(2):04017006.

Aslani, M., Asla, R. N., Oftadehb, R., and Panahic, M. S. (2010). A novel hybrid simplex-genetic algorithm for the optimum design of truss structures. In *Proceedings of the world congress on engineering*, volume 2.

Aslani, M., Ghasemi, P., and Gandomi, A. H. (2018). Constrained mean-variance mapping optimization for truss optimization problems. *Struct. Des. Tall Spec. Build.*, 27(6):e1449.

Bashin, A., Masad, E., Kutay, M. E., Buttlar, W., Kim, Y.-R., Marasteanu, M., Kim, Y. R., Schwartz, C. W., and Carvalho, R. L. (2012). Applications of advanced models to understand behavior and performance of asphalt mixtures. *Transp. Res. E-Circ.*, 1(E-C161).

Birgisson, B., Roque, R., Kim, J., and Pham, L. V. (2004). The Use of Complex Modulus To Characterize Performance of Asphalt Mixtures and Pavements in Florida. Technical Report September 2004, Florida Department of Transportation, Florida Department of Transportation.

Carpenter, W. and Barthelemy, J.-F. (1993). A comparison of polynomial approximations and artificial neural nets as response surfaces. *Struct. Multidiscip. Optim.*, 5(3):166–174.

Cheng, B. and Titterington, D. (1994). Neural networks: A review from a statistical perspective.

Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences*. Cengage learning.

Ebrahimi, M., ShafieiBavani, E., Wong, R. K., Fong, S., and Fiaidhi, J. (2017). An adaptive meta-heuristic search for the internet of things. *Future Generation Computer Systems*, 76:486–494.

Erlich, I., Venayagamoorthy, G. K., and Worawat, N. (2010). A mean-variance optimization algorithm. *IEEE Congr. Evol. Comput.*, pages 1–6.

Fodor, I. K. (2002). A Survey of Dimension Reduction Techniques. Technical report, U.S. Department of Energy, U.S. Department of Energy.

García-Segura, T., Yepes, V., and Frangopol, D. M. (2017). Multi-objective design of post-tensioned concrete road bridges using artificial neural networks. *Struct. Multidiscip. Optim.*, 56(1):139–150.

Ghasemi, P., Aslani, M., Rollins, D. K., Williams, R. C., and Schaefer, V. R. (2018). Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks. *Int. J. Pavement Res. Technol.*

Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., and Maggio, R. M. (2017). Use of principal component analysis (pca) and hierarchical cluster analysis (hca) for multivariate association

between bioactive compounds and functional properties in foods: A critical perspective. *Trends Food Sci. & Technol.*

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J. Am. Stat. Assoc.*, pages 1–11.

Hargrove, L., Scheme, E., Englehart, K., and Hudgins, B. (2017). Principal components analysis tuning for improved myoelectric control. *CMBES Proceedings*, 30(1).

Johnson, R. A., Wichern, D. W., et al. (2014). *Applied multivariate statistical analysis*, volume 4. Prentice-Hall New Jersey.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.*, 30(3):487.

Juang, C. H. and Chen, C. J. (1999). CPT-Based Liquefaction Evaluation Using Artificial Neural Networks. *Comput. -Aided Civ. Infrastruct. Eng.*, 14(3):221–229.

K Rollins Sr, D. (2015). A One-Dimensional PCA Approach for Classifying Imbalanced Data. *J. Comput. Sci. & Systems Biology*, 8(1):5–11.

Kaloush, K. E., Witczak, M., and Sullivan, B. W. (2003). Simple performance test for permanent deformation evaluation of asphalt mixtures. *Sixth Int. RILEM Symp. Perform. Test. Eval. Bitum. Mater.*, pages 498–505.

Kartam, N. (1994). Neural Netwroks in Civil Engineering: Systems and Application. *J. Comput. Civ. Eng.*, 8(2):149–162.

Kutner, M. H., Nachtsheim, C., and Neter, J. (2004). *Applied linear regression models*. McGraw-Hill/Irwin.

Kvasnak, A., Robinette, C., and Williams, R. (2007). A Statistical Development of a Flow Number Predictive Equation for the Mechanistic-Empirical Pavement Design Guide.CD-ROM . In Transportation Research Board, Washington, D., editor, *Transportation Research Board, Washington, DC.*, volume 3, pages 1–18, Transportation Research Board, Washington, DC. DC.Transportation Research Board, Washington.

Lever, J., Krzywinski, M., and Altman, N. (2017). Points of significance: Principal component analysis.

Liew, K., Tan, H., Ray, T., and Tan, M. (2004). Optimal process design of sheet metal forming for minimum springback via an integrated neural network evolutionary algorithm. *Struct. Multidiscip. Optim.*, 26(3):284–294.

Lim, J., Jang, Y. S., Chang, H. S., Park, J. C., and Lee, J. (2018). Role of multi-response principal component analysis in reliability-based robust design optimization: An application to commercial vehicle design. *Struct. Multidiscip. Optim.*, pages 1–12.

LOMBARDERO, A. C., Piqueras, V. Y., and Vidosa, F. G. (2012). Automatic design of concrete vaults using iterated local search and extreme value estimation. In *Latin American Journal of Solids and Structures*, volume 9, pages 675–689. Argentinean Association of Computational Mechanics, Brazilian Association of Computational Mechanics, Mexican Association of Numerical Methods in Engineering and Applied Sciences.

Malik, M. R., Isaac, B. J., Coussement, A., Smith, P. J., and Parente, A. (2018). Principal component analysis coupled with nonlinear regression for chemistry reduction. *Combust. Flame*, 187:30–41.

Marti-Vargas, J. R., Ferri, F. J., and Yepes, V. (2013). Prediction of the transfer length of prestressing strands with neural networks. *Comput. Concr.*, 12(2):187–209.

MATLAB (2012). Matheworks, T., 2012. Matlab 2012b. ,.

May, R., Dandy, G., and Maier, H. (2011). Review of input variable selection methods for artificial neural networks. In *Artificial neural networks-methodological advances and biomedical applications*. InTech.

Nicholas, P. E., Padmanaban, K., Vasudevan, D., and Ramachandran, T. (2015). Stacking sequence optimization of horizontal axis wind turbine blade using fea, ann and ga. *Struct. Multidiscip. Optim.*, 52(4):791–801.

Noh, Y., Choi, K., and Du, L. (2009). Reliability-based design optimization of problems with correlated input variables using a gaussian copula. *Struct. Multidiscip. Optim.*, 38(1):1–16.

Paya-Zaforteza, I., Yepes, V., González-Vidosa, F., and Hospitaler, A. (2010). On the weibull cost estimation of building frames designed by simulated annealing. *Meccanica*, 45(5):693–704.

Pellinen, T. and Witczak, M. (2002). Use of Stiffness of Hot-Mix Asphalt as a Simple Performance Test. *Transp. Res. Rec. : J. Transp. Res. Board*, 1789(02):80–90.

Rahami, H., Kaveh, A., Aslani, M., and Najian Asl, R. (2011). A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization. *Iran University of Science & Technology*, 1(1):29–46.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer.

Rodezno, M., Kaloush, K., and Corrigan, M. (2010). Development of a Flow Number Predictive Model. *Transp. Res. Rec. : J. Transp. Res. Board*, 2181(2181):79–87.

Rollins, D. K., Zhai, D., Joe, A. L., Guidarelli, J. W., Murarka, A., and Gonzalez, R. (2006). A novel data mining method to identify assay-specific signatures in functional genomic studies. *BMC bioinformatics*, 7:377.

Rueda, J. L. and Erlich, I. (2015). Testing MVMO on learning-based real-parameter single objective benchmark optimization problems. *2015 IEEE Congr. Evol. Comput. CEC 2015 - Proc.*, pages 1025–1032.

Su, C.-T. and Tong, L.-I. (1997). Multi-response robust design by principal component analysis. *Total Qual. Manag.*, 8(6):409–416.

Talatahari, S., Gandomi, A. H., Yang, X.-S., and Deb, S. (2015). Optimum design of frame structures using the eagle strategy with differential evolution. *Engineering Structures*, 91:16–25.

Timm, D., West, R., Priest, A., Powell, B., Selvaraj, I., Zhang, J., and Brown, R. (2006). Phase ii ncat test track results. *NCAT Report*, 6(05).

Todd, M. J. and Yıldırım, E. A. (2007). On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids. *Discret. Appl. Math.*, 155(13):1731–1744.

Witzcak, M. W. (2002). *Simple performance test for superpave mix design*, volume 465. Transportation Research Board.

Xiong, F., Wang, D., Zhang, S., Cai, K., Wang, S., and Lu, F. (2017). Lightweight optimization of the side structure of automobile body using combined grey relational and principal component analysis. *Struct. Multidiscip. Optim.*, 57:441–461.

Xiong, F., Wang, D., Zhang, S., Cai, K., Wang, S., and Lu, F. (2018). Lightweight optimization of the side structure of automobile body using combined grey relational and principal component analysis. *Struct. Multidiscip. Optim.*, 57(1):441–461.

Yi, S., Lai, Z., He, Z., Cheung, Y.-m., and Liu, Y. (2017). Joint sparse principal component analysis. *Pattern Recognition*, 61:524–536.

Zhang, J., Yang, J., and Kim, Y. R. (2015). Characterization of mechanical behavior of asphalt mixtures under partial triaxial compression test. *Constr. Build. Mater.*, 79:136–144.

CHAPTER 4. PRINCIPAL COMPONENT NEURAL NETWORKS FOR MODELING, PREDICTION, AND OPTIMIZATION OF HOT MIX ASPHALT DYNAMICS MODULUS

A paper published at the Journal of Infrastructure, MDPI

4.1 Abstract

The dynamic modulus of hot mix asphalt (HMA) is a fundamental material property that defines the stress-strain relationship based on viscoelastic principles and is a function of HMA properties, loading rate, and temperature. Because of the large number of efficacious predictors (factors) and their nonlinear interrelationships, developing predictive models for dynamic modulus can be a challenging task. In this research, results obtained from a series of laboratory tests including mixture dynamic modulus, aggregate gradation, dynamic shear rheometer (on asphalt binder), and mixture volumetric are used to create a database. The created database is used to develop a model for estimating the dynamic modulus. First, the highly correlated predictor variables are detected, then Principal Component Analysis (PCA) is used to first reduce the problem dimensionality, then to produce a set of orthogonal pseudo-inputs from which two separate predictive models were developed using linear regression analysis and Artificial Neural Networks (ANN). These models are compared to existing predictive models using both statistical analysis and Receiver Operating Characteristic (ROC) Analysis. Empirically-based predictive models can behave differently outside of the convex hull of their input variables space, and it is very risky to use them outside of their input space, so this is not common practice of design engineers. To prevent extrapolation, an input hyper-space is added as a constraint to the model. To demonstrate an application of the proposed framework, it was used to solve design-based optimization problems, in two of which optimal and inverse design are presented and solved using a mean-variance mapping optimization algorithm.

The design parameters satisfy the current design specifications of asphalt pavement and can be used as a first step in solving real-life design problems.

4.2 Introduction

The stress-strain relationship for asphalt mixtures under sinusoidal loading can be described by the dynamic modulus, $|E^*|$, a function of material's components properties, loading rate, and temperature Ghasemi (2018); Birgisson et al. (2004b). The dynamic modulus is one of the primary design inputs in Pavement Mechanistic-Empirical (M-E) Design to describe the fundamental linear viscoelastic material properties Arabali et al. (2017); Bozorgzad et al. (2017); Guide (2004), and is one of the key parameters used to evaluate rutting and fatigue cracking distress predictions in Mechanistic-Empirical Pavement Design Guide (MEPDG) Guide (2004); AASHTO (2008). Although $|E^*|$ has a significant role in pavement design, the associated test procedure is time-consuming and requires expensive equipments, so extensive effort has been extended to predict $|E^*|$ from hot mix asphalt (HMA) material properties Nobakht and Sakhaeifar (2018); Peng et al. (2019); Shu and Huang (2008).

Predictive modeling is a process of estimating outcomes from several predictor variables using data mining tools and probability theory. An initial model can be formulated using either a simple linear equation or a more sophisticated structure obtained through a complex optimization algorithm Devore (2011).

There are several well-known predictive models for dynamic modulus, some of them are regression models, and some more recent ones have used techniques that include Artificial Neural Networks (ANN) and genetic programming El-Badawy et al. (2018). Andrei et al. Andrei et al. (1999), used 205 mixtures with 2750 data points and revised the original Witczak model, and the developed model has subsequently been reformulated to use binder shear modulus rather than binder viscosity Bari and Witczak (2007). Christensen et al. Christensen Jr et al. (2003), developed a new $|E^*|$ predictive model based on the law of mixtures. The data base used for training the model contained 206 $|E^*|$ measurements from 18 different HMA mixtures. Jamrah et al. Jamrah

et al. (2014), attempted to develop improved $|E^*|$ predictive models for HMA used in the State of Michigan. They observed a significant difference between measured and fitted $|E^*|$ values, especially at high temperatures and low frequencies. Alkhateeb et al. Al-Khateeb et al. (2006), developed a new predictive model from the law of mixtures to be used over broader ranges of temperature and loading frequencies, including higher temperatures/lower frequencies. The predictor variables used in that model were Voids in Mineral Aggregate (VMA) and binder shear modulus (G^*).

Sakhaeifar et al. Sakhaeifar et al. (2017), developed individual temperature-based models for predicting dynamic modulus over a wide range of temperatures. The predictor variables used in their model were aggregate gradation, VMA, Voids Filled with Asphalt (VFA), air void (V_a), effective binder content (V_{beff}), G^*, and binder phase angel (δ).

The existing dynamic modulus predictive models in the literature typically use two or more predictors from the following list: aggregate gradation, volumetric properties, and binder shear properties. These predictor variables are not necessarily an independent set of variables and thus it may not be appropriate for use in developing models. Since cross-correlated inputs in a dataset can unfavorably affect the accuracy of a predictive model by unduly affecting the estimation of their causative effects on the response variable, a pre-processing step of data evaluation would be useful for studying the quality of the input variables and their pair-wise correlations Ghasemi et al. (2018b). Principal Component Analysis (PCA) is a multivariate statistical approach that not only reduces the dimensionality of the problem but also converts a set of correlated inputs to a set of orthogonal (pseudo-)inputs using an orthogonal transformation Ren et al. (2019). During such a transformation, PCA maximizes the amount of information of the original dataset $\mathbf{X}$ by using a smaller set of pseudo-variables Fodor (2002); Johnson et al. (2014). Another issue in all of the predictive models is extrapolation that can be risky because a model might behave differently outside of the convex hull that contains all of the data points used for its training. To avoid using points outside of this convex hull, a hyper-space containing all data points can be found and added as a constraint on the desired modeling problem.

Ghasemi et al. Ghasemi et al. (2018a), developed a methodology for eliminating correlated inputs and extrapolation in modeling; they created a laboratory database of accumulated strain values of several asphalt mixtures and used the resulting framework to estimate the amount of permanent deformation (rutting) in asphalt pavement. Following their new PCA-based approach, this study focuses on developing a machine-learning based framework for predicting the dynamic modulus of HMA using orthogonal pseudo-inputs obtained from principal component analysis. Unlike most of the existing $|E^*|$ predictive models, the proposed framework uses different data sets for model training and performance testing. To avoid extrapolation, an n-dimensional hyperspace is developed and added as a constraint to the modeling problem. This study also claims to determine the optimal HMA design and design variables for a pre-specified $|E^*|$ by applying framework using an evolutionary-based optimization algorithm. It is worth pointing out that, unlike other predictive models, the proposed framework is not site-specific and also not limited to the materials used in the American Association of State Highway and Transportation Officials (AASHTO) road test, i.e., this framework can adjust itself based on the dataset presented to the framework. The need for a more robust and general framework for performance prediction in asphalt pavement also stems from the availability of the vast amount of experimental data in this field. In this work, the developed framework operates in such a spirit and improves the accuracy of available models via machine learning-based approaches.

The remainder of the document is organized as follows: Section 4.3 presents material and methodology, followed by Section 4.4 that covers results and discussion. Two examples of the proposed framework's applications are discussed in Section 4.5, followed by conclusions presented in Section 4.6.

4.3 Material and Methodology

Twenty-seven specimens from nine different asphalt mixtures (three replicates for each mixture group) were used in this study. Using AASHTO TP 79-13 the dynamic modulus test was performed at three temperatures (0.4, 17.1, and 33.8 °C) and nine loading frequencies (25, 20, 10, 5, 2, 1, 0.5,

Table 4.1: General mixture properties of nine asphalt mixtures used in this study.

	Mix 1	**Mix 2**	**Mix 3**	**Mix 4**	**Mix 5**	**Mix 6**	**Mix 7**	**Mix 8**	**Mix 9**
Binder performance grade	58–28	58–28	58–28	58–34	58–34	58–34	64–28	64–34	64–28
% V_{beff}	4.20	4.10	4.10	3.90	3.50	4.30	4.20	4.00	4.60
%VMA	13.50	13.50	13.60	13.10	12.50	13.90	13.70	13.40	14.40
% VFA	70.30	70.40	70.60	69.60	68.10	71.20	70.80	70.20	72.30
G_{mb}	2.32	2.31	2.31	2.32	2.31	2.32	2.31	2.32	2.31
G_{mm}	2.41	2.46	2.51	2.48	2.64	2.46	2.48	2.51	2.44
% V_a	4.01	3.99	3.99	3.98	3.98	4.03	4	3.99	3.98
% passing $3/4''$	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
% passing $1/2''$	93.90	96.40	87.20	93.50	95.10	96.40	94.10	94.40	94.20
% passing $3/8''$	77.50	84.60	73.70	76.40	83.10	87.30	83.40	82.00	80.90
% passing #4	49.80	53.10	48.40	52.20	52.20	60.90	63.80	48.20	58.60
% passing #8	34.40	38.40	35.10	43.60	38.80	46.90	47.10	34.90	46.00
% passing #30	16.70	18.70	17.90	20.90	18.80	23.40	21.70	19.20	25.90
% passing #50	10.30	10.80	10.90	11.40	9.90	12.40	11.90	11.80	13.80
% passing #100	6.10	5.90	6.40	5.80	5.40	6.10	6.60	6.10	7.20
% passing #200	3.60	3.30	6.20	3.30	3.50	3.40	4.00	3.10	4.00

0.2, 0.1 Hz). The maximum theoretical specific gravity (G_{mm}), the bulk specific gravity (G_{mb}), and the effective binder content (V_{beff}) were determined and used to calculate other volumetric properties of the asphalt mixtures.Asphalt binder shear properties were obtained from a dynamic shear rheometer (DSR) test. Using ASTM D7552-09(2014) the test was performed over a wide range of temperatures (-10 to 54 °C) and frequencies (0.1 Hz to 25 Hz), the same test temperatures and loading frequencies used in the mixture dynamic modulus test. It is important to note that this study uses a consistent definition of frequency, and that in order to predict the dynamic modulus value of an asphalt mixture for example at 4 °C and 25 Hz, for example, one should use as a model input the complex shear modulus of asphalt binder, $|G^*|$, at 4 °C and 25 Hz. A summary of the nine different mixture properties is given in Table 4.1. Using the laboratory test results on 27 specimens, a database of 243 data points was created for use in further modeling.

4.3.1 Preliminary Processing Step: Input Variable Selection

A parsimonious set of input variables is required to develop a model Fodor (2002). For a common model structure, one can represent the expectation function of the response as $y_i = f_i(\mathbf{x}_i, \theta)$, where y_i is the expected response variable at the ith measurement, $i = 1, \ldots, n$, $\mathbf{x}_i$ is the input vector

at the ith measurement, and θ is the vector of unknown model parameters with $\theta = [\theta_1 \dots \theta_q]^T$. It is assumed that the element in the ith row and jth column of the Jacobian matrix , $\mathbf{J}$, is $\frac{\partial \eta_i}{\partial \theta_j}$ i.e., $\mathbf{J} = \left\{ \frac{\partial \eta_i}{\partial \theta_j} \right\}$. Note that the jth column represents θ_j and its column vector reflects the variation in the response space as θ_j varies over a specific set of experimental conditions. If j and k are two orthogonal columns, their correlation coefficient (r) must be zero, meaning that the information used to estimate θ_j is independent from the information used to estimate θ_k and vice versa. The benefit of using orthogonal input variables is that not only does it result in consolidation of causative effects of inputs on the output but it also maximizes parameter accuracy and therefore estimation accuracy of the predicted output.

According to the literature Andrei et al. (1999); Bari and Witczak (2007); Christensen Jr et al. (2003); Al-Khateeb et al. (2006); Kim et al. (2011); Sakhaeifar et al. (2017, 2015); El-Badawy et al. (2018), the stiffness characteristic of an asphalt mixture presented by a dynamic modulus can be estimated by its component properties. In this study, the input variables vector ($\mathbf{x}$) defines the asphalt mixture's component properties. A summary of the selected input variables and their ranges in the dataset is presented in Table 4.2 with the x_i's and y being the input and output variables, respectively.

Table 4.2: Selected input variables (x) and output variable (y)

Variable	**Identity**	**Min.**	**Max.**	**Ave.**	**Std. Dev.**
y	$Log\|E^*\|$	2.62	4.37	3.76	0.46
x_1	Cum. % retained on $3/4''$	3.60	13.00	6.11	2.63
x_2	Cum. % retained on $3/8''$	12.68	26.29	19.01	4.11
x_3	Cum. % retained on #4	36.20	51.76	45.86	5.319
x_4	Cum. % retained on #8	52.87	65.70	59.42	5.06
x_5	Cum. % retained on #30	74.06	83.30	79.63	2.76
x_6	Cum. % retained on #50	86.22	90.12	88.57	1.15
x_7	Cum. % retained on #100	92.81	94.59	93.83	0.48
x_8	% Passing from #200	3.07	6.18	3.81	0.89
x_9	$Log\|G^*\|$	-2.29	3.03	0.50	1.26
x_{10}	Phase angle (degree)	28.15	79.17	52.86	11.54
x_{11}	$\%V_{beff}$	3.50	4.60	4.10	0.29
x_{12}	%VMA	12.50	14.40	13.51	0.49
x_{13}	%VFA	68.10	72.30	70.40	1.08
x_{14}	$\%V_a$	3.98	4.01	3.99	0.01

Cross-correlation analysis is performed on the 14 selected predictor variables and the obtained pairwise correlation matrix is given in Table 4.3 along with the schematic heat map of the correlation matrix given by Figure 4.1. Correlation coefficients with absolute values of 0.5 or higher are displayed in bold red text. The corresponding cells in the correlation heat map are shown in dark blue and dark red as shown in Figure 4.1. According to Table 4.3, the absolute values of the 130 correlation coefficients are greater than 0.1, with 50 of them greater than 0.5, indicating that several of the input variables give an impression of being highly correlated. The correlation heat map also clearly indicates that a high level of correlation (dark blue and dark red cells) exists within the input

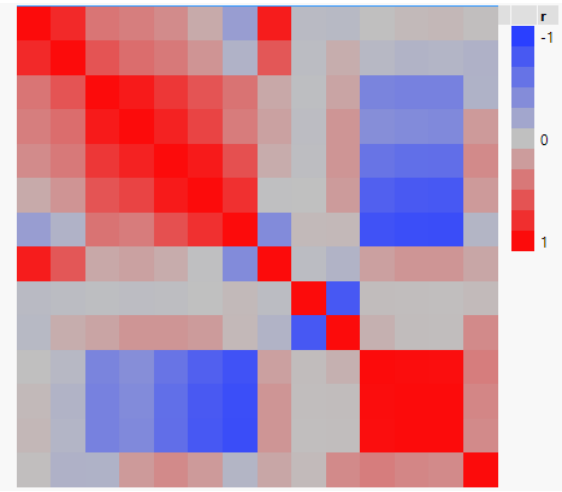

Figure 4.1: Heat map

variables. If the correlated input variables are detected, to enable accurate mapping of the inputs to the response variable, it would be useful to produce a smaller set of orthogonal pseudo-variables using the PCA method and use them in model development Fodor (2002).

4.3.2 Orthogonal Transformation Using PCA

In multivariate statistics, PCA is an orthogonal transformation of a set of (possibly) correlated variables into a set of linearly uncorrelated ones, and the uncorrelated (pseudo-) variables, called principal components (PCs), are linear combinations of the original input variables. This orthogonal transformation is performed such that the first principal component has the greatest possible variance (variation within the dataset). This procedure is then followed for the second component, then the third component, etc. This means that each succeeding component in turn has the highest variance when it is orthogonal to the preceding components Rollins et al. (2006); Jolliffe (2002); Kuźniar and Waszczyszyn (2006); Hua et al. (2007). To help visualize the PCA transformation, a schematic dataset with three input variables is presented in Figure 4.2 (left). As shown in this figure, in conducting PCA the data points are transferred from the original 3D original input space (on the left) to a 2D principal component space (on the right).

Table 4.3: Pairwise Correlation Matrix for the Selected Input Variables.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
x_1	**1**	**0.832**	0.412	0.366	0.294	0.119	−0.269	**0.905**	−0.044	−0.058	0.003	0.04	0.049	0.013
x_2	**0.832**	**1**	**0.597**	0.458	0.391	0.246	−0.109	**0.583**	−0.035	0.106	−0.061	−0.099	−0.089	−0.115
x_3	0.412	**0.597**	**1**	**0.918**	**0.756**	**0.596**	0.425	0.133	−0.019	0.154	−0.465	−0.485	−0.49	−0.111
x_4	0.366	0.458	**0.918**	**1**	**0.87**	**0.687**	0.375	0.169	−0.028	0.237	−0.388	−0.412	−0.424	0.212
x_5	0.294	0.391	**0.756**	**0.87**	**1**	**0.919**	**0.618**	0.112	−0.021	0.235	**−0.585**	**−0.631**	**−0.633**	0.3
x_6	0.119	0.246	**0.596**	**0.687**	**0.919**	**1**	**0.794**	−0.009	0.003	0.203	**−0.741**	**−0.796**	**−0.806**	0.209
x_7	−0.269	−0.109	0.425	0.375	**0.618**	**0.794**	**1**	−0.414	0.036	0.047	**−0.854**	**−0.886**	**−0.892**	−0.087
x_8	**0.905**	**0.583**	0.133	0.169	0.112	−0.009	−0.414	**1**	−0.032	−0.102	0.179	0.238	0.238	0.142
x_9	−0.044	−0.035	−0.019	−0.028	−0.021	−0.003	0.036	−0.032	**1**	**−0.808**	0.021	0.016	0.013	0.034
x_{10}	−0.058	0.106	0.154	0.237	0.235	0.203	0.047	−0.102	**−0.808**	**1**	0.09	0.024	0.014	0.3
x_{11}	0.003	−0.061	−0.465	−0.388	**−0.585**	**−0.741**	**−0.854**	0.179	0.021	0.09	**1**	**0.988**	**0.985**	0.372
x_{12}	0.04	−0.099	−0.485	−0.412	**−0.631**	**−0.796**	**−0.886**	0.238	0.016	0.024	**0.988**	**1**	**0.998**	0.321
x_{13}	0.049	−0.089	−0.49	−0.424	**−0.633**	**−0.806**	**−0.892**	0.238	0.013	0.014	**0.985**	**0.998**	**1**	0.301
x_{14}	0.013	−0.115	−0.111	0.212	0.3	0.209	−0.087	0.142	0.034	0.3	0.372	0.321	0.301	**1**

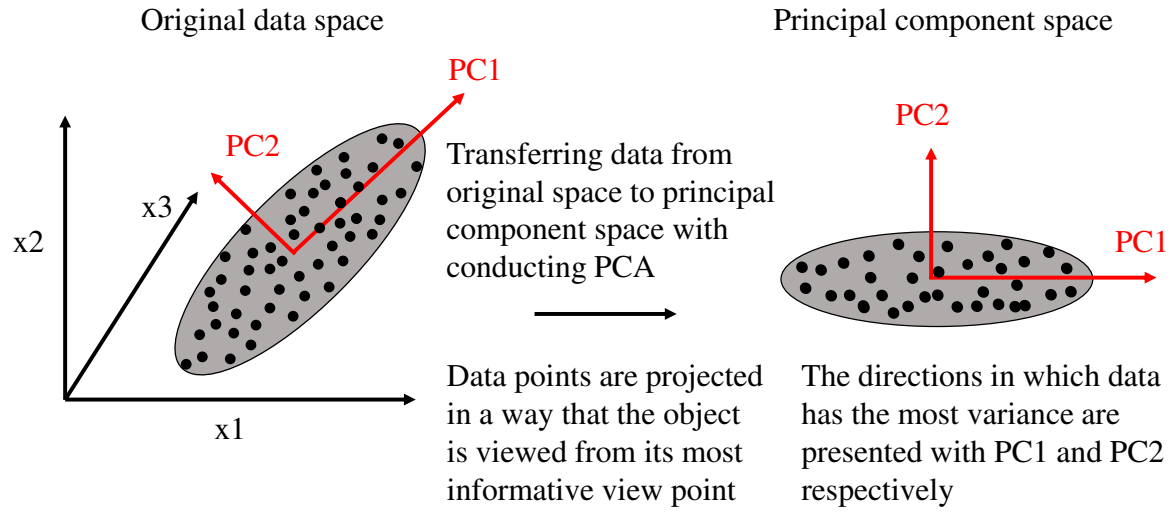

Figure 4.2: Schematic of the PCA transformation.

PCA can be performed either by eigenvalue decomposition of a data covariance (or correlation) matrix or by singular value decomposition. The process usually begins with mean centering the data matrix (and normalizing or using Z-scores) for each attribute as follows:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} \frac{x_{11}-\overline{x}_1}{s_1} & \frac{x_{12}-\overline{x}_2}{s_1} & \cdots & \frac{x_{1p}-\overline{x}_p}{s_p} \\ \frac{x_{21}-\overline{x}_1}{s_1} & \frac{x_{22}-\overline{x}_2}{s_2} & \cdots & \frac{x_{2p}-\overline{x}_p}{s_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x_{n1}-\overline{x}_1}{s_1} & \frac{x_{n2}-\overline{x}_2}{s_2} & \cdots & \frac{x_{np}-\overline{x}_p}{s_p} \end{bmatrix} \tag{4.1}$$

where for $k = 1$ to n and $j = 1$ to p, x_{kj} is the kth measurement for the jth variable, $\overline{x}_k$ is the sample mean for the kth variable, and s_k is sample standard deviation for the kth variable. As discussed in the previous section, highly correlated input variables lead to inflation of the standard error of estimate, negatively affecting the accuracy of the estimation. PCA will help us not only reduce the dimensionality of the modeling problem, but will also produce orthogonal pseudo-variables to be used in solving the problem. To perform PCA in this study we used eigenvalue decomposition of the correlation matrix of the data. The eigenvalues of the data correlation matrix are calculated, ranked, and sorted in descending order (representing their quota of the total variation within the dataset), as presented in Table 4.4. According to the eigenvalues, the first five PCs represent 95.8% of the existing variation within the dataset.

Recalling the fact that the PCs are linear combinations of the original input variables, the PCs can be defined as in Eq. (4.2):

$$pc_i = \sum_{j=1}^{14} \alpha_{ij} x_j + \beta_i \tag{4.2}$$

where $i = 1$ to 14 , the α_{ij} is the corresponding coefficients, the β_i are constants, and the x_j are the original input variables. Equation (4.2) can be stated in matrix notation as in Eq. (4.3):

$$\mathbf{p} = \mathbf{M}\mathbf{z} + \mathbf{n} \tag{4.3}$$

where

$$\mathbf{p} = \begin{bmatrix} pc_1 \\ pc_2 \\ pc_3 \\ pc_4 \\ pc_5 \end{bmatrix}$$

$$\mathbf{M}^T = \begin{bmatrix} 0.03 & 0.19 & -0.08 & -0.06 & -0.09 \\ 0.03 & 0.11 & -0.04 & -0.05 & 0.04 \\ 0.06 & 0.04 & 0.00 & 0.00 & 0.10 \\ 0.06 & 0.05 & 0.02 & 0.04 & 0.07 \\ 0.13 & 0.06 & 0.05 & 0.09 & -0.01 \\ 0.33 & 0.01 & 0.09 & 0.16 & -0.20 \\ 0.71 & -0.54 & 0.05 & 0.13 & -0.20 \\ -0.03 & 0.52 & -0.20 & -0.07 & -0.53 \\ -0.01 & -0.05 & -0.4 & 0.43 & 0.12 \\ 0.00 & 0.01 & 0.06 & -0.02 & 0.00 \\ -1.26 & 0.64 & 0.52 & 0.61 & 0.67 \\ -0.75 & 0.37 & 0.21 & 0.28 & 0.31 \\ -0.34 & 0.17 & 0.09 & 0.12 & 0.14 \\ -4.41 & 18.38 & 47.08 & 76.24 & -40.48 \end{bmatrix} \tag{4.4}$$

$$\mathbf{n} = \begin{bmatrix} -55.95 \\ -58.54 \\ -218.20 \\ -352.79 \\ 174.78 \end{bmatrix}$$

Further modeling efforts will be performed using the first five PCs.

Table 4.4: Eigenvalues of the normalized input variables matrix.

Number	Eigenvalue	Percent Variation	Cumulative Percent Variation
1	6.0225	43.018	43.018
2	3.2193	22.995	66.013
3	1.9746	14.104	80.118
4	1.4174	10.124	90.242
5	0.7850	5.607	95.848
6	0.3176	2.269	98.117
7	0.1091	0.779	98.896
8	0.0778	0.556	99.452
9	0.0549	0.392	99.844
10	0.0218	0.156	100

4.3.3 Holdout Cross Validation

In prediction problems, cross validation will be used to estimate model accuracy. Cross validation is a model validation technique that can be used to prevent overfitting as well as to assess how the results of a statistical analysis can be generalized to an independent dataset. In this study, a holdout cross validation technique is used in which the given dataset is randomly assigned to two subsets, d_0 and d_1, the training set and the test set, respectively. Since the training set contains 80% of the data points and the test set contains 20% of the data points, 80% of the data points are used to train the model and the remainder are used to evaluate the trained model's performance.

4.3.4 Principal Component Regression (PCR)

Linear regression attempts to model the relationship between response variables and explanatory variables by fitting a linear equation to observed data. In regression analysis, the least-squares method is used to calculate the best fitting line for the observed data by minimizing the sum of the squares of the residuals (differences between the measured responses and the fitted values by a linear function of parameters).

All possible regression structures were considered for representing the relationship between the response variable, $log|E^*|$, and predictor variables (pc_1, pc_2, pc_3, pc_4, and pc_5). To estimate the values of the unknown coefficients in the model, the least-squares criterion of minimizing the sum of

squared residuals (SSE) is used. Finally, after eliminating redundant terms, the reduced third order cubic and interaction terms were developed and selected as the best-fitted model. The developed model is called "Principal Component Regression (PCR)".

4.3.5 Principal Component Neural Network (PCNN)

A predictive model called "Principal Component Neural Network (PCNN)" is developed as briefly described in this section. ANNs are brain-inspired systems intended to replicate the way humans learn. Neural network structures consist of several layers, including input layers, output layers, and hidden layer(s), with nodes (neurons) in each layer Kartam (1994); Sanabria et al. (2017); Gong et al. (2018); Fathi et al. (2019). A three-layer feed-forward neural network is developed for this study. It consists of an input layer of five neurons (five input variables), a hidden layer of 10 neurons, and an output layer of one neuron (one response variable). A trial-and-error procedure of optimizing the computational time and cost function is used to choose the number of hidden neurons. In this study supervised learning is used in which a training dataset, including inputs and outputs, is presented to the network. The network adjusted its weights in such a way that the adjusted set of weights produces an input/output mapping resulting in the smallest error. This iterative process is carried on until the sum of square residuals (SSE) increases. After the learning or training phase, the performance of the trained network must be measured against an independent (unseen) testing data Cheng and Titterington (1994a); Kartam (1994). Let the input of each processing node be pc_i, the adjustable connection weight be w_{ij}, and let the bias at output layer be b_0, so that the network transfer (activation) function is $f(.)$. The jth output of the first layer can be obtained using Eq. (4.5)

$$\nu_j = f_1(pc_i, w_{ij}), \quad i = 1, ..., 5 \quad and \quad j = 1, ..., 10 \tag{4.5}$$

and the response will be

$$\hat{y} = f_2\left(f_1(pc_i, w_{ij})\right). \tag{4.6}$$

If we assume that

$$f_2(\nu_j, w_{Hj}) = b_0 + \sum_j \nu_j w_{Hj}, \tag{4.7}$$

and for each j,

$$f_1(pc_i, w_{ij}) = b_{Hj} + \sum_j pc_i w_{ij}, \tag{4.8}$$

then a feed-forward neural network can be formulated as follows:

$$\hat{y} = f_2 \left\{ b_0 + \sum_{j=1}^{n} \left[w_{Hj} \cdot f_1 \left(b_{Hj} + \sum_{i=1}^{m} pc_i w_{ij} \right) \right] \right\}, \tag{4.9}$$

where pc_i is pseudo input parameter i, w_{ij} is the weight of connection between input variable i (for $i = 1$ to 5) and neuron j of the hidden layer, b_0 is a bias at the output layer, w_{Hj} is the weight of connection between neuron j of the hidden layer and output layer neuron, b_{Hj} is a bias at neuron j of the hidden layer (for $j = 1$ to 10), and $f_1(t)$ and $f_2(t)$ are transfer functions of the hidden layer and output layer, respectively.

It should be pointed out that iteration proceeds until the convergence criterion is met. Thus, similar to the linear regression model, the validation set is not used. The Bayesian Regularization algorithm is used to achieve network training efficiency.

4.3.6 Effective Variable Space

It is widely known that the use of an empirical predictive model outside the convex hull containing the data points is prohibited. In this context, *effective variable space* is referred to the space where the uncertainty of the developed models is bound to their already calculated thresholds. In other words, outside of this region, the extrapolated behavior of the models may not be predictable. To guard against extrapolation, Ghasemi et al. Ghasemi et al. (2018a) concluded that the space containing input data could be interpreted as a symmetrical convex space, then demonstrated how this space can be used in the design procedure.

Following the approach in Ghasemi et al. (2018a), a normal distribution is assumed for each input variable (x_i), resulting in their joint distribution being bi-variate normal, and such distributions are usually represented in form of a contour diagram. Since a contour curve on such a diagram

contains the points on a surface with the same distance from the $x_i x_j$ plane, these points have a constant density function Kutner et al. (2004) (see Figure 4.3 for an example of such distribution). The cross section is obtained by slicing a bi-variate normal surface at a constant distance from the $x_i x_j$ plane. As indicated in Figure 4.3, the n-dimensional hyperspace is a hyper-ellipsoid with minimum volume (to avoid any gaps in the edges).

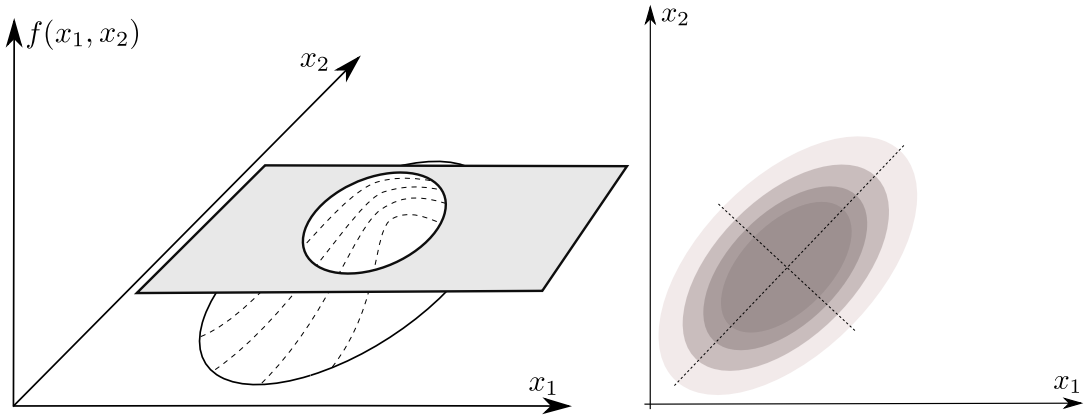

Figure 4.3: A schematic of a bivariate normal distribution.

Khachiyan's work Todd and Yıldırım (2007) formulates the problem of finding an approximate minimum volume enclosing ellipsoid ($\mathcal{E}$) given p data points in n-dimensions as an optimization problem. In Ghasemi et al. Ghasemi et al. (2018a), the authors detailed the derivation of a procedure for solving this problem and obtaining its effective variable space. For brevity, the flowchart in Figure 4.4 summarizes this iterative method for finding the minimum volume enclosing ellipsoid. This algorithm was used to find two enclosing ellipsoids in the primary space (14-dimensional) and the pseudo space (5-dimensional) of the dataset. It should be pointed out that this space is independent of the predictive models and is used only to solve the optimal (and inverse) design problems.

4.3.7 Guideline for Implementation

A summary of the methodologies used to develop the framework is presented in Figure 4.5. The procedure begins with collecting experimental data from the laboratory, followed by the pre-

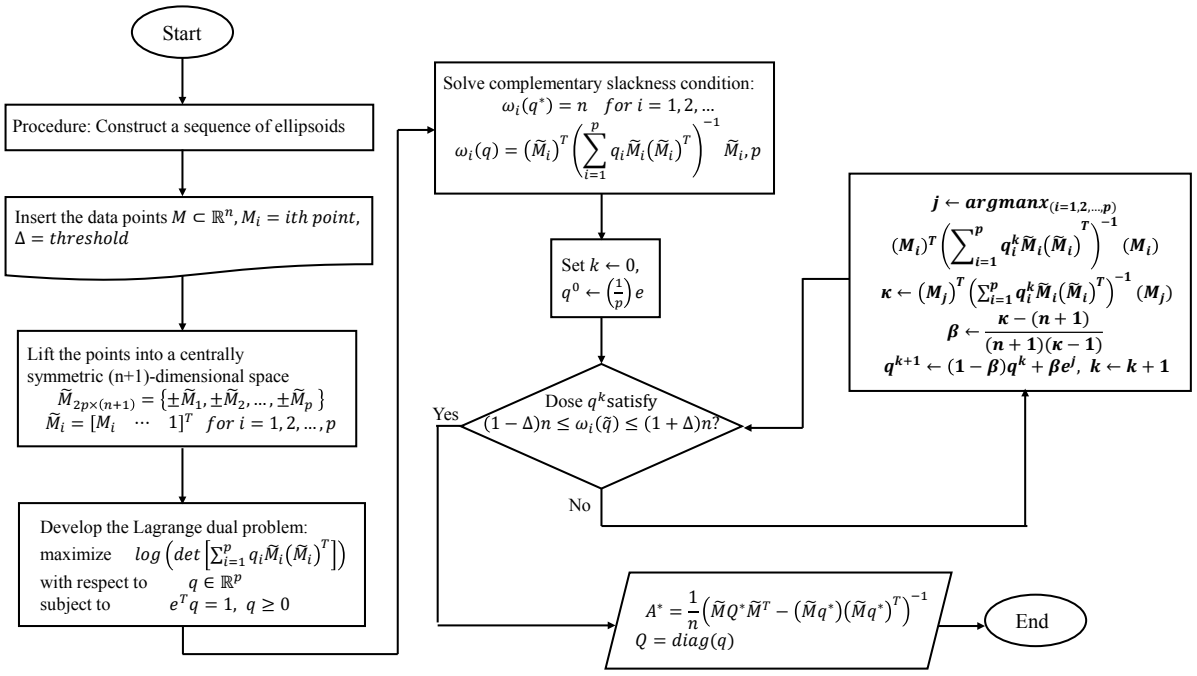

Figure 4.4: Iterative method to solve the problem

processing step of input variable evaluation. The flowchart continues with the model development and the addition of a constraint on the n-dimensional input variable hyperspace to the modeling problem. The developed models can then be used to predict pavement performance, solve design-based optimization problems, etc. There are a number of aspects of the proposed framework that can be achieved using free and commercially available software like *MATLAB®, Python, and R packages*, and one may implement many parts of the framework in the language of their interest. For example, the algorithm to find the n-dimensional hyper-ellipsoid is very straightforward using the flowchart in Figure 4.4.

4.4 Developed Model Results, Performance, and Validation

The results produced by the developed models are presented in this section, and their capability to use empirical data to estimate the dynamic modulus of asphalt mixtures is evaluated.

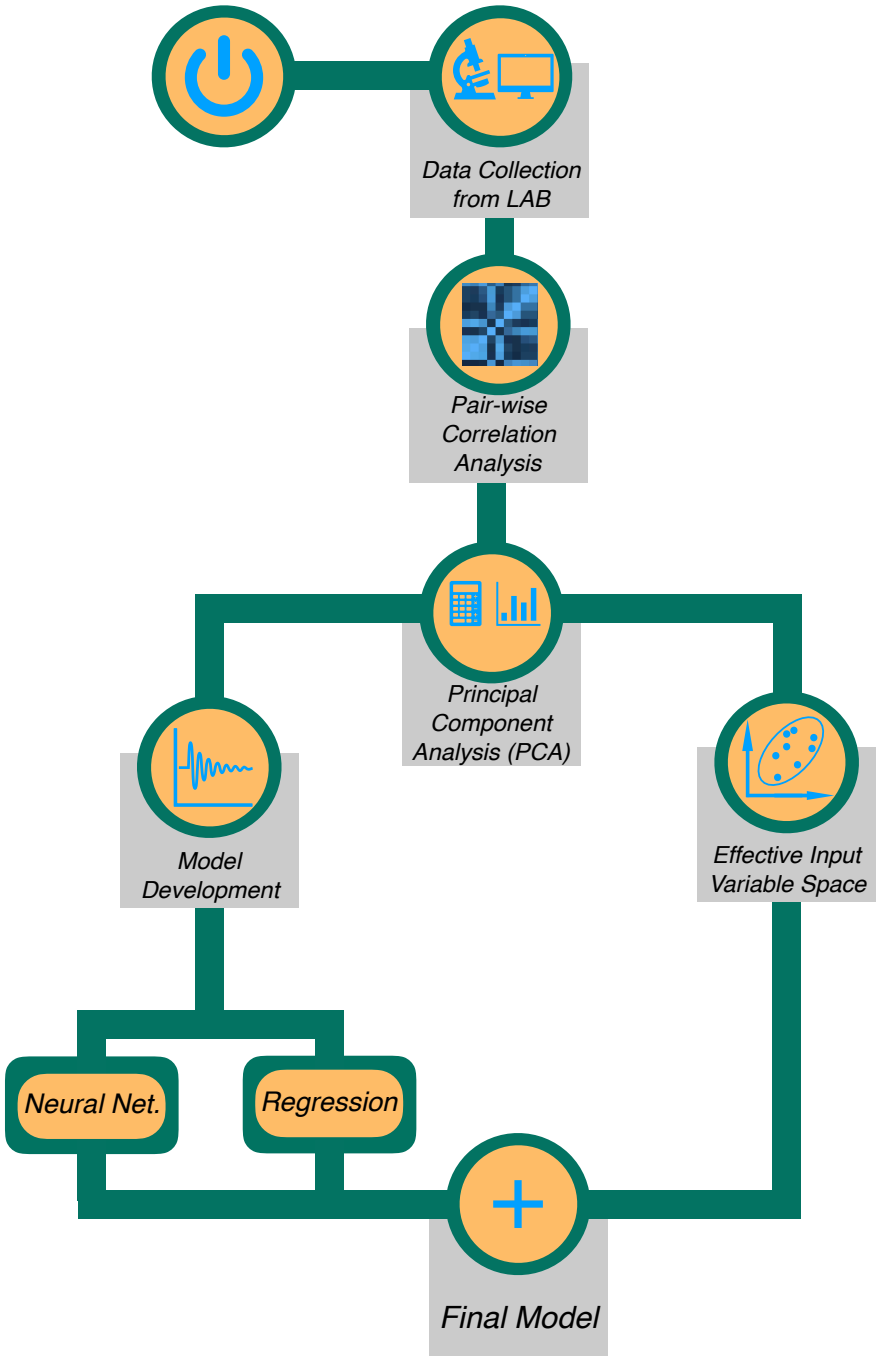

Figure 4.5: A summary of the methodologies used.

4.4.1 Model Performance

The performance is first compared with the existing predictive models; modified Witczak, Hirsch, and Alkhateeb models selected from the literature are presented in Equations (4.10)–(4.13), respectively Bari and Witczak (2007); Christensen Jr et al. (2003); Al-Khateeb et al. (2006)

$$\begin{aligned}\log|E^*| = -0.349 + 0.754(|G_b^*|^{-0.0052})(6.65 - 0.032\rho_{200} - 0.0027(\rho_{200})^2 + \\ 0.011\rho_4 - 0.0001(\rho_4)^2 + 0.006\rho_{3/8} - 0.00014(\rho_{3/8})^2 - 0.08V_a - 1.06(\frac{V_{beff}}{V_{beff}+V_a})) \\ + \frac{2.558 + 0.032V_a + 0.713(\frac{V_{beff}}{V_{beff}+V_a}) + 0.0124\rho_{3/8} - 0.0001(\rho_{3/8})^2 - 0.0098\rho_{3/4}}{1 + exp(-0.7814 - 0.5785\log|G_b^*| + 0.8834\log\delta_b)}\end{aligned} \tag{4.10}$$

where $|E^*|$ is dynamic modulus in psi; $|G^*|$ is the binder shear modulus in psi; δ_b is the binder phase angle in degrees; $\rho_{3/4}$ is the cumulative percent aggregate retained on the $3/4''$ sieve (19 mm); $\rho_{3/8}$ is the cumulative percent aggregate retained on the $3/8''$ sieve (9.5 mm); ρ_4 is the cumulative percent aggregate retained on the No. 4 sieve (4.75 mm); ρ_{200} is the percent aggregate passing the No. 200 sieve (0.075 mm); V_a is the percent air void in the mix; V_{beff} is the effective asphalt content; VMA is the percent of voids in the mineral aggregate, and VFA is the percent voids filled with asphalt,

$$|E_m^*| = P_c\left(4,200,000(1 - \frac{VMA}{100}) + 3|G^*|_b(\frac{VFA \times VMA}{10,000})\right) + \frac{(1-P_c)}{\frac{\frac{1-VMA}{100}}{4,200,000} + \frac{VMA}{3|G_b^*|(VFA)}}} \tag{4.11}$$

where

$$P_c = \frac{(20 + 3|G_b^*|(VFA)/(VMA))^{0.58}}{650 + (3|G^*|_b(VFA)/(VMA))^{0.58}} \tag{4.12}$$

and $|E^*|_m$ is dynamic modulus of HMA in psi; P_c is the aggregate contact volume; VMA is the percentage of mineral aggregate voids in compacted mixture; and VFA is the percentage of voids filled with asphalt in the compacted mixture,

$$|E_m^*| = 3(\frac{100 - VMA}{100})(\frac{(90 + 1.45\frac{|G^*|_b}{VMA})^{0.66}}{1100 + (0.13\frac{|G_b^*|}{VMA})^{0.66}})|G_g^*| \tag{4.13}$$

where $|E_m^*|$, $|G_b^*|$, and $|G_g^*|$ (the complex shear modulus of binder in the glassy state, assumed to be 10^9 Pa.) are in Pa. Equation (4.14) shows the best reduced third-order (linear) regression model (PCR) fitting the measured response:

$$\begin{aligned}\hat{y} =& c_0 + c_1 pc_1 + c_2 pc_2 + c_3 pc_3 + c_4 pc_4 + c_5 pc_5 \\ &+ c_6 pc_1 pc_2 + c_7 pc_1 pc_3 + c_8 pc_1 pc_4 + c_9 pc_1 pc_5 \\ &+ c_{10} pc_2 pc_3 + c_{11} pc_2 pc_4 + c_{12} pc_2 pc_5 + c_{13} pc_3 pc_4 \\ &+ c_{14} pc_3 pc_5 + c_{15} pc_4 pc_5 + c_{16} pc_1 pc_2 pc_3 + c_{17} pc_1 pc_2 pc_4 \\ &+ c_{18} pc_1 pc_2 pc_5 + c_{19} pc_1 pc_3 pc_4 + c_{20} pc_1 pc_3 pc_5 + c_{21} pc_2 pc_3 pc_4 \\ &+ c_{22} pc_1 pc_4 pc_5 + c_{23} pc_2 pc_4 pc_5 + c_{24} pc_3 pc_4 pc_5\end{aligned} \tag{4.14}$$

where, $c_0 = 6.59$; $c_1 = 2.58$; $c_2 = 4.4$; $c_3 = -0.36$; $c_4 = 0.49$; $c_5 = 1.93$; $c_6 = -0.33$; $c_7 = -0.77$; $c_8 = -1.69$; $c_9 = 0.15$; $c_{10} = -1.65$; $c_{11} = -4.68$; $c_{12} = 4.81$; $c_{13} = 0.7$; $c_{14} = -0.85$; $c_{15} = -1.58$; $c_{16} = -0.17$; $c_{17} = -0.79$; $c_{18} = 1.83$; $c_{19} = 0.04$; $c_{20} = 0.18$; $c_{21} = 0.42$; $c_{22} = 0.05$; $c_{23} = 0.32$; $c_{24} = 0.06$. The trained three-layer ANN (PCNN) presented in Equation (4.9) contains the following connection weights and biases:

$$\mathbf{W}^T = \begin{bmatrix} -0.511 & 0.134 & 0.654 & -1.064 & -0.267 \\ -0.315 & -0.147 & -0.267 & 0.177 & -1.047 \\ -0.060 & -1.266 & 0.759 & -1.248 & -0.331 \\ -0.075 & 0.022 & 0.208 & 0.015 & 0.167 \\ -0.074 & 0.022 & 0.206 & 0.015 & 0.165 \\ 0.103 & -0.177 & 1.253 & -1.045 & 0.535 \\ 0.078 & -0.020 & -0.231 & -0.014 & -0.172 \\ 0.238 & 0.070 & -0.885 & 0.848 & 0.943 \\ 0.123 & 0.456 & -0.387 & 1.547 & -0.017 \\ -0.079 & 0.020 & 0.213 & 0.014 & 0.173 \end{bmatrix}$$

$$\mathbf{W}_H = \begin{bmatrix} 0.869 \\ -0.886 \\ 0.632 \\ -0.291 \\ -0.288 \\ -0.859 \\ 0.299 \\ 0.556 \\ 0.971 \\ -0.299 \end{bmatrix}, \ \mathbf{B}_{\mathrm{H}} = \begin{bmatrix} 0.162 \\ 0.710 \\ 0.319 \\ -0.008 \\ -0.009 \\ -0.570 \\ 0.007 \\ 0.290 \\ -0.373 \\ -0.007 \end{bmatrix}, \ \mathbf{B}_0 = [0.148]$$

Figure 4.6 presents the performance of the developed models in terms of measured values of dynamic modulus versus the fitted dynamic modulus values. The measured and fitted values are fairly close to the line of equality, indicating that the fitted values are highly correlated with the measured ones.

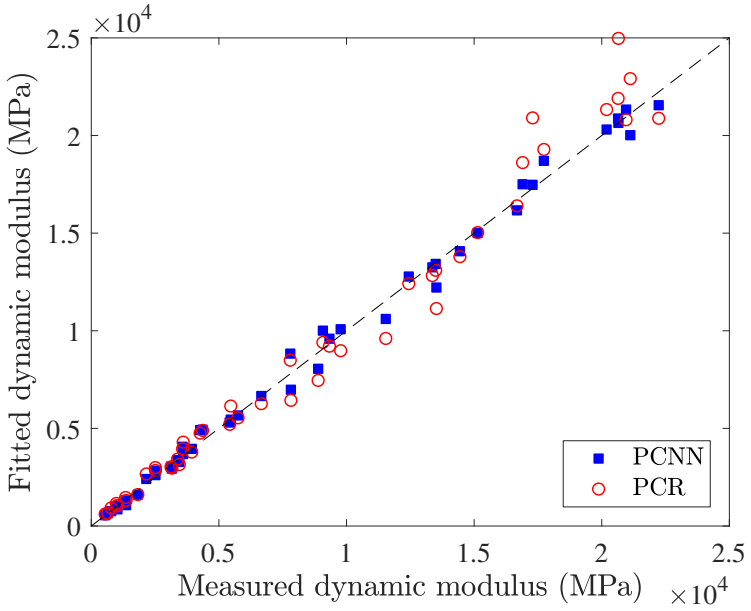

Figure 4.6: Measured values of dynamic modulus .

Table 4.5: Statistics which are used to compare model performance.

Statistical Component	**Formula**	**Definition**
Average difference (AD)	$AD = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$	An estimate of systematic model bias
Average absolute difference (AAD)	$AAD = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$	Average closeness of the fitted and measured values of response
r_{fit}	$r_{fit} = \frac{n\sum_{i=1}^{n} y_i\hat{y}_i - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n}\hat{y}_i)}{\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}\sqrt{n\sum_{i=1}^{n}\hat{y}_i^2 - (\sum_{i=1}^{n}\hat{y}_i)^2}}$	Correlation of the measured and fitted values of response
Coefficient of determination (R^2)	$R^2 = 1 - \frac{SS_{res}}{SS_{total}}$	Portion of the response variation elucidated by regressors in the fitted model in linear models

Comparisons of PCR and PCNN performance to that of the existing predictive models are conducted based on three statistics: average difference (AD), average absolute difference (AAD), and correlation between measured and fitted values of response (r_{fit}). A summary of the definitions of these statistical components and their formulas is presented in Table 4.5. In the formulas presented in Table 4.5, y_i is the ith measured response, $\hat{y}_i$ is the ith fitted response, and n is the number of data points.

The results of the comparison are presented in Table 4.6. According to the values of r_{fit} in Table 4.6, the estimated dynamic modulus values obtained form PCR and PCNN models are highly correlated with measured values according to the values, showing that the both PCR and PCNN performed well in terms of modeling the response variable.

Although the corresponding values of r_{fit} for modified Witczak, Hirsch, and Alkhateeb models are 0.93, 0.95, and 0.95, respectively, the average difference and average absolute difference with respect to the measured response are significantly higher than those of PCR and PCNN. This means that the fitted values by the modified Witczak, Hirsch, and Alkhateeb models are not close as those fitted by PCR and PCNN to the response value. In other words, r_{fit}, which reflects the correlation between response and estimated response (if one goes up the other one goes up), could be biased, and in this situations other statistics (AD, and AAD) could be used to evaluate the goodness of fit. The dynamic modulus measured and predicted values are presented in Figure 4.7 for four asphalt mixtures. According to the presented master curves the current study (PCNN model) provides the

Table 4.6: Performance comparison of the developed and existing models.

		Average Difference (MPa)	Average Absolute Difference (MPa)	r_{fit}	R^2
PCR	Training	3.9	575.3	0.996	0.99
	Testing	−162.3	718.9	0.995	na
PCNN	Training	13.2	380.7	0.997	na
	Testing	9.7	337.5	0.997	na
Modified Witczak		−2460	3152.1	0.93	0.88
Hirsch		1241.6	1785.7	0.95	0.91
Alkhateeb		2844.5	2984.5	0.95	0.90

closest values of E^* to the measurements for all of three test temperatures, while, the conventional models either overestimate or underestimate the response variable.

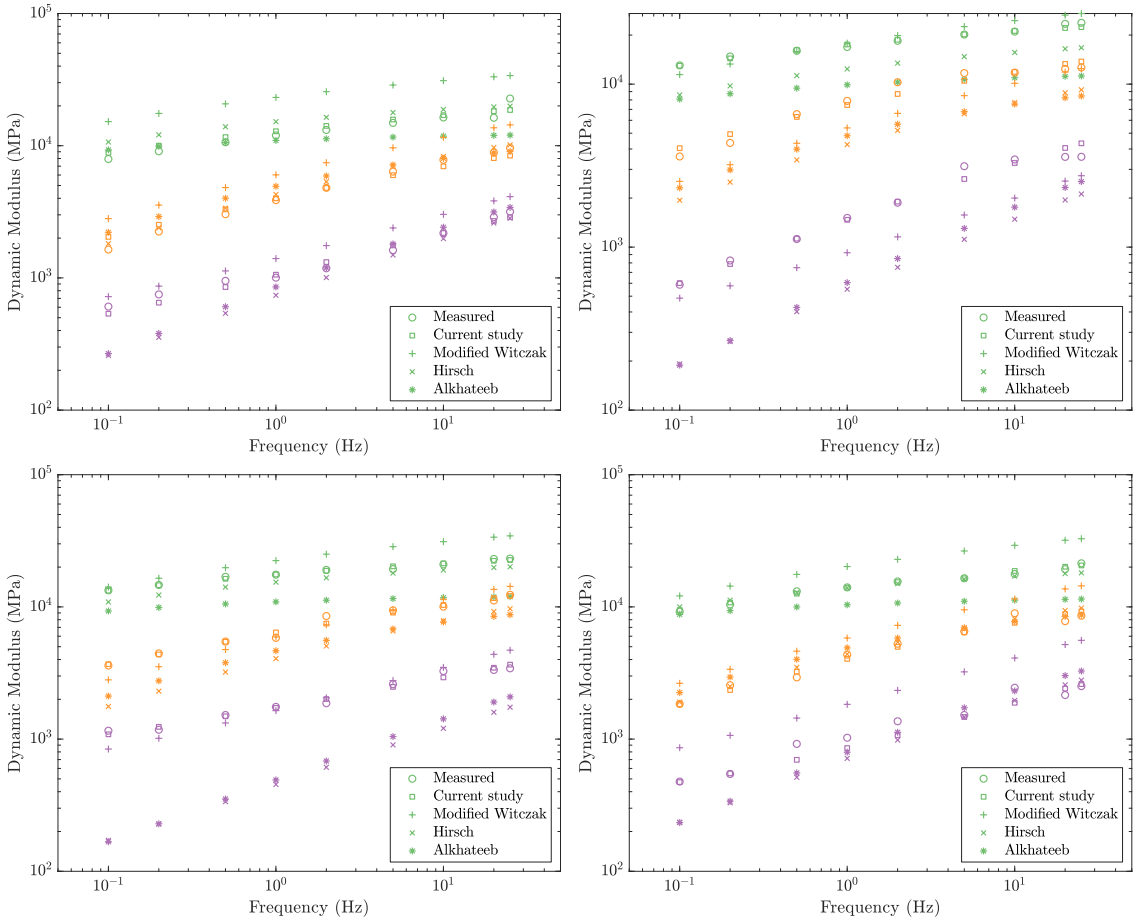

Figure 4.7: Comparing the measured and predicted dynamic modulus.

A graphical comparison of the PCR and PCNN performance and that of the existing models is presented in the following section.

4.4.2 Receiver Operating Characteristic Analysis (ROC)

A receiver operating characteristic (ROC) graph is a technique for visualizing, organizing, and selecting classifiers based on their performance. ROC graphs are widely used in medical decision-making as well as in machine learning and data-mining research Fawcett (2006). True ROC curves plot the false positive rate (probability of false alarm) on the x-axis and the true positive rate (probability of detection) on the y-axis. A classifier is said to perform well if the ROC curve climbs rapidly towards the upper left-hand corner. The more the curve deviates from $y = x$ behavior, the more accurate the prediction is Bi and Bennett (2003). We can borrow from the concept of ROC curve to obtain a measurement of fit for the competing models and a ROC graph for this study is presented in Figure 4.8 for this study. As described in Equation (4.15), the x-axis indicates the standardized residuals ordered from the lowest to the highest (e_i^{**}). Residuals are sorted in ascending order and divided by the largest one that belongs to the Hirsch model.

$$e_i^{**} = \frac{|e_i^*|}{|e_{i_{MAX}}^*|} \tag{4.15}$$

The y-axis indicates the fraction of points whose standardized residuals are less than e_i^{**} Marti-Vargas et al. (2013). Although the curves obtained for all of the models are monotonically non-decreasing and climb towards the upper left-hand corner (the desired situation shows that the predictive models perform well), PCNN and PCR curves were the highest, proving their better performance to be better than that of existing predictive models.

A convenient global measure of the goodness-of-the-fit is the Area Under the Curve (AUC). To compare classifiers, it is more desirable to reduce ROC performance, to a single scalar value representing expected performance. Since the AUC is a portion of the area of the unit square, its value will always lie between 0 and 1. The AUC values for the PCNN, PCR, Alkhateeb, modified Witczak, and Hirsch models are 0.9864, 0.9717, 0.8746, 0.7609, and 0.6320, respectively. One

can use the ROC and AUC analysis results and rank the predictive models according to their performances. In this study, the PCNN model reflected the highest performance in predicting the dynamic modulus value, while the Hirsch model ranked the lowest among all the models.

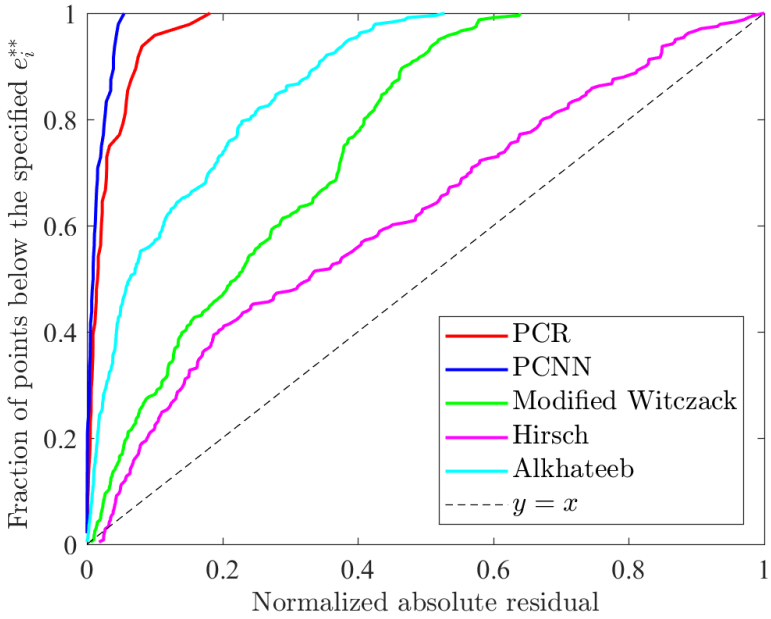

Figure 4.8: ROC curves for the developed and the existing predictive models.

4.4.3 Model Validation

The current regression model is presented in the following general form as

$$y_i = f_i(\mathbf{Z_i}, \boldsymbol{\theta}) + e_i^*. \tag{4.16}$$

In the above equation f_i is the ith expectation function, $\boldsymbol{\theta}$ is the vector of parameters, and e_i^* is a random deviation of y_i from f_i. This term is assumed to be independent and normally distributed with a mean of zero and unknown variance σ^2 for $i = 1, \cdots, n$, where n is the number of input vectors. If the above assumptions are violated, the results of the analysis could be misleading or erroneous. These assumptions can be testified by examining residuals as defined by

$$e_i^* = y_i - \hat{y}_i. \tag{4.17}$$

The assumption of independency holds when the residuals plot does not reflect a trivial pattern. The normality assumption is assessed by creating a normal probability plot of the residuals. When the error has a normal distribution, this plot will appear as a straight line Devore (2011). These assumptions were checked for PCR and PCNN, as presented in Figure 4.9. The assumption of equal variances does not appear to be violated because there are no trivial pattern in this plot. Figure 4.9 presents the normal probability of the residuals in which it can be seen that the data points are close to the straight line and the normality assumption is validated.

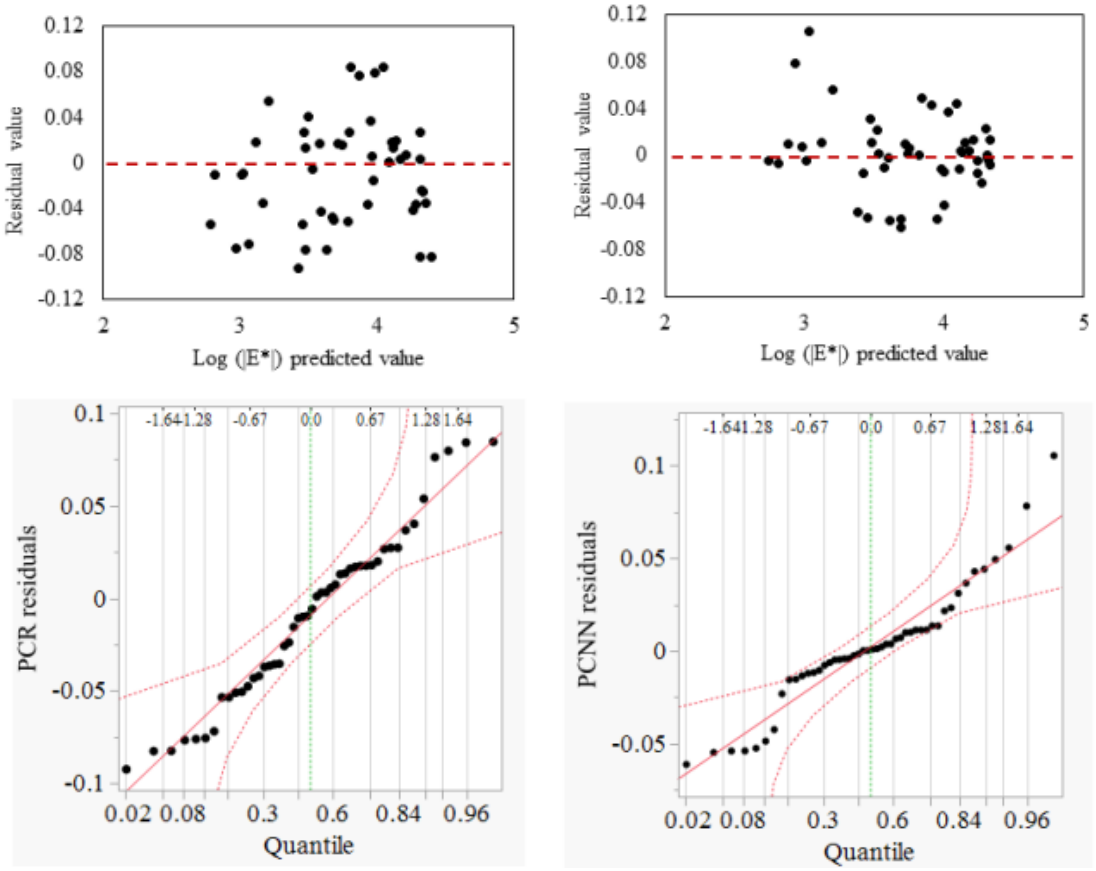

Figure 4.9: Checking the assumptions of independency models.

4.5 Application of the Framework: Flexible Pavement Design and Optimization

The above framework is used along with an optimization algorithm to answer the following two central questions:

- what design parameters result in the maximum $|E^*|$?
- what design parameters result in a pre-specified $|E_0^*|$?

One can see that the first item corresponds to the optimal design problem while the second one corresponds to the so-called inverse design.

Since it was shown through multiple statistical measurements that PCNN had the best prediction capability, this model is used in the following section to solve the optimization problems. The ANN used in PCNN is essentially an interconnected nonlinear function, and this necessitates the application of a global optimizer. Moreover, the effective variable space enters the problem as a series of constraints and further restricts the available algorithms. The optimal design problem is formulated as follows:

$$\begin{aligned} \text{maximize} \quad & |E^*| = F_{ANN}(\mathbf{x}) \\ \text{with respect to} \quad & \mathbf{x} = (x_1, ..., x_{14}) \\ \text{subject to} \quad & (\mathbf{x} - \mathbf{v})^T A(\mathbf{x} - \mathbf{v}) \leq \mathbf{1}, \\ & (\mathbf{x}_{pca} - \mathbf{v}')^T A'(\mathbf{x}_{pca} - \mathbf{v}') \leq \mathbf{1}, \end{aligned} \tag{4.18}$$

where the vector of fourteen variables is $\mathbf{x}$, and $(\mathbf{x} - \mathbf{v})^T A(\mathbf{x} - \mathbf{v}) \leq \mathbf{1}$ are the enclosing ellipsoid constraint equations for the original and PCA-based variables. A penalty function approach is used to convert the above constrained problem to an unconstrained one Rahami et al. (2011a). In this case, when the penalty function is active, it decreases (increases) the objective function when the problem is one of maximization (minimization), and the degree of penalty is based on the closeness of the solution to the corresponding constraint.

Since the inverse design problem aims at finding the specification of a predefined goal, it is defined as a minimization problem as follows:

$$\begin{aligned} \text{minimize} \quad & error = ||E^*| - |E_0^*|| \\ \text{with respect to} \quad & \mathbf{x} = (x_1, ..., x_{14}) \\ \text{subject to} \quad & (\mathbf{x} - \mathbf{v})^T A(\mathbf{x} - \mathbf{v}) \leq \mathbf{1}, \\ & (\mathbf{x}_{pca} - \mathbf{v}')^T A'(\mathbf{x}_{pca} - \mathbf{v}') \leq \mathbf{1}, \end{aligned} \tag{4.19}$$

where $|E_0^*|$ is the desired (goal) dynamic modulus. Although a similar penalization method can also be used to address the constraints in this case, for the above problem the constraints will penalize the objective function when they are active.

Reliable solution of the above problems requires the application of a gradient-free optimization algorithm. Gradient-based optimization algorithms are not applicable in this case because of the network-based nature of the ANNs. Evolutionary-based algorithms are potentially easy-to-use algorithms in the above problems. Novel algorithms have been used to solve complex optimization problems in recent years He and Yao (2002); Cai et al. (2006), and in this case, Mean-Variance Mapping Optimization (MVMO), an in-house optimization algorithm based on the work by Elrich et al. Rueda and Erlich (2015); Erlich et al. (2010), is used. The constraints are handled using the approach described in Aslani et al. Aslani et al. (2018), in which, the convergence rate of a constrained MVMO was compared to the already-developed methodologies using benchmark structural problems. Authors in Ghasemi et al. (2018a) indicated that a constrained MVMO is capable accurately identifying an optimal value with a minimum number of simulations. It should be noted that the choice of optimization algorithm is not the principal focus of this study.

Figure 4.10 (left) depicts the convergence achieved for the first design problem by the constrained MVMO algorithm. The initial data points are random making it heavily penalized, and then the objective function increases as the algorithms evolves. Exploration-exploitation behavior is achieved using adaptive strategies in the course of optimization for MVMO. $\delta = 0.05$ is used as the threshold

in Figure 4.4. Solving the maximization problem resulted in $|E^*_{max}|$ = 53,703 MPa. The optimal design parameters are presented in the first column of Table 4.7.

To find the maximum amount of dynamic modulus one could design for without low temperature failure in the asphalt binder, the maximization problem was solved one more time with an additional constraint of $G^* sin\delta \leq 5000$, resulting in $|E^*_{max}|$ = 36,307 MPa. Corresponding design parameters are presented in the second column of Table 4.7 as the optimal design 2.

Figure 4.10 (right) shows the convergence of the algorithm for the inverse design problem after starting randomly from three different initial points, with the algorithm is terminated when the error reaches about 10^{-9}. A pre-specified $|E^*_0|$ of 20,417 MPa is considered and the inverse problem of finding the corresponding design parameters is solved. Because of non-linearity of the function, the problem has no unique solution. Three of the possible solutions are presented as designs 1 to 3 in Table 4.7.

Finally, the five sets of design parameters are compared with current design specification, with the results shown in Table 4.7. The percentage of aggregate passing by each sieve size is within the acceptable range of the gradation specification. Gradation charts are presented in Figure 4.11. The obtained percentages of air voids are 4%, which is the target value in the design specification. The obtained values for VMA are slightly less than 14% for a nominal maximum aggregate size (NMAS) of 12.5 mm because the VMA values of the nine mixtures used to train the PCNN are slightly less that 14% (see Table 4.1). The acceptable range for VFA varies with the amount of traffic load measured in million Equivalent Single Axle Loads (ESALs) as follows:

- traffic loading < 0.3 $\rightarrow 70 <$ VFA < 80
- $0.3 <$ traffic loading $< 3.0 \rightarrow 65 <$ VFA < 78
- traffic loading > 3.0 $\rightarrow 65 <$ VFA < 75

The VFAs obtained for all of the five sets of design are satisfied for all of the traffic categories.

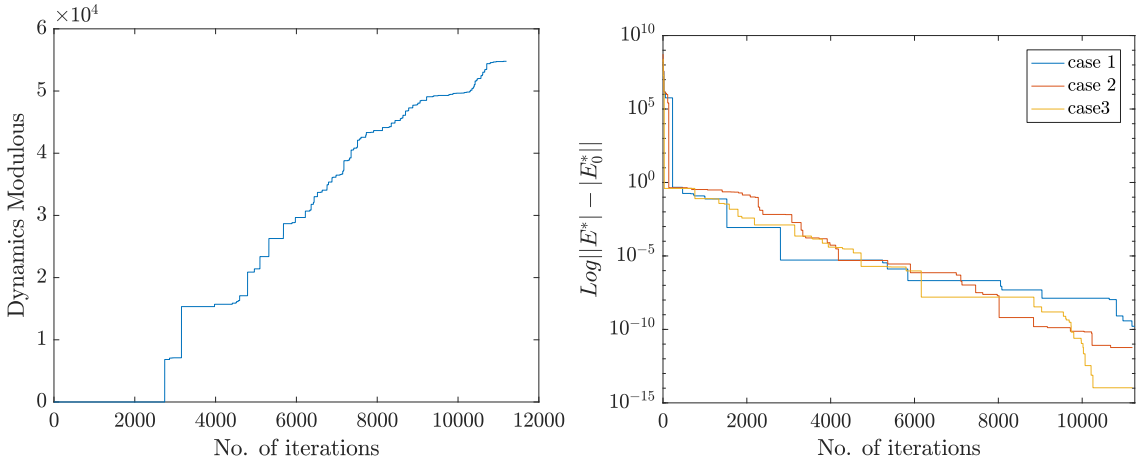

Figure 4.10: Convergence for the optimal design problem.

Table 4.7: Th corresponding design parameters.

Identity	**Optimal Design 1**	**Optimal Design 2**	**Design 1**	**Design 2**	**Design 3**	**Design Specification**			
						Control Points		**Restricted Zone**	
						Lower	**Upper**	**Lower**	**Upper**
%Passing from 3/4″	100	100	100	100	100	-	100	-	-
%Passing from 1/2″	93.38	94.03	92.25	91.88	91.80	90	100	-	-
%Passing from 3/8″	81.74	81.72	79.57	79.92	80.70	-	90	-	-
%Passing from #4	53.00	53.90	55.36	55.23	54.39	-	-	-	-
%Passing from #8	39.56	40.51	41.37	41.08	40.92	28	58	39.1	39.1
%Passing from #30	20.75	20.68	21.02	20.87	20.83	-	-	19.1	23.1
%Passing from #50	11.66	11.60	12.08	11.81	12.02	-	-	15.5	15.5
%Passing from #100	6.22	6.21	6.52	6.38	6.40	-	-	-	-
%Passing from #200	4.10	3.85	4.38	4.58	4.56	2	10	-	-
G* (Mpa)	103.13	7.81	133.51	30.20	11.82	-	-	-	-
Phase angle (degree)	35.71	39.60	47.69	47.27	44.77	2	8	-	-
Vbeff%	4.11	4.18	4.02	4.06	4.05	-	-	-	-
VMA	13.47	13.56	13.41	13.45	13.44	-	-	-	-
VFA	70.29	70.50	70.11	70.24	70.24	-	-	-	-
Va%	4.00	4.00	3.99	4.00	4.01	4		-	-

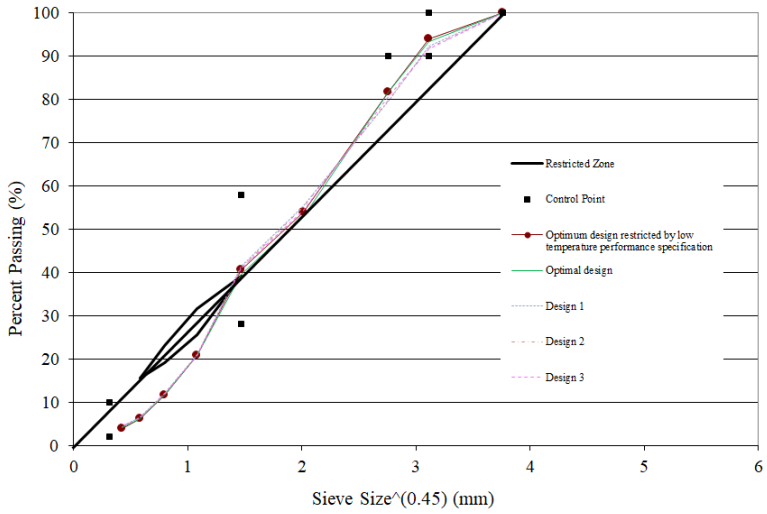

Figure 4.11: Aggregate gradation graphs.

4.6 Conclusions

This study used the HMA dynamic modulus data and focused to evaluate the quality of predictor variables to be used in a procedure of model development. Correlation analysis is performed to identify cross-correlated input variables, and correlated inputs are replaced by orthogonal pseudo-inputs (PCs) obtained using PCA. Two separate models are developed using multivariate regression and ANN (called PCR and PCNN, respectively). Extrapolation in empirical modeling is addressed by adding the constraint of an n-dimensional enclosing ellipsoid to the modeling problem. Performances of the proposed models were compared to existing predictive models using both statistical analysis and ROC analysis. The models developed satisfactorily estimated the dynamic modulus value, with PCNN indicating remarkably better performance when fitted to the test data than the existing predictive models from the literature. These PCA-based approaches are thus highly recommended as precise modeling strategies in this application. Moreover, these methodologies appear to be capable of modeling other material properties and future investigation in this regard is recommended. To determine this framework's application in pavement design, two optimization problems including optimal design and inverse design have been presented and solved using

a mean-variance mapping optimization algorithm. The results for the two problems are in a good agreement with the HMA mix design specification and thus could be a reasonable starting point in solving real-life design problems. Although, the developed models as well as obtained optimal design parameters are based on the empirical database created in this study, the suggested framework has the capability of being re-trained and adjusted to fit new data. For obtaining more reliable and applicable results, a larger empirical database would be required.

4.7 References

AASHTO, A. (2008). Mechanistic-empirical pavement design guide: A manual of practice. *Aaoshat Off. Ed.*

Al-Khateeb, G., Shenoy, A., Gibson, N., and Harman, T. (2006). A new simplistic model for dynamic modulus predictions of asphalt paving mixtures. *J. Assoc. Asph. Paving Technol.*, 75.

Andrei, D., Witczak, M., and Mirza, M. (1999). Development of a revised predictive model for the dynamic (complex) modulus of asphalt mixtures. *Dev. 2002 Guide Des. New Rehabil. Pavement Struct. NCHRP*.

Arabali, P., Sakhaeifar, M. S., Freeman, T. J., Wilson, B. T., and Borowiec, J. D. (2017). Decision-making guideline for preservation of flexible pavements in general aviation airport management. *J. Transp. Eng. Part B: Pavements*, 143(2):04017006.

Aslani, M., Ghasemi, P., and Gandomi, A. H. (2018). Constrained mean-variance mapping optimization for truss optimization problems. *Struct. Des. Tall Spec. Build.*, 27(6):e1449.

Bari, J. and Witczak, M. (2007). New predictive models for viscosity and complex shear modulus of asphalt binders: For use with mechanistic-empirical pavement design guide. *Transp. Res. Rec. : J. Transp. Res. Board*, 1(2001):9–19.

Bi, J. and Bennett, K. P. (2003). Regression error characteristic curves. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 43–50.

Birgisson, B., Roque, R., Kim, J., and Pham, L. V. (2004). The use of complex modulus to characterize the performance of asphalt mixtures and pavements in florida. Technical report.

Bozorgzad, A. et al. (2017). Consistent distribution of air voids and asphalt and random orientation of aggregates by flipping specimens during gyratory compaction process. *Constr. Build. Mater.*, 132:376–382.

Cai, Z., Cai, Z., and Wang, Y. (2006). A Multiobjective Optimization-Based Evolutionary Algorithm for Constrained Optimization. *Evol. Comput. IEEE Trans.*, 10(6):658–675.

Cheng, B. and Titterington, D. (1994). Neural networks: A review from a statistical perspective.

Christensen Jr, D., Pellinen, T., and Bonaquist, R. (2003). Hirsch model for estimating the modulus of asphalt concrete. *J. Assoc. Asph. Paving Technol.*, 72.

Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences*. Cengage learning.

El-Badawy, S., Abd El-Hakim, R., and Awed, A. (2018). Comparing artificial neural networks with regression models for hot-mix asphalt dynamic modulus prediction. *J. Mater. Civ. Eng.*, 30(7):04018128.

Erlich, I., Venayagamoorthy, G. K., and Worawat, N. (2010). A mean-variance optimization algorithm. *IEEE Congr. Evol. Comput.*, pages 1–6.

Fathi, A., Mazari, M., Saghafi, M., Hosseini, A., and Kumar, S. (2019). Parametric study of pavement deterioration using machine learning algorithms. In *Airfield and Highway Pavements 2019: Innovation and Sustainability in Highway and Airfield Pavement Technology*, pages 31–41. American Society of Civil Engineers Reston, VA.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognit. Lett.*, 27(8):861–874.

Fodor, I. K. (2002). A Survey of Dimension Reduction Techniques. Technical report, U.S. Department of Energy, U.S. Department of Energy.

Ghasemi, P. (2018). Performance evaluation of coarse-graded field mixtures using dynamic modulus results gained from testing in indirect tension mode of testing.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. (2018a). Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: Elimination of correlated inputs and extrapolation in modeling. *Struct. Multidiscip. Optim.*, pages 1–19.

Ghasemi, P., Aslani, M., Rollins, D. K., Williams, R. C., and Schaefer, V. R. (2018b). Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks. *Int. J. Pavement Res. Technol.*

Gong, H., Sun, Y., Mei, Z., and Huang, B. (2018). Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks. *Constr. Build. Mater.*, 190:710–718.

Guide, N. D. (2004). Guide 1-37a, guide for mechanistic-empirical design of new and rehabilitated pavement structures, national cooperative highway research program. *Transp. Res. Board Natl. Res. Counc.*

He, J. and Yao, X. (2002). From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. *IEEE Trans. Evol. Comput.*, 6(5):495–511.

Hua, X., Ni, Y., Ko, J., and Wong, K. (2007). Modeling of temperature–frequency correlation using combined principal component analysis and support vector regression technique. *J. Comput. Civ. Eng.*, 21(2):122–135.

Jamrah, A., Kutay, M. E., and Ozturk, H. I. (2014). Characterization of asphalt materials common to michigan in support of the implementation of the mechanistic-empirical pavement design guide. Technical report.

Johnson, R. A., Wichern, D. W., et al. (2014). *Applied multivariate statistical analysis*, volume 4. Prentice-Hall New Jersey.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.*, 30(3):487.

Kartam, N. (1994). Neural Netwroks in Civil Engineering: Systems and Application. *J. Comput. Civ. Eng.*, 8(2):149–162.

Kim, Y. R., Underwood, B., Far, M. S., Jackson, N., and Puccinelli, J. (2011). Ltpp computed parameter: Dynamic modulus. Technical report.

Kutner, M. H., Nachtsheim, C., and Neter, J. (2004). *Applied linear regression models*. McGraw-Hill/Irwin.

Kuźniar, K. and Waszczyszyn, Z. (2006). Neural networks and principal component analysis for identification of building natural periods. *J. Comput. Civ. Eng.*, 20(6):431–436.

Marti-Vargas, J. R., Ferri, F. J., and Yepes, V. (2013). Prediction of the transfer length of pre-stressing strands with neural networks. *Comput. Concr.*, 12(2):187–209.

Nobakht, M. and Sakhaeifar, M. S. (2018). Dynamic modulus and phase angle prediction of laboratory aged asphalt mixtures. *Constr. Build. Mater.*, 190:740–751.

Peng, C., Feng, J., Feiting, S., Changjun, Z., and Decheng, F. (2019). Modified two-phase micromechanical model and generalized self-consistent model for predicting dynamic modulus of asphalt concrete. *Constr. Build. Mater.*, 201:33–41.

Rahami, H., Kaveh, A., Aslani, M., and Asl, R. N. (2011). A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization. *Int. J. Optim. Civ. Eng.*, 1(January):29–46.

Ren, R., Han, K., Zhao, P., Shi, J., Zhao, L., Gao, D., Zhang, Z., and Yang, Z. (2019). Identification of asphalt fingerprints based on atr-ftir spectroscopy and principal component-linear discriminant analysis. *Constr. Build. Mater.*, 198:662–668.

Rollins, D. K., Zhai, D., Joe, A. L., Guidarelli, J. W., Murarka, A., and Gonzalez, R. (2006). A novel data mining method to identify assay-specific signatures in functional genomic studies. *BMC bioinformatics*, 7:377.

Rueda, J. L. and Erlich, I. (2015). Testing MVMO on learning-based real-parameter single objective benchmark optimization problems. *2015 IEEE Congr. Evol. Comput. CEC 2015 - Proc.*, pages 1025–1032.

Sakhaeifar, M. S., Richard Kim, Y., and Garcia Montano, B. E. (2017). Individual temperature based models for nondestructive evaluation of complex moduli in asphalt concrete. *Constr. Build. Mater.*, 137:117–127.

Sakhaeifar, M. S., Richard Kim, Y., and Kabir, P. (2015). New predictive models for the dynamic modulus of hot mix asphalt. *Constr. Build. Mater.*, 76:221–231.

Sanabria, N., Valentin, V., Bogus, S., Zhang, G., and Kalhor, E. (2017). Comparing neural networks and ordered probit models for forecasting pavement condition in new mexico. Technical report.

Shu, X. and Huang, B. (2008). Dynamic modulus prediction of hma mixtures based on the viscoelastic micromechanical model. *J. Mater. Civ. Eng.*, 20(8):530–538.

Todd, M. J. and Yıldırım, E. A. (2007). On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids. *Discret. Appl. Math.*, 155(13):1731–1744.

CHAPTER 5. DEVELOPING A ROBUST MODELING APPROACH FOR PAVEMENT PERFORMANCE PREDICTION AND OPTIMIZATION

A paper published at the Journal of Association of Asphalt Pavement Technologists (AAPT)

5.1 Abstract

In pavement technology, performance models are mathematical expressions that relate pavement condition, surface distresses and structural properties as response variables to a set of predictors including material properties, traffic loading, environmental factors, etc. In the existence of numerous important predictors and their interrelationships, developing a predictive model for pavement performance is not a trivial task. In this study, a machine learning-based framework is developed for predicting pavement performance. The framework starts with a preparation step of data pre-processing and data wrangling. After removing outliers, the framework will conduct Principal Component Analysis (PCA) to firstly reduce the dimensionality of the problem and secondly eliminate the pairwise correlation between the inputs by producing orthogonal pseudo-inputs. These pseudo-inputs are used to develop two predictive models using multivariate regression analysis and Artificial Neural Networks (ANN). In empirical predictive models, mapping input space to response space can be threatened by extrapolation. However, it is often disregarded by design engineers. In this study to confront extrapolation, a method is implemented to determine a hyperspace based on the inputs. The hyperspace determines where the predictive model is valid up to given thresholds and is then added as a constraint to the modeling problem. Two of the performance-related characteristics of asphalt mixture including rut resistance and dynamic modulus are considered to examine the robustness of the proposed approach. The developed predictive models then compared to the conventional models for each case and indicated superior performance r_{fit} of 0.97 and 0.99 for rutting and dynamic modulus, respectively). A global variable importance analysis is also conducted to obtain the most effective variable in each case. Percent air void and binder shear

properties appeared to be the most effective variables in predicting rutting and dynamic modulus, respectively. To indicate an application of the developed framework in asphalt pavement design, for each of these two cases a design-related optimization problem is defined and solved using mean-variance mapping optimization (MVMO) algorithm. The obtained optimal design parameters are within the acceptable range of current asphalt pavement design specifications and thus can be used as an appropriate starting point in design procedure.

5.2 Introduction

Maintenance, rehabilitation, and reconstruction of the highway system are the major expenses in a state general expenditure. Therefore, seeking to develop an accurate and efficient performance model to predict the remaining service life of a pavement and to provide its rehabilitation or reconstruction requirements punctually is beneficial. Relating pavement condition, surface distresses, and structural properties, to a set of predictors including material properties, traffic loading, environmental factors, etc. via mathematical expressions is called performance modeling Jalali et al. (2019); Hosseini et al. (2020). According to the American Association of State Highway and Transportation Officials (AASHTO), pavement performance is the pavement ability to serve traffic over time sufficiently. To measure and predict pavement performance, a reproducible, authoritative, and field calibrated condition evaluating system is required. Several researchers tempted to develop pavement performance predictive models but almost all of the performance models are site specific and also restricted to the materials used in the AASHTO road test.

One of the performance related properties of asphalt pavement is its resistance to rutting. Rutting or permanent deformation often happens under the wheel- path and appears as a depression worn into a pavement with uplift occurring along the sides Ghasemi et al. (2018b); Notani et al. (2019). To analyze asphalt mixture rut susceptibility, performance testing along with mechanistic-empirical regression-based modeling appear to be a common approach Bashin et al. (2012). To simulate rutting in laboratory a rut resistance index called Flow Number (FN) is defined. In a repeated loading and unloading test FN is the point at which the strain rate starts to increase

with loading. This parameter has demonstrated a strong correlation with rutting that happens in asphalt pavement due to traffic in field. In asphalt pavement design procedure, the amount of rutting should generally be limited to 0.4 inches (10.16 mm) regarding the total deformation of a pavement structure.

It has been demonstrated that the amount of rutting is a function of binder viscosity, volumetric properties of asphalt mixture, and testing temperature Kaloush et al. (2003); Witzcak (2002). Kvasnak et al.Kvasnak et al. (2007), proposed a list of the efficacious factors in rut susceptibility of asphalt mixture. The list includes nominal maximum aggregate size (NMAS), voids in mineral aggregate (VMA), percentage aggregate passing through sieve sizes No.4, No.16, No.200, binder grade, binder viscosity, asphalt content, testing temperature, and the number of gyrations. Rodezno et al. Rodezno et al. (2010) selected 12 parameters, i.e., testing temperature, maximum shear stress, normal stress, binder viscosity, percentage aggregate passing through sieve sizes 3/4-inch, 3/8-inch, and No.4, percentage air voids, effective binder content, binder content, VMA, and voids filled with asphalt (VFA) to be important in estimating asphalt pavement rutting behavior. It is illustrated by Apeagyei, et al.Apeagyei (2011) that dynamic modulus test results at specific temperature and loading frequencies along with aggregate gradation appears to have strong correlation with FN test results. However, there are some discrepancies on the existence of correlation between rut susceptibility of asphalt mixture and its dynamic modulus value Birgisson et al. (2005); Pellinen and Witczak (2002); Timm et al. (2006).

Another widely used pavement performance characteristic is dynamic modulus, $|E^*|$, which defines stress-strain relationship of asphalt mixtures under sinusoidal loading. Dynamic modulus represents the stiffness characteristic of asphalt mixture and it has a significant role in pavement design. Therefore, several researchers have been attempted to predict asphalt mixture dynamic modulus as a function of material's components properties, loading rate, and temperature Nobakht and Sakhaeifar (2018); Peng et al. (2019); Shu and Huang (2008).

There are several well-known predictive models for dynamic modulus amongst which, some use regression analysis while newer ones use other techniques including genetic programming and

artificial neural networks Ziari et al. (2018). Wiczak et al. developed a predictive model using material components properties including binder viscosity. Andrei et al Andrei et al. (1999) modified the original Witczak model. The developed model has then been modified to use binder shear modulus instead of binder viscosity Bari and Witczak (2007). Christensen et al. Christensen Jr et al. (2003) created a predictive model based on the law of mixtures. Alkhateeb et al. Al-Khateeb et al. (2006) created a model from the law of mixtures to be used over a wide range of temperatures and loading frequencies. Sakhaeifar et al.Sakhaeifar et al. (2017) created separate temperature-based models that can predict dynamic modulus over a wide range of temperature. The predictor variables of the aforementioned models are selected from the following list: cumulative percentage aggregate retained on sieve sizes 3/4-inch, 3/8-inch, No.4, and percent aggregate passing the No.200, VMA, VFA, percentage air voids, effective binder content, binder shear modulus ($|G^*|$), and binder phase angle (δ).

The predictor variables used in the conventional performance predictive models, e.g., rutting and dynamic modulus, are not admitted being independent set of variables and therefore may not be suitable to be used in modeling. Pairwise correlated predictors in the data set can decrease the estimation accuracy of their effects on the response variable. Therefore, a data preparation step is useful to assure that the input variables are qualified to be used in model development K Rollins Sr (2015). The process of transforming a raw data into another format with the intention of making it more appropriate and valuable for the main analysis is called data wrangling. One of the data wrangling techniques is reducing the dimension of the data especially when predictor variables are highly correlated. There are several dimensionality reduction techniques including exploratory factor analysis (EFA), variable clustering and principal component analysis (PCA). To reduce the dimension of the data, variable clustering and EFA tend to eliminate some of the predictors Thompson (2004) while PCA transforms the data into a new coordinate system and preserves more information of observed variables with less tendency to information loss. Therefore, PCA is preferred over other techniques for this particular application presented in this study.

PCA is a multivariate statistical procedure that reduces the size of a data set by transforming a large set of variables into a smaller set of orthogonal, (i.e. zero correlation) pseudo-variables called principal components (PCs). The produced pseudo-inputs can be used as input variables in developing predictive models. They not only make the prediction analysis easier but also contain most of the information of the large set Ghasemi et al. (2018b).

Another issue within performance models is that since they are developed based on empirical data, they can be prone to extrapolation which is defined as the process of estimating beyond the original observation range. In case of extrapolation, predictive models are subject to major uncertainty and high risk of producing meaningless results. To prevent extrapolation from happening, a hyper-space containing all the data points can be found and added as a constraint to the desired modeling problem.

The focus of this study is developing a machine learning-based framework to predict pavement performance using orthogonal pseudo-inputs obtained from principal component analysis. Unlike most of the conventional performance models, the proposed framework utilizes different data for model training and performance testing. An n-dimensional hyperspace is determined and added as a constraint to the modeling problem to guard against extrapolation. The authority of the proposed framework is illustrated by solving two separate problems of predicting rutting behavior and predicting dynamic modulus of asphalt mixtures. In order to find the most effective variable, global variable importance analysis is performed on the developed models. This research also claims to determine the optimal design and a successful approach to perform inverse design of asphalt mixture as some of the applications of the framework using a state-of-the-art evolutionary optimization algorithm.

5.3 Methodology

5.3.1 2.1. Data preprocessing

Data preprocessing is a crucial task in every machine learning and data mining project. Irrelevant and redundant information as well as correlated and unreliable predictor variables can

produce misleading results. Therefore, the representation and quality of data should be verified prior running the main analysis.

A parsimonious model is a model that accomplishes a great explanatory predictive power with as few predictor variables as possible. The obstacles in creating such model can be the availability of numerous, highly correlated, and weakly related or unrelated predictor variables Fodor (2002). In a general model, the expectation function is given by

$$\eta_i = f(x_i, \theta) \tag{5.1}$$

where vector x_i is input values at the i^{th} sampling time, η_i is the expected value of the response, and $\theta = [\theta_1, \ldots, \theta_q]^T$ is a vector containing unknown parameters of the model. The corresponding element of the i^{th} row and j^{th} column in the Jacobian matrix, **J,** is defined by $\frac{\partial \eta_i}{\partial \theta_j}$. In this case if two columns (e.g., m and n) are orthogonal, their correlation coefficient should be zero. In other words, if the two columns are independent, the information used to estimate θ_m is separate from the information used to estimate θ_n. Such approach will solidify the inputs-output relationship which leads to enhancing the accuracy of analysis. Correlated columns in the Jacobian Matrix are a consequence of pairwise correlated inputs. Therefore, to decrease standard parameter errors and increase accuracy of input-output mapping, the pairwise correlation of inputs could be eliminated by implementing PCA Fodor (2002).

5.3.2 Data Wrangling: Dimensionality Reduction using Principal Component Analysis (PCA)

Data wrangling is a main step in creating a machine learning model. During this step the data will be converted into a suitable format which can be used by any machine learning algorithm. During data wrangling step PCA can be implemented to remove correlated predictors, improve algorithm performance, and reduce overfitting.

PCA is a statistical technique often used to reduce the dimensionality of the data by selecting the most important features that capture maximum information about the data set. In other words, PCA is a technique of extracting important features (in the form of PCs) from a large set of

available features in a data set. The features are selected based on variance that they cause in the response. During this orthogonal transformation, original inputs of the data set are converted to the principal components (PCs) which are linear combinations of the original inputs Rollins et al. (2006). PCA works in a way that the variable that causes highest variance is the first PC, the variable responsible for the second highest variance is considered the second PC, and so on.

PCA can be implemented by either eigenvalue decomposition of a data covariance (or correlation) matrix or singular value decomposition (SVD), often after normalizing the data matrix (mean centering) for each procedure Jolliffe (2002). In the present study, the eigenvalue decomposition of the data correlation matrix is used. PCA reduces the dimension of a given data set, $\mathbf{X}$, by representing the l original variables ($x1$, . . . , xl) as p new pseudo-variables (PCs), where $p < l$. For a given dataset, $\mathbf{X}$, this analysis is summarized in the following steps:

- Standardize original data, X, by transforming it to Z using the following equations:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x_{11}} & \mathbf{x_{12}} & \cdots & \mathbf{x_{1l}} \\ \mathbf{x_{21}} & \mathbf{x_{22}} & \cdots & \mathbf{x_{2l}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x_{n1}} & \mathbf{x_{n2}} & \cdots & \mathbf{x_{nl}} \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} \mathbf{z_1} & \mathbf{z_2} & \ldots & \mathbf{z_l} \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{x_{11}}-\overline{\mathbf{x_1}}}{\mathbf{s_1}} & \frac{\mathbf{x_{12}}-\overline{\mathbf{x_2}}}{\mathbf{s_2}} & \ldots & \frac{\mathbf{x_{1l}}-\overline{\mathbf{x_l}}}{\mathbf{s_l}} \\ \frac{\mathbf{x_{21}}-\overline{\mathbf{x_1}}}{\mathbf{s_1}} & \frac{\mathbf{x_{22}}-\overline{\mathbf{x_2}}}{\mathbf{s_2}} & \ldots & \frac{\mathbf{x_{2l}}-\overline{\mathbf{x_l}}}{\mathbf{s_l}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathbf{x_{n1}}-\overline{\mathbf{x_1}}}{\mathbf{s_1}} & \frac{\mathbf{x_{n2}}-\overline{\mathbf{x_2}}}{\mathbf{s_2}} & \ldots & \frac{\mathbf{x_{nl}}-\overline{\mathbf{x_l}}}{\mathbf{s_l}} \end{bmatrix}$$

[2]

where, for $k = 1$ to n and $j = 1$ to l, x_{kj} is the k^{th} measurement for the j^{th} variable, x_k is sample mean for the k^{th} variable, and s_k is sample standard deviation for the k^{th} variable.

- Determine the unit eigenvectors e_1 , . . . , e_l of $\mathbf{Z}$.
- Determine the corresponding eigenvaluesλ_1, . . . , λ_l.
- Rank the eigenvectors (in descending order) according to their eigenvalues.
- Select the p PCs according to their eigenvalues.

Considering the number of eigenvalues (or components) versus their fraction of total represented variance is used to select the appropriate number of PCs. One might stop adding more PCs when

little variance is gained by retaining additional eigenvalues. The selected PCs will be used in the main analysis.

5.3.3 Cross Validation

To assure that the results of a statistical analysis can generalize to an independent dataset a model validation technique called cross validation should be implemented. This technique is mainly used to prevent overfitting in prediction problems, where a model is usually trained with a dataset called training data and is tested against a first-seen dataset called testing data Refaeilzadeh et al. (2009). According to the size of an available database and the desired computational time, the most suitable cross validation technique should be selected to be used in the modeling.

5.3.4 Principal Component Regression (PCR)

Recalling from section 5.3.2, the selected PCs are used as new predictors in the modeling procedure. All possible regression structures should be considered and examined for mapping the predictors(*pc*'s) to the response variable (pavement performance). To estimate the values of unknown coefficients of the model the least squares criterion of minimizing the sum of squared residuals (SSE) is implemented. Finally, after eliminating the redundant terms the reduced model is developed and selected as the best fitted model. The developed model is called "Principal Component Regression (PCR)".

5.3.5 Principal Component Neural Network (PCNN)

A predictive model called, "Principal Component Neural Network (PCNN)" is developed using artificial neural networks (ANN). ANNs consist of a collection of connected units or nodes called artificial neurons which learn to perform tasks by considering examples Cheng and Titterington (1994b). A three-layer feed-forward neural network, consisting of an input layer of n neurons, with n being the number of principal components (pc's), a hidden layer of 10 neurons, and an output layer of one neuron, which is the pavement performance, is developed using the MATLAB

software. The number of hidden neurons is selected to make balance between the cost function and computational time using a trial and error procedure.

To initiate the training process the input of each processing neuron, pc_i, is multiplied by a randomly assigned (and adaptable) connection weight w_{ij} and the weighted inputs are summed and added to a threshold value, b_0. The result crossed through a nonlinear transfer function (sigmoid in this study) resulted the output of the first layer, ν_i, which establishes the input for the next layer. During each iteration in training process, the network adjusts its weights and biases to minimize the loss function which is the difference between the predicted and observed values of response variable. The iterative procedure continues till the convergence criterion is satisfied. The performance of the trained network is then validated against an unseen set of data (test data set). For network training efficiency the Bayesian Regularization algorithm is implemented.

The output νj from the jth hidden node is given by

$$\nu_j = f_1(pc_i, w_{ij}), \quad i = 1, ..., m \quad and \quad j = 1, ..., 10 \tag{5.2}$$

and the single output $\widehat{y}$ is:

$$\hat{y} = f_2\left(f_1(pc_i, w_{ij})\right). \tag{5.3}$$

Then the expression of $\widehat{y}$ as a function of PC becomes a complicated nonlinear regression function with the j sets of weights, as parameters. It is assumed that

$$f_2(\nu_j, w_{Hj}) = b_0 + \sum_j \nu_j w_{Hj}, \tag{5.4}$$

and for each j,

$$f_1(pc_i, w_{ij}) = b_{Hj} + \sum_j pc_i w_{ij}. \tag{5.5}$$

so a general form of the feed forward neural network is described in Eq. 5.6

$$\hat{y} = f_2 \left\{ b_0 + \sum_{j=1}^{n} \left[w_{Hj} \cdot f_1 \left(b_{Hj} + \sum_{i=1}^{m} pc_i w_{ij} \right) \right] \right\} \tag{5.6}$$

where, b_0 is bias at output layer; wHj is weight of connection between neuron j of the hidden layer and output layer neuron; bH j is bias at neuron j of the hidden layer (for $j = 1$ to 10); $w$ij is weight of connection between input variable i (for $i = 1\ to\ m$) and neuron j of the hidden layer; pc$i$ is pseudo input parameter i; $f_1(t)$ is transfer function of the hidden layer, and $f_2(t)$ is transfer function of the output layer.

For an arbitrary variable t The transfer functions used in the network, $f_1(t)$ and $f_2(t)$, are defined in Eq. 5.7,

$$f_k(t) = \frac{1}{1 + e^{-t}} \quad for \quad k = 1, 2. \tag{5.7}$$

5.3.6 Effective Variable Space

A major shortcoming in most of research articles where a predictive model is developed is the space (range of input variables) where the developed empirical model is valid. The answer to this question is important because the behavior of such empirical model is arbitrary when the inputs are not inside the n-dimensional hyper ellipsoid covering the original data (in this case the training dataset). A number of valid assumptions make the above conclusion possible. One of these assumptions is based on the normal distribution of the input variables and their joint distribution which is bi-variate normal Devore (2011). In this way, any slice of such a domain (at a constant density function) will result in a n-dimensional hyper ellipsoid. Thus, one should perform a test to determine if the desired data is inside this appropriate space. This concept is better visualized in Figure 5.1Kutner et al. (2004).

The n-dimensional space can be located using a number of approaches Todd and Yıldırım (2007). The problem of finding the n-dimensional ellipsoid that contains m-dimensional data has been a subject to a thorough classical computational complexity analysis Sun and Freund (2004) and is not the purpose of this article.

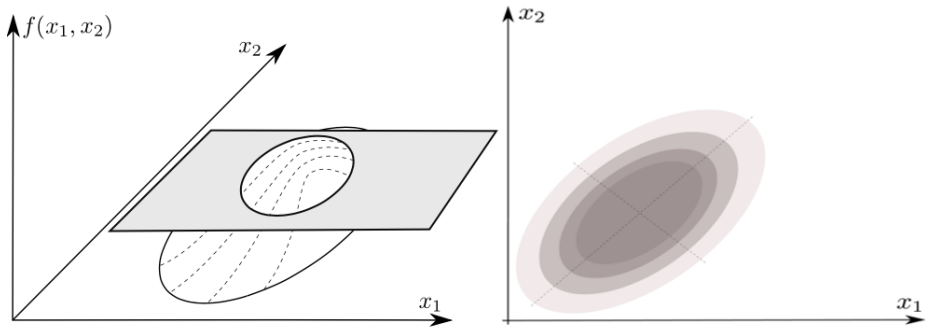

Figure 5.1: Schematics of a bi-variate distribution.

Ghasemi et al.Ghasemi et al. (2018a) showed that an iterative scheme where a modified dual optimization problem is solved numerically to find the details of the desired n-dimensional hyperspace (in this case an n-dimensional ellipsoid) is an effective solution to this problem. This approach is only dependent on an inverse matrix step which might become numerically cumbersome for large data sets.

Another approach to this problem is possible through the usage of interior-point algorithms as described in Sun and Freund (2004). To summarize the approach used in this article, we start by defining the following optimization (minimization) problem:

$$\begin{aligned} \text{minimize} \quad & Vol(\mathcal{E}) \\ \text{with respect to} \quad & \mathbf{v} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{n \times n} \\ \text{subject to} \quad & (\mathbf{M}_i - \mathbf{v})^T \mathbf{A} (\mathbf{M}_i - \mathbf{v}) \leq \mathbf{1} \quad for \quad i = 1, 2, ..., p \end{aligned} \tag{5.8}$$

where, p is the number of points given in n dimensions and M_i is the i$^{\text{th}}$ point. The unknowns are vector v and the matrix A. The above problem is reformulated in order to fit an interior point type algorithm as follows:

$$\begin{aligned} \text{minimize} \quad & -\ln(\det(M)) \\ \text{with respect to} \quad & \omega_i, y_i \in \mathbb{R}^p \\ \text{subject to} \quad & Ma_i - z - y_i = 0, i = 1, \ldots, p \\ & \omega_i = 1, i = 1, \ldots, p \end{aligned} \tag{5.9}$$

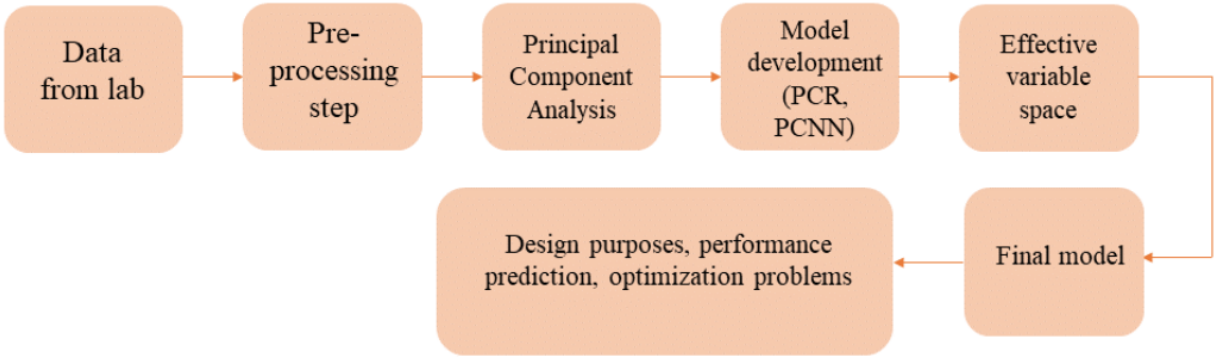

Figure 5.2: A summary of the sequential tasks .

To use an available interior-point method, one should reformulate the above problem to a barrier function type:

$$B = -\ln\det M - \theta\sum_{i}^{m}\mathrm{n}\left(\omega_i^2 - y_i^T y_i\right) \tag{5.10}$$

An interior point method is used where at each step Schur-complement matrix of size $p^2\,(n+1)^2$ is factorized.

The above algorithm is used to determine the p-dimensional and d-dimensional hyper ellipsoid in the original and pseudo space of the data set, respectively.

A summary of the proposed methodology is presented in Figure 5.2. For every performance prediction problem, the framework starts with creating an empirical data base using field or laboratory-produced data. The framework continues with a data preprocessing and data wrangling step to make the data more appropriate and valuable for the main analysis which is developing the predictive models (PCR and PCNN). The framework confront extrapolation with using the developed predictive models over their allowed variable space. The developed models can be used in further performance prediction, design and optimization problems, etc.

The rest of the paper focuses on using the developed framework to define and solve two separate performance related problems: (1) predicting rutting behavior (2) predicting dynamic modulus of asphalt mixture. These two examples are presented, solved, and discussed in section 3 of this paper.

5.4 Results and discussion

The problem of predicting rutting behavior, as one of the most important performance related properties of asphalt mixture, is defined and solved in this section.

5.4.1 Problem (1): Predicting Rutting Behavior

5.4.1.1 Material and laboratory testing

To create a data set for predicting rutting behavior of asphalt mixture, specimens are collected from different locations in the State of Wisconsin. Eighty-three specimens from 21 different mixture are used to perform laboratory testing.

Asphalt mixtures were collected directly from the back of the delivery trucks at the plant site, and for each asphalt mixture the corresponding asphalt binder was sampled during mix plant production. Maximum theoretical specific gravity (G_{mm}) is measured in accordance with AASHTO T209/ASTM D2041. The measured G_{mm} is used to obtain other volumetric properties of the asphalt mixtures. Specimens are compacted using a superpave gyratory compactor to the following dimensions: 150 mm in diameter by 170 mm in height. Specimens are compacted to three different air voids including 4.0%, 7.0%, and 10.0%. Following AASHTO T166/ASTM D2726, the bulk specific gravity values of specimens are determined.

To conduct dynamic modulus test, a 100-mm diameter by 150-mm height cylindrical specimen is cored out of the laboratory compacted specimens. The specimens are trimmed and prepared for dynamic modulus test. The specimens are tested at an effective test temperature of 36.6 °C under repeated sinusoidal load with 25, 10, 1, and 0.1 Hz loading frequency following AASHTO TP-13. The same specimens are used to perform flow number tests under repeated haversine load with the load being applied for a duration of 0.1 second and a dwell period of 0.9 second. The same effective temperature of 36.6°C is selected for flow number test. No confining pressure is applied, and the axial stress is similar to the deviator stress (600 kPa). The accumulated strain at the FN is considered to be the response variable during the modeling procedure.

To study binder shear properties, complex shear modulus test was performed using Dynamic Shear Rheometer (DSR) based on ASTM D7552-09 at 36.6 °C and similar loading frequencies used in the dynamic modulus tests (25, 10, 1, and 0.1 Hz) .

It is worth pointing out that 36.6°C is selected as the temperature that all the laboratory tests are performed at based on Midwestern area of the United States climate condition. In other words, this temperature is considered as a reasonable temperature at which permanent deformation happens in this area, equivalent to a seasonal correction throughout the year. However, what makes the machine learning based models so special is that the model can be retrained and modified based on any test temperatures suiting any climate conditions. Once the data is fed into the framework it will learn the data pattern and the network will modify its weights and biases to fit the new data.

5.4.1.2 Step 1: Data preprocessing

According to the literature Kaloush et al. (2003); Kvasnak et al. (2007); Rodezno et al. (2010); Witzcak (2002), the rutting behavior of an asphalt mixture can be accurately estimated as a function of its component properties. The input variables are selected amongst those properties that have already proven to be important in predicting rutting behavior. However, their importance was also re-examined before selection by performing a multi-factor analysis of variance (ANOVA). The selected material component properties and their ranges that are measured and used in this section are presented in Table 5.1. These properties are selected based on the existing literature and used as the original input variables to predict accumulated strain value at the FN.

To study the quality of the input variables and their interrelationships, the correlation analysis is performed, and the cross-correlation matrix of the input variables is obtained and presented by Table 5.2. Within the matrix, there are 273 elements which their absolute values are greater than 0.1. This means that the corresponding variables are not independent. Besides, there are 41 elements that their absolute values are greater than 0.5 (elements in bold and red text). This means that several on the input variables appear to have strong correlation. Therefore, to eliminate the existing pairwise correlation PCA should be implemented.

Table 5.1: Original input variables of problem (1): predicting rutting behavior.

Variable	Identity	Min.	Max.	Ave.	Std. Ave.
x_1	Binder %	3.400	6.600	5.093	0.773
x_2	G*	210800.543	1163559.917	612179.576	265903.652
x_3	NMAS	12.500	25.000	15.922	3.761
x_4	Passing 3/4”	81.300	100.000	98.554	4.085
x_5	Passing 1/2”	38.300	98.800	87.130	15.199
x_6	Passing 3/8”	34.100	89.900	76.340	15.099
x_7	Passing #4	26.200	72.500	56.248	13.741
x_8	Passing #8	17.500	54.000	42.249	10.512
x_9	Passing #16	14.200	47.400	32.178	8.740
x_{10}	Passing #30	9.600	47.400	32.178	8.740
x_{11}	Passing #50	5.700	18.600	12.022	3.179
x_{12}	Passing #100	3.700	9.800	6.187	1.424
x_{13}	Passing #200	2.800	8.500	4.322	1.115
x_{14}	VMA	10.323	21.000	16.452	2.502
x_{15}	VFA	46.450	91.719	65.189	9.062
x_{16}	Va%	1.019	9.825	5.868	2.088
x_{17}	E*	395.700	2299.400	869.410	411.524

Table 5.2: Pairwise correlation matrix.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	1	0.33	-0.73	0.6	0.71	0.63	0.47	0.45	0.46	0.45	0.41	0.43	0.42	0.61	-0.18	0.35	-0.43
x_2	0.33	1	-0.11	0.08	-0.03	-0.23	-0.38	-0.32	-0.18	-0.06	0.01	0.23	0.35	0.02	-0.02	0.02	0.24
x_3	-0.73	-0.11	1	-0.64	-0.76	-0.69	-0.51	-0.41	-0.39	-0.32	-0.25	-0.30	-0.39	-0.53	0.13	-0.30	0.37
x_4	0.6	0.08	-0.64	1	0.88	0.78	0.58	0.51	0.42	0.32	0.25	0.23	0.24	0.46	-0.21	0.32	-0.45
x_5	0.71	-0.03	-0.76	0.88	1	0.95	0.77	0.71	0.61	0.49	0.41	0.39	0.36	0.56	-0.33	0.44	-0.51
x_6	0.63	-0.23	-0.69	0.78	0.95	1	0.92	0.86	0.75	0.61	0.49	0.30	0.16	0.53	-0.27	0.38	-0.49
x_7	0.47	-0.38	-0.51	0.58	0.77	0.92	1	0.95	0.84	0.69	0.49	0.11	-0.14	0.43	-0.16	0.26	-0.37
x_8	0.45	-0.32	-0.41	0.51	0.71	0.86	0.95	1	0.95	0.83	0.59	0.11	-0.15	0.42	-0.16	0.26	-0.41
x_9	0.46	-0.18	-0.39	0.42	0.61	0.75	0.84	0.95	1	0.96	0.74	0.18	-0.11	0.40	-0.16	0.25	-0.41
x_{10}	0.45	-0.06	-0.32	0.32	0.49	0.61	0.69	0.83	0.96	1	0.84	0.28	-0.06	0.35	-0.15	0.22	-0.37
x_{11}	0.41	0.01	-0.25	0.25	0.43	0.49	0.49	0.59	0.74	0.84	1	0.59	0.17	0.21	-0.20	0.20	-0.25
x_{12}	0.43	0.23	-0.30	0.23	0.39	0.30	0.11	0.11	0.18	0.28	0.59	1	0.82	0.24	-0.28	0.28	-0.15
x_{13}	0.42	0.35	-0.39	0.24	0.36	0.16	-0.14	-0.15	-0.11	-0.06	0.17	0.82	1	0.29	-0.30	0.32	-0.22
x_{14}	0.61	0.02	-0.53	0.46	0.56	0.53	0.43	0.42	0.40	0.35	0.21	0.24	0.29	1	-0.62	0.83	-0.71
x_{15}	-0.18	-0.02	0.13	-0.21	-0.33	-0.27	-0.16	-0.16	-0.16	-0.15	-0.20	-0.28	-0.30	-0.62	1	-0.94	0.52
x_{16}	0.35	0.02	-0.30	0.32	0.44	0.38	0.26	0.26	0.25	0.22	0.20	0.28	0.32	0.83	-0.94	1	-0.63
x_{17}	-0.43	0.24	0.37	-0.45	-0.51	-0.49	-0.37	-0.41	-0.41	-0.37	-0.25	-0.15	-0.22	-0.71	0.52	-0.63	1

Table 5.3: Eigenvalues of the normalized matrix

Number	Eigenvalue	Percent Variance	Cumulative Percent
1	7.979	46.935	46.935
2	2.876	16.920	63.856
3	1.947	11.456	75.311
4	1.539	9.052	84.363
5	0.912	5.363	89.726
6	0.581	3.419	93.145
7	0.421	2.475	95.620
8	0.260	1.529	97.149
9	0.198	1.165	98.314
10	0.152	0.892	99.206
11	0.063	0.371	99.577
12	0.038	0.224	99.801
13	0.016	0.095	99.896
14	0.008	0.049	99.945
15	0.005	0.028	99.973
16	0.003	0.016	99.989
17	0.002	0.011	100.000

5.4.1.3 Step 2: Data wrangling: Dimensionality reduction using PCA

For the pairwise correlation matrix the eigenvalues and their corresponding percent variance are calculated and presented in Tables 5.3 which indicates the fraction of total variation in the data expressed by each eigenvalue. In Table 5.3 the eigenvalues are sorted in descending order meaning that the first eigenvalue represents the highest portion of the total variation, the second one has the second highest portion and so on. By selecting the first five PCs, 89.72% of the variation in the original data will be represented. Adding the sixth PC has an insignificant impact on the overall represented variation. Thus, the first five PCs are selected to be used as the pseudo-input variables.

PCs can be obtained using Eq. 5.11

$$pc_i = \sum_{j=1}^{17} \alpha_{ij} x_j + \beta_i \tag{5.11}$$

where $i = 1, \ldots, 5$, the αij is the corresponding coefficients, the βi are constants, and the xj's are the original input variables. Eq. 5.11 can be presented in matrix notation as in Eq. 5.12

$$\mathbf{p} = \mathbf{M}\mathbf{z} + \mathbf{n} \tag{5.12}$$

where

$$\mathbf{p} = \begin{bmatrix} pc_1 \\ pc_2 \\ pc_3 \\ pc_4 \\ pc_5 \end{bmatrix}$$

$$\mathbf{M}^T = \begin{bmatrix} 3.41 \times 10^{-1} & -2.04 \times 10^{-1} & 2.89 \times 10^{-1} & -1.94 \times 10^{-1} & 3.71 \times 10^{-1} \\ -7.60 \times 10^{-8} & -1.17 \times 10^{-6} & 1.31 \times 10^{-6} & 1.05 \times 10^{-7} & 2.47 \times 10^{-6} \\ -6.59 \times 10^{-2} & 3.22 \times 10^{-2} & -4.53 \times 10^{-2} & 9.10 \times 10^{-2} & -7.97 \times 10^{-3} \\ 6.34 \times 10^{-2} & -1.31 \times 10^{-2} & 1.86 \times 10^{-2} & -9.38 \times 10^{-2} & -7.13 \times 10^{-3} \\ 2.11 \times 10^{-2} & -2.33 \times 10^{-3} & 5.26 \times 10^{-3} & -1.63 \times 10^{-2} & -9.03 \times 10^{-3} \\ 2.17 \times 10^{-2} & 7.23 \times 10^{-3} & 1.55 \times 10^{-3} & -1.23 \times 10^{-2} & -1.13 \times 10^{-2} \\ 2.13 \times 10^{-2} & 2.01 \times 10^{-2} & -2.82 \times 10^{-3} & -7.27 \times 10^{-3} & -7.11 \times 10^{-3} \\ 2.79 \times 10^{-2} & 2.89 \times 10^{-2} & -2.98 \times 10^{-3} & 3.79 \times 10^{-3} & 1.41 \times 10^{-3} \\ 3.28 \times 10^{-2} & 3.22 \times 10^{-2} & 4.43 \times 10^{-3} & 2.35 \times 10^{-2} & 1.71 \times 10^{-2} \\ 3.73 \times 10^{-2} & 3.39 \times 10^{-2} & 1.53 \times 10^{-2} & 4.98 \times 10^{-2} & 3.14 \times 10^{-2} \\ 7.08 \times 10^{-2} & 3.30 \times 10^{-2} & 7.38 \times 10^{-2} & 1.54 \times 10^{-1} & -9.05 \times 10^{-3} \\ 1.08 \times 10^{-1} & -2.03 \times 10^{-1} & 2.33 \times 10^{-1} & 2.30 \times 10^{-1} & -2.72 \times 10^{-1} \\ 9.18 \times 10^{-2} & -4.06 \times 10^{-1} & 2.29 \times 10^{-1} & 6.69 \times 10^{-2} & -3.18 \times 10^{-1} \\ 9.98 \times 10^{-2} & -8.77 \times 10^{-2} & -1.12 \times 10^{-1} & -1.14 \times 10^{-2} & 8.85 \times 10^{-2} \\ -1.73 \times 10^{-2} & 3.21 \times 10^{-2} & 4.57 \times 10^{-2} & -2.69 \times 10^{-2} & -9.43 \times 10^{-4} \\ 9.75 \times 10^{-2} & -1.42 \times 10^{-1} & -1.95 \times 10^{-1} & 7.31 \times 10^{-2} & 4.70 \times 10^{-2} \\ -5.50 \times 10^{-4} & 2.73 \times 10^{-4} & 8.20 \times 10^{-4} & -8.92 \times 10^{-6} & 1.36 \times 10^{-4} \end{bmatrix}$$

$$\mathbf{n} = \begin{bmatrix} -1.72 \times 10 \\ 5.87 \times 10^{-1} \\ -8.21 \\ 7.36 \\ -4.85 \times 10^{-1} \end{bmatrix}$$

The PCs obtained in this way will be used in the main analysis.

5.4.1.4 Step 3: Model development

To validate the stability of the machine learning model and examine how well it would generalize to new data, cross validation technique is implemented. In this case of having limited amount of data (sample size of 83), k-fold cross validation technique leads to a less biased model compare to other methods. Because it ensures that every observation from the original data set has the chance of appearing in training and test sets. K-fold cross validation technique works in a way that the given data set is randomly partitioned into k subsets. K-1 of these subsets are used to train the model and the remaining subset is used to test the model. The process is repeated k times and each subset can be used as test data exactly once.

Based on the size of rutting data set (83 data points), the data set is randomly partitioned into 3 folds. Results of the developed models are presented in this section, and their performance in predicting the response variable over the defined effective variable space are examined and discussed.

After examining all the possible regression structures, the second-order quadratic linear regression model fitting the measured response the best, is presented in Eq. 5.13

$$\begin{aligned} \widehat{y} = c_0 \ + \ c_1 * pc_1 \ + \ c_2 * pc_2 \ + \ c_3 * pc_3 \ + \ c_4 * pc_4 \ + \ c_5 * pc_5 + \ c_6 * pc_1 * pc_2 \\ + \ c_7 * pc_2 * pc_4 \ + \ c_8 * pc_1 * pc_3 \ + \ c_9 * pc_2 * pc_3 \ + \ c_{10} * pc_3 * pc_5 \end{aligned} \tag{5.13}$$

where, $c_0 = 1.64 \times 10^4$, $c_1 = -8.89 \times 10^2$, $c_2 = -1.24 \times 10^3$, $c_3 = -1.24 \times 10^3$, $c_4 = -92.41$, $c_5 = 6.55 \times 10^2$, $c_6 = 1.58 \times 10^2$, $c_7 = -4.35 \times 10^2$, $c_8 = 2.35 \times 10^2$, $c_9 = 3.74 \times 10^2$, and $c_{10} = -8.28 \times 10^2$.

Using Eq. 5.6 the developed network's connection weights and biases are presented by the following matrices:

$$W^T = \begin{bmatrix} -0.45 & 1.70 & -0.81 & -0.85 & -1.30 \\ 0.01 & -1.46 & 0.28 & 0.11 & -1.93 \\ 1.32 & -1.39 & -0.03 & -0.22 & 0.90 \\ -0.09 & 0.18 & -0.14 & -0.32 & -0.15 \\ 0.09 & -0.18 & 0.14 & 0.32 & 0.15 \\ -0.61 & 0.26 & -1.19 & -0.81 & 0.12 \\ 0.09 & -0.18 & 0.14 & 0.32 & 0.15 \\ 0.44 & -1.50 & -0.41 & 1.05 & -0.64 \\ 0.28 & -0.63 & 0.83 & 0.74 & 1.16 \\ -0.00 & 0.17 & 0.00 & -1.38 & -0.29 \end{bmatrix}$$

$$W_H = \begin{bmatrix} 0.67 \\ 0.35 \\ -1.25 \\ -0.29 \\ 0.29 \\ -1.71 \\ 0.29 \\ -0.26 \\ 0.97 \\ 0.99 \end{bmatrix} \qquad B_H = \begin{bmatrix} -0.08 \\ -1.21 \\ -1.74 \\ -1.53 \\ 0.57 \\ -0.57 \\ -1.32 \\ -0.57 \\ 1.46 \\ -1.75 \end{bmatrix} \qquad , \; B_0 = [1.33]$$

Using several statistics, the performance results of the PCR and PCNN models are discussed and presented in Table 5.4. The first statistic is the "average difference (AD)" defined as

$$AD = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i). \tag{5.14}$$

AD is an estimate of systematic model bias, n is the number of input vectors, yi is the ith measured response value, and $\widehat{y}$ is the ith fitted response value. The second statistical component

Table 5.4: Statistical Analysis of PCR and PCNN Modeling (na*: not applicable)

		PCR			PCNN		
	Statistics	Fold 1	Fold 2	Fold 3	Fold 1	Fold 2	Fold 3
Training	AD	0	0	0	34.99	-242.99	46.19
	AAD	1497.02	1705.55	1514.59	729.41	1350.87	944.83
	r_{fit}	0.83	0.82	0.85	0.96	0.87	0.94
	R^2	0.69	0.68	0.72	na*	na*	na*
Testing	AD	626.73	-129.91	-226.1	-98.24	149.2	-169.6
	AAD	2007.47	1515.74	2110.64	694.53	719.9	1037.28
	r_{fit}	0.79	0.8	0.73	0.97	0.95	0.92
	R^2	na*	na*	na*	na*	na*	na*

is the "average absolute difference (AAD)" which shows the average closeness of the fitted value to the measured response value. AAD is defined as

$$AAD = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}|. \tag{5.15}$$

The third statistical component, $r$fit, is the correlation of y and $\widehat{y}$ defined as

$$r_{fit} = \frac{n\sum_{i=1}^{n} y_i\hat{y}_i - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n}\hat{y}_i)}{\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}\sqrt{n\sum_{i=1}^{n}\hat{y}_i^2 - (\sum_{i=1}^{n}\hat{y}_i)^2}} \tag{5.16}$$

Higher r_{fit} (with maximum value of 1) indicates better fit. The last statistical component is R-squared (R^2) also known as the coefficient of determination. R^2 is the portion of the variance in the dependent variable that can is predictable from the independent variables and can be explicated by the fitted model. It is applicable to the PCR (training set) since it is linear in its parameters and not to PCNN due to the non-linear nature of its parameters.

Researchers have developed several predictive models for rutting in asphalt mixture Witzcak (2002); Kaloush et al. (2003); Rodezno et al. (2010); Kvasnak et al. (2007); Andrei et al. (1999). A summary of the most well-known conventional models, their parameters and prediction accuracy expressed in R^2 is presented in Figure 5.3.

Kaloush et.al.	$\log\frac{\epsilon_p}{\epsilon_r} = -31.5552 + 1.734\log T + 0.39937\log N$ $R^{2^*} = 0.664$ $\log\frac{\epsilon_p}{\epsilon_r} = 0.3082 + 0.3534\log N$ $R^{2^*} = 0.55$	ϵ_p = accumulated permanent strain ϵ_r = resilient strain N = number of load repetitions T = mix temperature (°F)
Leahy et. al.	$\log\frac{\epsilon_p}{\epsilon_r} = -6.631 + 0.435\log N + 2.767\log T + 0.11\log S + 0.118\log\eta + 0.93\log V_{beff} + 0.501\log V_a$ $R^{2^*} = 0.76$	ϵ_p = accumulated permanent strain ϵ_r = resilient strain N = number of load repetitions T = mix temperature (°F) S = deviatoric stress (psi_ η = viscosity at70 °F (10^6 *poise*) V_{beff} = effective binder content (%volume) V_a = percent air voids
Ayres et. al.	$\log\frac{\epsilon_p}{\epsilon_r} = -4.80661 + 2.58155\log T + 0.42956\log N$ $R^{2^*} = 0.725$	ϵ_p = accumulated permanent strain ϵ_r = resilient strain N = number of load repetitions T = mix temperature (°F)
Kaloush et. al.	$FN = (432367000)T^{-2.215}Visc^{0.312}V_{beff}^{-2.6604}V_a^{-0.1525}$ $\frac{S_e}{S_y} = 0.534,\ R^{2^*} = 0.72$	T = test temperature (°F) $Visc$ = binder viscosity at 70°F (10^6 *poise*) V_{beff} = effective asphalt content (%volume) V_a = air voids (%)
Kvasnak et.al.	$logFN = 2.866 + 0.00613\,Gyr + 3.86\,Visc - 0.072VMA + 0.282\,P_4 - 0.051\,P_{16} + 0.075\,P_{200}$ $R^{2^*} = 0.91$	Gyr = number of gyrations $Visc$ = binder viscosity at test temperature (0.02 *to* 0.286 * 10^6 *poise*) VMA = voids in mineral aggregates P_4, P_{16}, P_{200} = percent aggregate passing from sieve sizes #4, #16, and #200
2002 AASHTO design guide	$\frac{\epsilon_p}{\epsilon_r} = a\,T^b N^c$	ϵ_p = accumulated plastic strain at N repetition of load ϵ_r = resilient strain of the asphalt material as a function of mix properties, temperature, and time rate of loading N = number of load repetitions a, b, c = non-linear regression coefficients

Figure 5.3: Summary of the well-known existing models

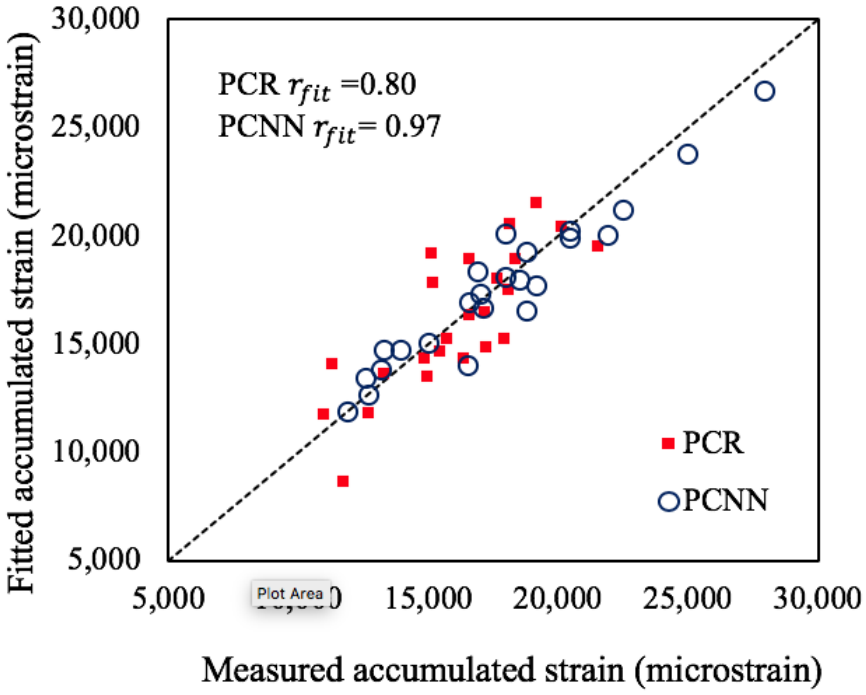

Figure 5.4: Measured values of accumulated strain

* It should be mentioned that the ability of the developed model in fitting to the empirical data should be expressed in terms of the r_{fit} value and not the R^2 value.

Although the reported R^2 values for Leahy, Ayres, and Kaloush models seem reasonable, they did not use separate data sets for training and testing possibly resulting in biased and over fitted models. Comparing the results obtained from the PCR and PCNN models with the previous prediction models used in the AASHTO design procedure, the developed models are performing significantly better with $r_{\text{fit}} = 0.8$ for PCR model and $r_{\text{fit}} = 0.97$ for PCNN model. In Comparison between PCR and PCNN, one can see that although the PCR works well in predicting the response variable, PCNN provides the best fit for both training and test sets. Measured values of accumulated strain and the fitted values are presented in Figure 5.4. The measured and fitted values are close to the line of equality meaning that the fitted values by PCR and PCNN have a strong correlation with the measured one.

5.4.2 Problem (2): Predicting dynamic modulus

The problem of predicting dynamic modulus, as one of the most important performance characteristics of asphalt mixture, is solved and discussed in this section.

5.4.2.1 Material and laboratory testing

To create a data set for dynamic modulus prediction, 27 specimens from nine different asphalt mixtures are selected and used to perform laboratory testing. AASHTO T209/ASTM D2041 is used to measure maximum theoretical values of specific gravity (G_{mm}). The measured G_{mm} values are used to obtain other volumetric properties of the asphalt mixtures.

The samples are 150 mm in diameter and 38 mm in thickness cut from suparpave gyratory compacted specimens. The dynamic modulus test in indirect tension mode of testing is performed at three temperatures (0.4, 17.1, and 33.8 °C) and nine loading frequencies (25, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1 Hz) in accordance with AASHTO TP 62-07. Based on ASTM D7552-09 dynamic shear rheometer (DSR) test is conducted to measure complex shear modulus of asphalt binder. The test is performed at a wide variety of temperatures (-10 to 54 °C) and frequencies (0.1 Hz to 25 Hz) including the exact test temperatures and loading frequencies which are used in the mixture dynamic modulus test. It is important to mention that, the present research uses a consistent definition of frequency, meaning that to predict the dynamic modulus value of an asphalt mixture for example at 4 °C and 25 Hz, one should input in the model the complex shear modulus of asphalt binder, $|G^*|$, at 4 °C and 25 Hz. A summary of the nine different mixtures' properties is presented in Table 5.5. Using the laboratory test results on 27 specimens, a database of 243 data points is created to be used in further modeling.

5.4.2.2 Step 1: Data preprocessing

According to the literature dynamic modulus of asphalt mixture can be estimated by the material components properties Andrei et al. (1999); Bari and Witczak (2007); Christensen Jr et al. (2003); Al-Khateeb et al. (2006); Sakhaeifar et al. (2017). The input variables are selected amongst

Table 5.5: Properties of nine asphalt mixtures

	Mix 1	Mix 2	Mix 3	Mix 4	Mix 5	Mix 6	Mix 7	Mix 8	Mix 9
Binder performance grade	58-28	58-28	58-28	58-34	58-34	58-34	64-28	64-34	64-28
% Vbeff	4.20	4.10	4.10	3.90	3.50	4.30	4.20	4.00	4.60
%VMA	13.50	13.50	13.60	13.10	12.50	13.90	13.70	13.40	14.40
% VFA	70.30	70.40	70.60	69.60	68.10	71.20	70.80	70.20	72.30
Gmb	2.32	2.31	2.31	2.32	2.31	2.32	2.31	2.32	2.31
Gmm	2.41	2.46	2.51	2.48	2.64	2.46	2.48	2.51	2.44
% VA	4.01	3.99	3.99	3.98	3.98	4.03	4	3.99	3.98
% passing 3/4”	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
% passing 1/2”	93.90	96.40	87.20	93.50	95.10	96.40	94.10	94.40	94.20
% passing 3/8”	77.50	84.60	73.70	76.40	83.10	87.30	83.40	82.00	80.90
% passing #4	49.80	53.10	48.40	52.20	52.20	60.90	63.80	48.20	58.60
% passing #8	34.40	38.40	35.10	43.60	38.80	46.90	47.10	34.90	46.00
% passing #30	16.70	18.70	17.90	20.90	18.80	23.40	21.70	19.20	25.90
% passing #50	10.30	10.80	10.90	11.40	9.90	12.40	11.90	11.80	13.80
% passing #100	6.10	5.90	6.40	5.80	5.40	6.10	6.60	6.10	7.20
% passing #200	3.60	3.30	6.20	3.30	3.50	3.40	4.00	3.10	4.00

those properties that have already proven to be important in predicting dynamic modulus. However, their importance was also re-examined before selection by performing a multi-factor analysis of variance (ANOVA).The selected component properties and their ranges that are obtained from laboratory testing and used in the present study are summarized in Table 5.6.

To evaluate the quality of the predictors, correlation analysis is performed, and the result is presented by Table 5.7. Fifty Elements of the pairwise correlation matrix with absolute value of greater than 0.5 are shown in bold red text indicating that there is strong correlation between the predictors. Therefore, to eliminate the existing correlation PCA is implemented.

5.4.2.3 Step 2: Data wrangling: Dimensionality reduction using PCA

The eigenvalues of the correlation matrix are calculated and presented in Table 5.8. The eigenvalues and their corresponding contribution to the total variance are sorted in descending order. According to the table, 95.8% of the variation in the original data is expressed by the first five PCs. Adding the sixth PC has an insignificant impact on the overall represented variation. Therefore, the first five PCs are selected to be used as input variables in the modeling problem.

Table 5.6: Original Input Variables of Problem (1): Predicting Dynamic Modulus

Variable	Identity	Min.	Max.	Ave.	Std. Dev.
x_1	Cum. % retained on 3/4"	3.60	13.00	6.11	2.63
x_2	Cum. % retained on 3/8"	12.68	26.29	19.01	4.11
x_3	Cum. % retained on #4	36.20	51.76	45.86	5.319
x_4	Cum. % retained on #8	52.87	65.70	59.42	5.06
x_5	Cum. % retained on #30	74.06	83.30	79.63	2.76
x_6	Cum. % retained on #50	86.22	90.12	88.57	1.15
x_7	Cum. % retained on #100	92.81	94.59	93.83	0.48
x_8	% Passing from #200	3.07	6.18	3.81	0.89
x_9	$Log\|G^*\|$	-2.29	3.03	0.50	1.26
x_{10}	Phase angle (degree)	28.15	79.17	52.86	11.54
x_{11}	$V_{beff}\%$	3.50	4.60	4.10	0.29
x_{12}	VMA	12.50	14.40	13.51	0.49
x_{13}	VFA	68.10	72.30	70.40	1.08
x_{14}	Va%	3.98	4.01	3.99	0.01

Table 5.7: Correlation Matrix for the Input Variables

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
x_1	**1**	**0.832**	0.412	0.366	0.294	0.119	-0.269	**0.905**	-0.044	-0.058	0.003	0.04	0.049	0.013
x_2	**0.832**	**1**	**0.597**	0.458	0.391	0.246	-0.109	**0.583**	-0.035	0.106	-0.061	-0.099	-0.089	-0.115
x_3	0.412	**0.597**	**1**	**0.918**	**0.756**	**0.596**	0.425	0.133	-0.019	0.154	-0.465	-0.485	-0.49	-0.111
x_4	0.366	0.458	**0.918**	**1**	**0.87**	**0.687**	0.375	0.169	-0.028	0.237	-0.388	-0.412	-0.424	0.212
x_5	0.294	0.391	**0.756**	**0.87**	**1**	**0.919**	**0.618**	0.112	-0.021	0.235	**-0.585**	**-0.631**	**-0.633**	0.3
x_6	0.119	0.246	**0.596**	**0.687**	**0.919**	**1**	**0.794**	-0.009	0.003	0.203	**-0.741**	**-0.796**	**-0.806**	0.209
x_7	-0.269	-0.109	0.425	0.375	**0.618**	**0.794**	**1**	-0.414	0.036	0.047	**-0.854**	**-0.886**	**-0.892**	-0.087
x_8	**0.905**	**0.583**	0.133	0.169	0.112	-0.009	-0.414	**1**	-0.032	-0.102	0.179	0.238	0.238	0.142
x_9	-0.044	-0.035	-0.019	-0.028	-0.021	-0.003	0.036	-0.032	**1**	**-0.808**	0.021	0.016	0.013	0.034
x_{10}	-0.058	0.106	0.154	0.237	0.235	0.203	0.047	-0.102	**-0.808**	**1**	0.09	0.024	0.014	0.3
x_{11}	0.003	-0.061	-0.465	-0.388	**-0.585**	**-0.741**	**-0.854**	0.179	0.021	0.09	**1**	**0.988**	**0.985**	0.372
x_{12}	0.04	-0.099	-0.485	-0.412	**-0.631**	**-0.796**	**-0.886**	0.238	0.016	0.024	**0.988**	**1**	**0.998**	0.321
x_{13}	0.049	-0.089	-0.49	-0.424	**-0.633**	**-0.806**	**-0.892**	0.238	0.013	0.014	**0.985**	**0.998**	**1**	0.301
x_{14}	0.013	-0.115	-0.111	0.212	0.3	0.209	-0.087	0.142	0.034	0.3	0.372	0.321	0.301	**1**

Table 5.8: Eigenvalues of the **Z** Matrix and the Corresponding Percent Variance

Number	Eigenvalue	Percent Variation	Cumulative Percent Variation
1	6.0225	43.018	43.018
2	3.2193	22.995	66.013
3	1.9746	14.104	80.118
4	1.4174	10.124	90.242
5	0.7850	5.607	95.848
6	0.3176	2.269	98.117
7	0.1091	0.779	98.896
8	0.0778	0.556	99.452
9	0.0549	0.392	99.844
10	0.0218	0.156	100

The PCs are obtained using Eq. 5.11 and 5.12 where matrices **M** and **n** are as follows

$$M^T = \begin{bmatrix} 0.03 & 0.19 & -0.08 & -0.06 & -0.09 \\ 0.03 & 0.11 & -0.04 & -0.05 & 0.04 \\ 0.06 & 0.04 & 0.00 & 0.00 & 0.01 \\ 0.06 & 0.05 & 0.02 & 0.04 & 0.07 \\ 0.13 & 0.06 & 0.05 & 0.09 & -0.01 \\ 0.33 & 0.01 & 0.09 & 0.16 & -0.20 \\ 0.71 & -0.54 & 0.05 & 0.13 & -0.20 \\ -0.03 & 0.52 & -0.20 & -0.70 & -0.53 \\ -0.01 & -0.05 & -0.40 & 0.43 & 0.12 \\ 0.00 & 0.01 & 0.06 & -0.02 & 0.00 \\ -1.26 & 0.64 & 0.52 & 0.61 & 0.67 \\ -0.75 & 0.37 & 0.21 & 0.28 & 0.31 \\ -0.34 & 0.17 & 0.09 & 0.12 & 0.14 \\ -4.41 & 18.38 & 47.08 & 76.24 & -40.48 \end{bmatrix}, n = \begin{bmatrix} -55.95 \\ -58.54 \\ -218.20 \\ -352.79 \\ 174.78 \end{bmatrix}$$

5.4.2.4 Step 3: Model development

To examine the stability of the developed model against of an unseen data set and based on the size of dataset, a holdout cross validation technique is used. During this procedure the given dataset is randomly assigned to two subsets, d0 and d1, called the training set and the test set, respectively. The training set contains 80% of the data points and the test set contains 20% of the data points which means 80% of the data points are used to train the model and the rest is used to evaluate the performance of the trained model.

After examining all the possible regression structures, the best reduced third-order cubic linear regression model fitting the measured response, y, is presented in Eq. 5.17

$$\begin{aligned}\widehat{y} = c_0 + c_1\ pc_1 + c_2\ pc_2 + c_3\ pc_3 + c_4\ pc_4 + c_5\ pc_5 + c_6\ pc_1\ pc_2 + c_7\ pc_1\ pc_3 + c_8\ pc_1\ pc_4 \\ +c_9\ pc_1\ pc_5 + c_{10}\ pc_2\ pc_3 + c_{11}\ pc_2\ pc_4 + c_{12}\ pc_2\ pc_5 + c_{13}\ pc_3\ pc_4 + c_{14}\ pc_3\ pc_5 + \\ c_{15}\ pc_4\ pc_5 + c_{16}\ pc_1\ pc_2\ pc_3 + c_{17}\ pc_1\ pc_2\ pc_4 + c_{18}\ pc_1\ pc_2\ pc_5 + c_{19}\ pc_1\ pc_3\ pc_4 \\ +c_{20}\ pc_1\ pc_3\ pc_5 + c_{21} pc_2\ pc_3\ pc_4 + c_{22}\ pc_1\ pc_4\ pc_5 + c_{23}\ pc_2\ pc_4\ pc_5 + c_{24}\ pc_3\ pc_4\ pc_5\end{aligned} \tag{5.17}$$

where, $c_0 = 6.59$; $c_1 = 2.58$; $c_2 = 4.4$; $c_3 = -0.36$; $c_4 = 0.49$; $c_5 = 1.93$; $c_6 = -0.33$; $c_7 = -0.77$; $c_8 = -1.69$; $c_9 = 0.15$; $c_{10} = -1.65$; $c_{11} = -4.68$; $c_{12} = 4.81$; $c_{13} = 0.7$; $c_{14} = -0.85$; $c_{15} = -1.58$; $c_{16} = -0.17$; $c_{17} = -0.79$; $c_{18} = 1.83$; $c_{19} = 0.04$; $c_{20} = 0.18$; $c_{21} = 0.42$; $c_{22} = 0.05$; $c_{23} = 0.32$; $c_{24} = 0.06$.

The networks weights and biases are presented in the following matrices:

$$W^T = \begin{bmatrix} -0.511 & 0.134 & 0.654 & -1.064 & -0.267 \\ -0.315 & -0.147 & -0.267 & 0.177 & -1.047 \\ -0.060 & -1.266 & 0.759 & -1.248 & -0.331 \\ -0.075 & 0.022 & 0.208 & 0.015 & 0.167 \\ -0.074 & 0.022 & 0.206 & 0.015 & 0.165 \\ 0.103 & -0.177 & 1.253 & -1.045 & 0.535 \\ 0.078 & -0.020 & -0.231 & -0.014 & -0.172 \\ 0.238 & 0.070 & -0.885 & 0.848 & 0.943 \\ 0.123 & 0.456 & -0.387 & 1.547 & -0.017 \\ -0.079 & 0.020 & 0.213 & 0.014 & 0.173 \end{bmatrix}$$

$$W_H = \begin{bmatrix} 0.869 \\ -0.886 \\ 0.632 \\ -0.291 \\ -0.288 \\ -0.859 \\ 0.299 \\ 0.556 \\ 0.971 \\ -0.299 \end{bmatrix} \quad B_H = \begin{bmatrix} 0.162 \\ 0.710 \\ 0.319 \\ -0.008 \\ -0.009 \\ 0.570 \\ 0.007 \\ 0.290 \\ -0.373 \\ -0.007 \end{bmatrix} \quad B_0 = [0.148]$$

Performance results of the PCR and PCNN models are compared with three of the well-known predictive models for dynamic modulus base on the statistics introduced in section 5.4.1.4. A summary of these models equations and parameters definition as well as their performance in comparison with PCR and PCNN are presented by Figure 5.5 Bari and Witczak (2007); Christensen Jr et al. (2003); Al-Khateeb et al. (2006) and Table 5.9, respectively.

Modified Witczak	$$\log\lvert E^*\rvert = -0.349 + 0.754(\lvert G_b^*\rvert^{-0.0052})(6.65 - 0.032\,\rho_{200} - 0.0027\,(\rho_{200})^2 + 0.011\rho_4 - 0.0001\,(\rho_4)^2 + 0.006\,\rho_{\frac{3}{8}} - 0.00014\left(\rho_{\frac{3}{8}}\right)^2 - 0.08\,V_a - 1.06\,(\frac{V_{beff}}{V_{beff}+V_a})) + \frac{2.558+0.032\,V_a+0.713\left(\frac{V_{beff}}{V_{beff}+V_a}\right)+0.0124\,\rho_{\frac{3}{8}}-0.0001\left(\rho_{\frac{3}{8}}\right)^2-0.0098\,\rho_{\frac{3}{4}}}{1+\exp(-0.7814-0.5785\log\lvert G_b^*\rvert+0.8835\log\delta_b)}$$	$\lvert E^*\rvert$ = dynamic modulus (psi) $\lvert G^*\rvert$ = binder shear modulus (psi) δ_b = binder phase angle (degree) $\rho_{\frac{3}{4}}$, $\rho_{\frac{3}{8}}$, ρ_4 = cumulative percent aggregate retained on sieve seizes ¾'', 3/8" and no.4 ρ_{200} = percent aggregate passing from sieve no.200 V_a = percent air voids V_{beff} = effective binder content
Hirsch	$$\lvert E^*\rvert = P_C\left[4200000\left(1-\frac{VMA}{100}\right)+3\lvert G_b^*\rvert\left(\frac{VFA*VMA}{1000}\right)\right] + \frac{1-P_c}{\frac{1-VMA}{100}+\frac{VMA}{3\lvert G_b^*\rvert(VFA)}}$$ $$P_c = \frac{\left(20+\frac{3\lvert G_b^*\rvert(VFA)}{VMA}\right)^{0.58}}{650+\left(\frac{3\lvert G_b^*\rvert(VFA)}{VMA}\right)^{0.58}}$$	$\lvert E^*\rvert$ = dynamic modulus (psi) $\lvert G^*\rvert$ = binder shear modulus (psi) VMA = voids in mineral aggregate VFA = voids filled with asphalt Pc = aggregate contact volume
Alkhateeb	$$\lvert E^*\rvert = 3\left(\frac{100-VMA}{100}\right)\left(\frac{\left(90+1.45\,\frac{\lvert G^*\rvert}{VMA}\right)^{0.66}}{1100+\left(0.13\,\frac{\lvert G^*\rvert}{VMA}\right)^{0.66}}\right)\lvert G_g^*\rvert$$	$\lvert E^*\rvert$ = dynamic modulus (psi) $\lvert G^*\rvert$ = binder shear modulus (psi) VMA = voids in mineral aggregate $\lvert G_g^*\rvert$ = complex shear modulus of binder in glassy state

Figure 5.5: Summary of the well-known conventional dynamic modulus

Table 5.9: Statistical comparison of PCR and PCNN

		Average difference (MPa)	Average absolute difference (MPa)	r_{fit}	R^2
PCR	Training	3.9	575.3	0.996	0.99
	Testing	-162.3	718.9	0.995	na
PCNN	Training	13.2	380.7	0.997	na
	Testing	9.7	337.5	0.997	na
Modified Witczak		-2460	3152.1	0.93	0.88
Hirsch		1241.6	1785.7	0.95	0.91
Alkhateeb		2844.5	2984.5	0.95	0.90

According to the obtained r_{fit} values, the predicted values of dynamic modulus by PCR and PCNN models have a strong correlation (0.99) with the measured ones thus, both PCR and PCNN are capable of predicting the response. The corresponding values of rfit for modified Witczak, Hirsch, and Alkhateeb models are 0.93, 0.95, and 0.95 respectively, which seem reasonable. However, r_{fit} can be biased since it shows if the measured response increases then the predicted response will increase and vice versa. Considering other statistics, AD and AAD, shows that these values are significantly higher for modified Witczak, Hirsch, and Alkhateeb models. Meaning that the fitted values by these models are not as close as the ones fitted by PCR and PCNN to the response value. In other words, AD and AAD indicate that the Hirsch and Alkhateeb models are overpredicting and the modified Witczak model is underpredicting the response variables. Measured values of dynamic modulus and the fitted values are presented in Figure 5.7. The measured and fitted values are close to the line of equality meaning that the fitted values by PCR and PCNN have a strong correlation with the measured one.

Figure 5.6: Summary of the well-known conventional dynamic modulus

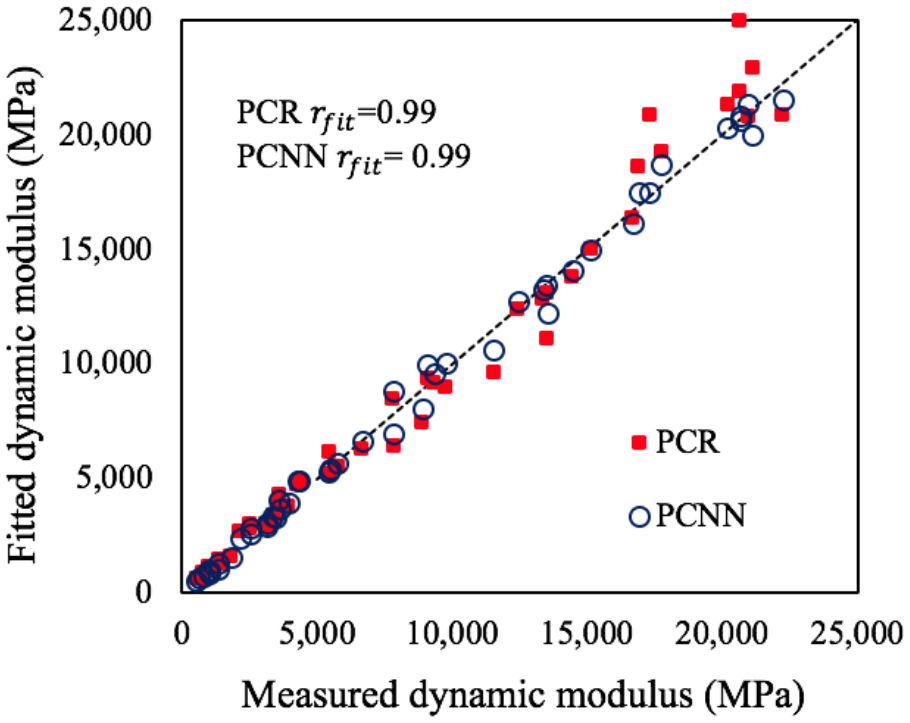

Figure 5.7: Measured values of dynamic modulus versus fitted values .

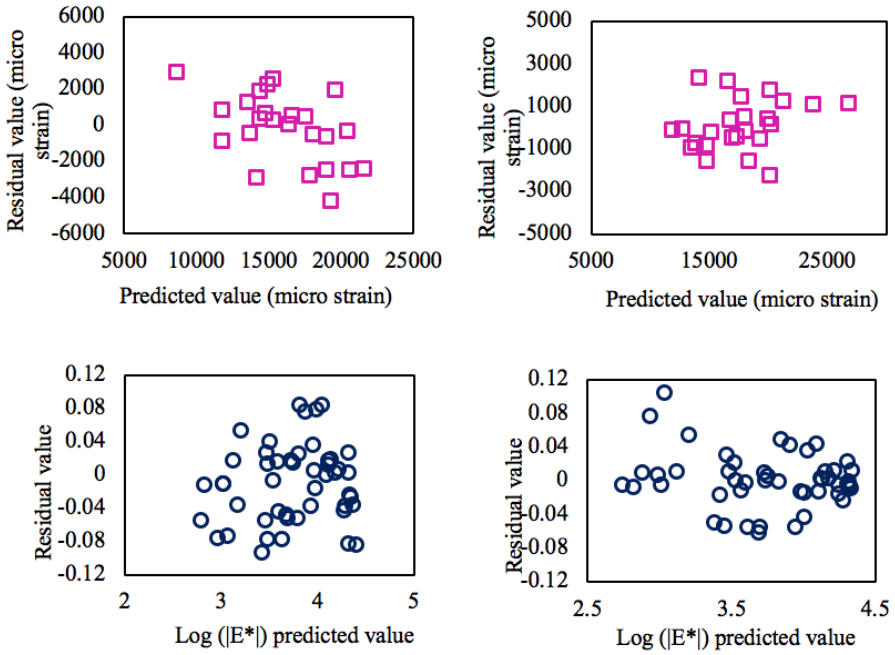

Figure 5.8: Plot of the residuals.

5.5 Model Validation

The task of confirming that the outputs of a statistical model have enough fidelity to the output of data generating process is called model validation. The difference between estimated and measured values of the response (residual) is assumed to be a random error which is normally distributed with a mean of zero and unknown variance Devore (2011). To examine the adequacy of these assumptions, two sets of residual diagnostic analyses are implemented and presented in Figure 5.8. If the residuals are random error terms, the residual plot should contain no obvious pattern. According to the residual plot this assumption is satisfied for both PCR and PCNN in rutting as well as dynamic modulus prediction models. The assumption of normality can be checked by normal probability plot in a way that if the residual distribution is normal, their plot will resemble a straight line. According to Figure 5.9 the data points are located around a straight line. Therefore, the normality assumption does not appear to be violated.

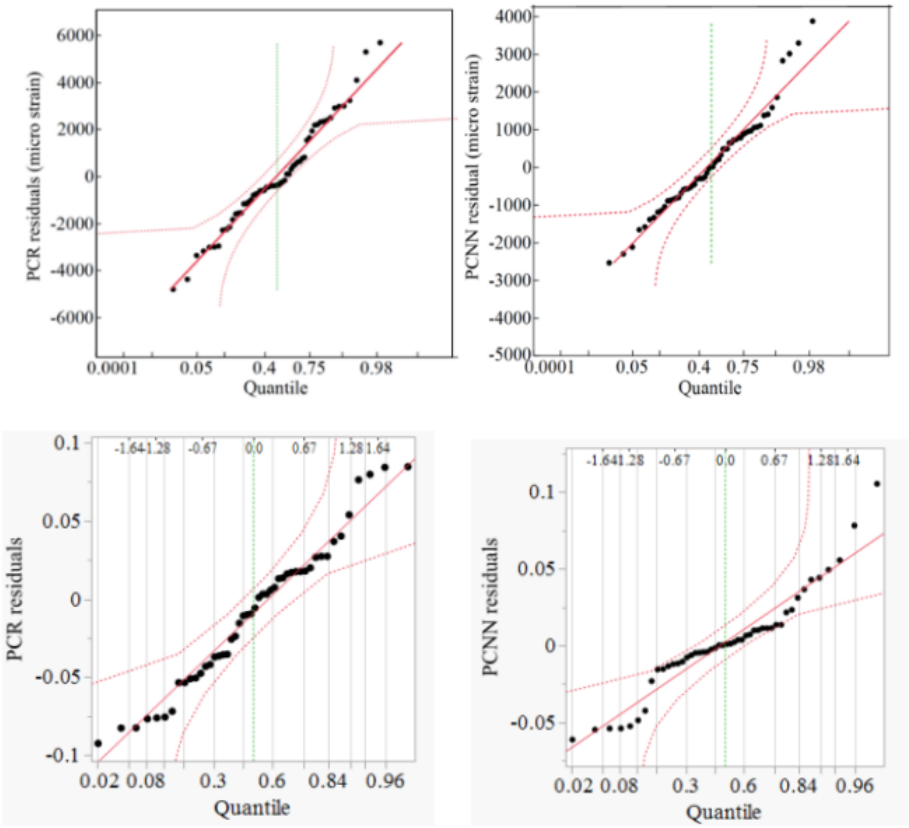

Figure 5.9: Normal probability plot.

5.6 Variable Importance Analysis (VIA)

Once the performance model has been identified, design engineers may desire to know which factor (predictor variable) included in the model has the strongest influence on the response variable. If variation of a specific factor causes high variability in the response, that effect is important relative to the model. Variable importance analysis is one of the global sensitivity analysis methods and is based on the functional decomposition idea by Sobol Sobol' (1990). He proved that one can decomposed a function $y = f(\mathbf{x}_1, \ldots, \mathbf{x}_p) = f_0 + f_1(\mathbf{x}_1) + \ldots + f_p(\mathbf{x}_p) + f_{12}(\mathbf{x}_1, \mathbf{x}_2) + \ldots$ into the sum of lower dimensional functions. The variability of these lower dimensional functions assesses the importance of input variables in terms of effect indexes which indicate relative importance of the variables and are presented as the main effect

$$S_i = \frac{V_i}{V(y)} = \frac{V(E(y|x_i))}{V(y)} \tag{5.18}$$

and the total effect

$$S_{\mathrm{Ti}} = \frac{V_{\mathrm{Ti}}}{V(y)} = \frac{V(y) - V(E(y|x_{-i}))}{V(y)} \tag{5.19}$$

where $\mathbf{x}_i$ represents a random input (vector) variable and $\mathbf{x}_{-i}$ indicates all input variables except $\mathbf{x}_i$. The unique contribution of the input variable $\mathbf{x}_i$ to the total variation of the response, y, is represented by main effect, while the total effect represents the overall contribution of $\mathbf{x}_i$ on y which includes all interaction terms Wei et al. (2015).

In the case of correlated factors (inputs), the contribution of an individual input variable to the variation of the response should be divided into two parts of uncorrelated and correlated contributions. In order to account for correlation, in the present study, factor values are constructed from observed combinations using a k-nearest neighbors (KNN) approach. Observed variance and co-variance are treated as representative of the co-variance structure of the factors. For rutting and dynamic modulus prediction models, VIA is conducted, and the results are presented in this section. Table 5.10 and Figure 5.10 indicate the main and total effects for rutting prediction model. For the developed rutting prediction model, percent air void, percent passing from #200 sieve, VFA

Table 5.10: Variable Importance Analysis Results

Variable	Main Effect	Total Effect
Va%	0.155	0.195
Passing #200	0.024	0.181
VFA	0.155	0.155
VMA	0.128	0.134
Passing #16	0.042	0.119
Passing #100	0.033	0.112
E*	0.091	0.091
NMAS	0.068	0.068
Passing #50	0.03	0.067
Passin 3/4”	0.052	0.052
Passing #8	0.043	0.043
G*	0.03	0.043
Passing 1/2”	0.041	0.041
Passing 3/8”	0.035	0.035
Passing #30	0.03	0.03
Passing #4	0.02	0.025
Binder%	0.021	0.021

and VMA are in turn the most effective variables, while for dynamic modulus prediction model the most effective variables are in turn complex shear modulus and phase angle as presented in Table 5.11 and Figure 5.11.

5.7 Application of the framework

In this section we present two applications of the developed framework in the design and optimization of flexible pavement.

5.7.1 Problem 1: Minimizing accumulated strain

A proper formulation of minimizing accumulated strain in the flexible pavement design can be formulated based on the above framework. The objective function is the output of the trained

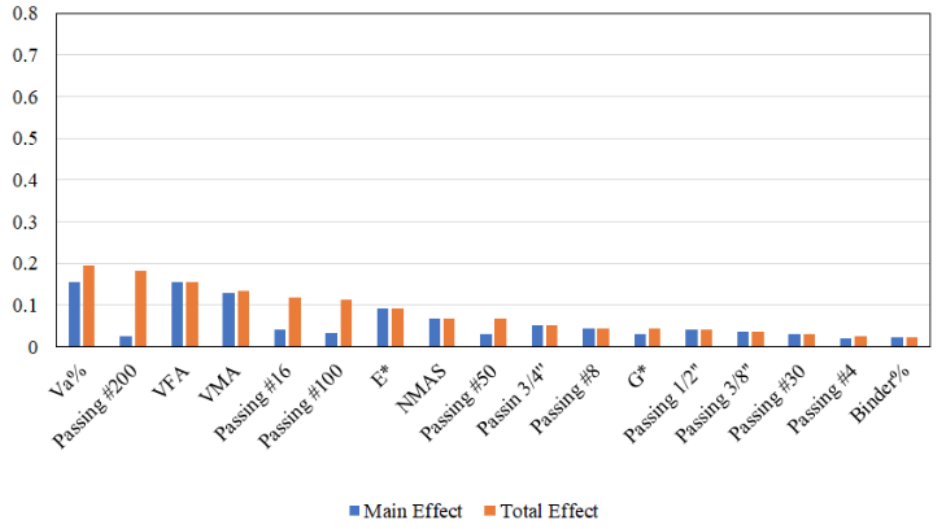

Figure 5.10: Variable importance analysis results

Table 5.11: Variable importance analysis results

Variable	Main Effect	Total Effect
Log —G*—	0.221	0.685
Phase angle	0.156	0.322
VMA	0.057	0.057
Cumulative%Retainedon1/2”	0.052	0.052
Cumulative%Retainedon 3/8”	0.052	0.052
Cumulative%Retained on #4	0.052	0.052
Cum. Retained on #8	0.052	0.052
Cum.Retained on #30	0.052	0.052
Cum.Retained on #50	0.052	0.052
Cum.Retained on #100	0.052	0.052
Passing #200	0.052	0.052
VFA	0.052	0.052
%Va	0.052	0.052
%Vbeff	0.042	0.042

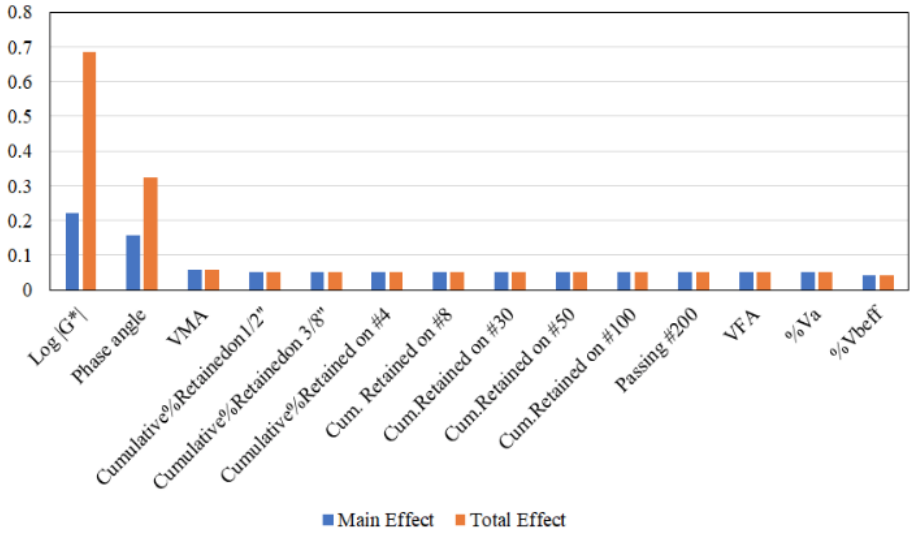

Figure 5.11: Variable importance analysis results.

network based on the laboratory data as explained in the previous section. The next step is adding the appropriate constraints to this problem. These constraints ensure that the desired point remains in the space where the trained dataset were originally located and that the model remains reliable in this domain. It should be noted that the ANN network makes a nonlinear objective function and unlike convex function, their behavior tends to promote multiple local minimizing points. The problem becomes more complicated by the introduction of constraints.

Since the overall optimization problems is not convex, gradient-based algorithms are prone to being trapped in a local optimal point. Many researchers have shows that a proper application of evolutionary-based algorithms for engineering-based problems can result into desirable solutions in a finite amount of computational time Cai et al. (2006); He and Yao (2002); Zitzler and Thiele (1999). In this study, we use a variant of an evolutionary-based search algorithm called Mean Variance Mapping Optimization (MVMO). MVMO Erlich et al. (2010) employs a number of operators to locate the global optimum of a given function. More specifically, MVMO preserves an archive of points during its evolution and extracts information to proceed forward from this archive. MVMO also behaves adaptively so that it can explore the domain initially and exploit the specific domain of interest to better locate the global optimum.

MVMO is originally designed to solve unconstrained problems. In an attempt to optimize truss structures with multiple constraints, Aslani et al. Aslani et al. (2018) provided an adaptive penalty function which can transfer a given constrained problem into an unconstrained one. This strategy has been shown to be robust, especially in the context of the problems with many constraints involved. In the problem of minimizing accumulated strain, the constrains limiting the space to an enclosing ellipsoid enter the problem through constrains and then are added to the objective function through the penalty approach that was described above. Thus, the problem is formulated as following:

$$\begin{aligned} \text{minimize} \quad & \epsilon = \epsilon_{ANN}(\mathbf{x}) \\ \text{with respect to} \quad & \mathbf{x} = (x_1, ..., x_{17}) \\ \text{subject to} \quad & (\mathbf{x} - \mathbf{v})^T \mathbf{A} (\mathbf{x} - \mathbf{v}) \leq \mathbf{1}, \\ & (\mathbf{x}_{pca} - \mathbf{v}')^T \mathbf{A}' (\mathbf{x}_{pca} - \mathbf{v}') \leq \mathbf{1}, \end{aligned} \tag{5.20}$$

where ε is the objective function (in this case, it is the accumulated strain) which is a function of material property indicated by x. Finally, the set of constraints are represented by $(x - v)^T A(x - v) \leq 1$. The uncertainty in the process of finding the enclosing can be found as a function of principal semi-axes of the ellipsoid (s_i):

$$(x - v)^T \, A \;\; (x - v) \leq 1 - |\Delta| \;\; max(s_i) \tag{5.21}$$

where Δ is pre-defined threshold in the effective variable space section. As indicated before, the constrained problem is changed to an unconstrained one through application of a penalty function.

MVMO convergence plot is shown in Figure 5.12. The initial point is random, and the algorithm is evolved through the application multiple routines that model exploration initially and later on exploitation to narrow down the domain of interest.

PCNN predictive model used in the objective function along with the constraint results in a minimum of 1772 micro strain. Table 5.12 summarizes the material properties associated with this solution. The obtained aggregate gradation graph in presented in Figure 5.13.

Table 5.12: Optimal design parameters

Variable	Identity	Optimal values from PCNN	Current design specification			
			Control points		Restricted zone	
			Lower	Upper	Lower	Upper
x_1	Binder %	4	-	-	-	-
x_2	G* (Pa)	270190	-	-	-	-
x_3	NMAS	19	-	-	-	-
x_4	Passing 3/4”	92	90	100	-	-
x_5	Passing 1/2”	66	-	90	-	-
x_6	Passing 3/8”	65	-	-	-	-
x_7	Passing #4	58	-	-	-	-
x_8	Passing #8	50	23	49	34.6	34.6
x_9	Passing #16	39	-	-	22.3	28.3
x_{10}	Passing #30	27	-	-	16.7	20.7
x_{11}	Passing #50	9	-	-	13.7	13.7
x_{12}	Passing #100	4	-	-	-	-
x_{13}	Passing #200	3	2	8	-	-
x_{14}	VMA	16	13	-	-	-
x_{15}	VFA	76	65	80	-	-
x_{16}	Va%	4	4	4	-	
x_{17}	E*(Mpa)	713	-	-	-	-

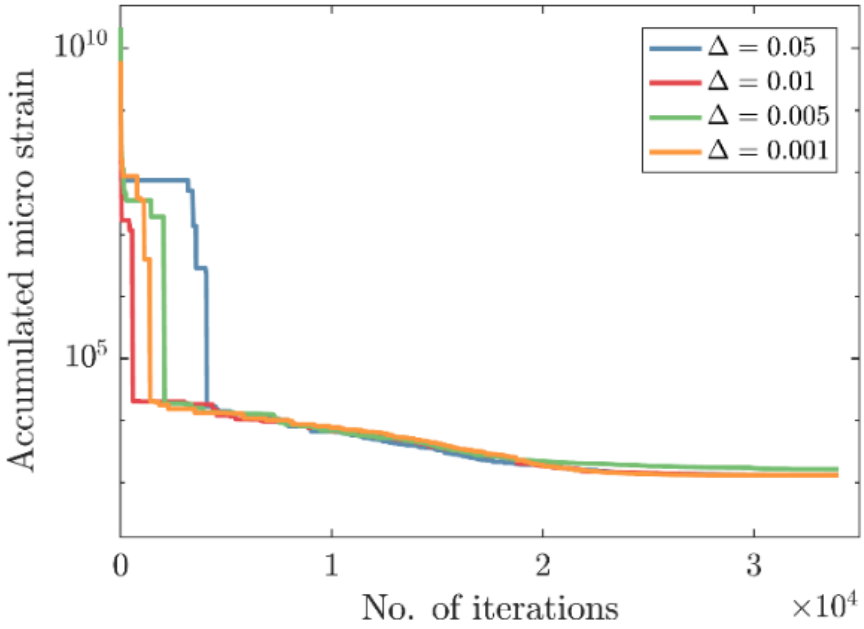

Figure 5.12: Convergence plot.

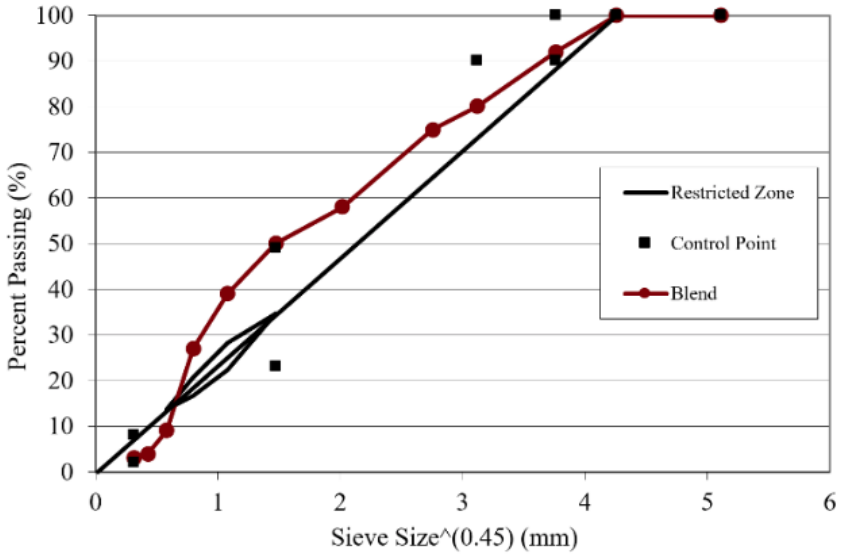

Figure 5.13: Optimal aggregate gradation graph.

5.7.2 Problem 2

The developed predictive model for dynamic modulus is used along with an optimization algorithm to answer the following two central questions:

- what design parameters result in the maximum $|E^*|$?
- what design parameters result in a pre-specified $|E^*|$?

One can see that the first item corresponds to the optimal design problem where

as the second one corresponds to the so-called inverse design. Recalling the PCNN's higher prediction capability, solving optimization problems based on PCNN would be more reliable. The optimal design problem is formulated as follows:

$$\begin{aligned} \text{maximize} \quad & |E^*| = F_{\text{ANN}}(x) \\ \text{with respect to} \quad & x = (x_1,\ \ldots,\ x_{14}) \\ \text{subject to} \quad & (x-v)^T A\ (x-v) \leq 1 \\ & \left(x_{\text{pca}} - v'\right)^T A' \left(x_{\text{pca}} - v'\right) \leq 1 \\ & (X-v)^T A\ (x-v) \leq 1 - |\Delta| \max(s_i) \end{aligned} \tag{5.22}$$

Similar to the rutting predicting problem, the vector $\mathbf{x}$ represents the predictor variables, and $(x-v)^T A(x-v) \leq 1$ is the constrain equation. This optimization problem will be solved by implementing a penalty function that penalizes (decreases, in the case of the maximization problem) the objective value for each constraint regarding to its degree of closeness/violation of the corresponding constraint.

Compared to optimal design problem, the inverse design problem is a minimization problem defined as follows:

$$\begin{aligned} \text{minimize} \quad & \text{error} = ||E^*| - |E_0^*|| \\ \text{with respect to} \quad & x = (x_1,\ \ldots,\ x_{14}) \\ \text{subject to} \quad & (x-v)^T \,\mathrm{A}\ (x-v) \leq 1 \\ & \left(x_{\text{pca}} - v^{'}\right)^T A^{'} \left(x_{\text{pca}} - v^{'}\right) \leq 1, \end{aligned} \tag{5.23}$$

where $|E_0^*|$ is the desired dynamic modulus. A similar penalization method is used to address the constraint in this case as well. As mentioned in previous section, constrained MVMO algorithm is implemented to solve the problem.

Figure 5.14 (left) shows the convergence plot for the optimal design problem solved using the constrained MVMO algorithm. The algorithm initiates with a random initial point (heavily penalized as it can be seen from the graph) and the objective function increases with every iteration. It should be noted that $\Delta = 0.05$ is used as the threshold in Figure 5.12. Solving the maximization problem resulted in $|E^* \max| = 53,703\ MPa$. The obtained optimal design parameters are presented in the first column of Table 5.13. The maximization problem is solved one more time with an additional constraint of $G^* \sin(\delta) \leq 5000kPa$. to find the maximum amount of dynamic modulus one could design for without low temperature failure in the asphalt binder. Solving this problem resulted in $|E \max| = 36,307\ MPa$. Corresponding design parameters are presented in the second column of Table 5.13 as the optimal design 2. Figure 5.14 (right) shows the convergence of the algorithm for the inverse design problem which was started randomly from three different initial points. The algorithm is terminated when the error reaches around 10^{-9}. A pre-specified $|E_0^*|$ of 20,417 MPa. is considered and the inverse problem of finding the corresponding design parameters is solved. Due to the non-linearity of the function, the problem dose not have a unique solution. Three of the possible solutions are presented as designs 1 to 3 in Table 5.13. The five sets of design parameters are compared with current design specification in Table 5.13. The percentage of aggregate passing by sieve size located in the allowable range of the gradation specification. Gradation charts are presented in Figure 5.15. The obtained percentage for air void is 4% which is the target value of the design specification. The obtained values for VMA are slightly less than

Table 5.13: Design parameters associated with maximum

Identity	Optimal Design 1	Optimal Design 2	Design 1	Design 2	Design 3	Design Specification			
						Control Points		Restricted Zone	
						Lower	Upper	Lower	Upper
%Passing from 3/4"	100	100	100	100	100	-	100	-	-
%Passing from 1/2"	93.38	94.03	92.25	91.88	91.80	90	100	-	-
%Passing from 3/8"	81.74	81.72	79.57	79.92	80.70	-	90	-	-
%Passing from #4	53.00	53.90	55.36	55.23	54.39	-	-	-	-
%Passing from #8	39.56	40.51	41.37	41.08	40.92	28	58	39.1	39.1
%Passing from #30	20.75	20.68	21.02	20.87	20.83	-	-	19.1	23.1
%Passing from #50	11.66	11.60	12.08	11.81	12.02	-	-	15.5	15.5
%Passing from #100	6.22	6.21	6.52	6.38	6.40	-	-	-	-
%Passing from #200	4.10	3.85	4.38	4.58	4.56	2	10	-	-
G* (Mpa)	103.13	7.81	133.51	30.20	11.82	-	-	-	-
Phase angle (degree)	35.71	39.60	47.69	47.27	44.77	2	8	-	-
Vbeff%	4.11	4.18	4.02	4.06	4.05	-	-	-	-
VMA	13.47	13.56	13.41	13.45	13.44	-	-	-	-
VFA	70.29	70.50	70.11	70.24	70.24	-	-	-	-
Va%	4.00	4.00	3.99	4.00	4.01	4		-	-

14% for nominal maximum aggregate size (NMAS) of 12.5 mm. The reason is that VMA values of the nine mixtures used to train the PCNN are slightly less that 14% (see Table 5.5). As discussed in previous section, the acceptable range for VFA depends on the amount of traffic. The obtained VFAs for all of the five sets of design are satisfied for all of the traffic categories.

It should be mentioned that these are examples of the framework's application in pavement design indicating how to apply appropriate constraints to the modeling problem. Other example can be finding optimal design parameters when a specific source of aggregate (or specific aggregate size) is missing. Multi-objective design problem (minimizing rut depth while maximizing pavement fatigue life) can also be solved with this framework.

5.8 Conclusions and recommendations

The purpose of this study was to develop a predictive framework for pavement performance that could be reproducible and easy to use for every data base. The proposed framework in a data preprocessing step, evaluates and qualifies the input variables to use them in further analysis. It identifies cross correlated input variables using correlation analysis and substitutes them by orthogonal pseudo-variables (PCs) using a multivariate technique called PCA. This transformation

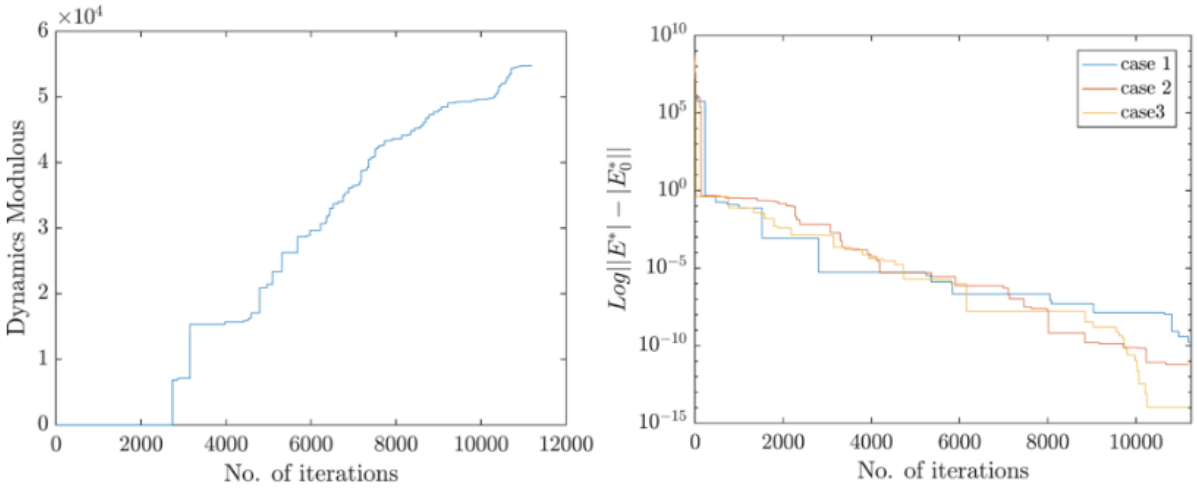

Figure 5.14: Convergence plot.

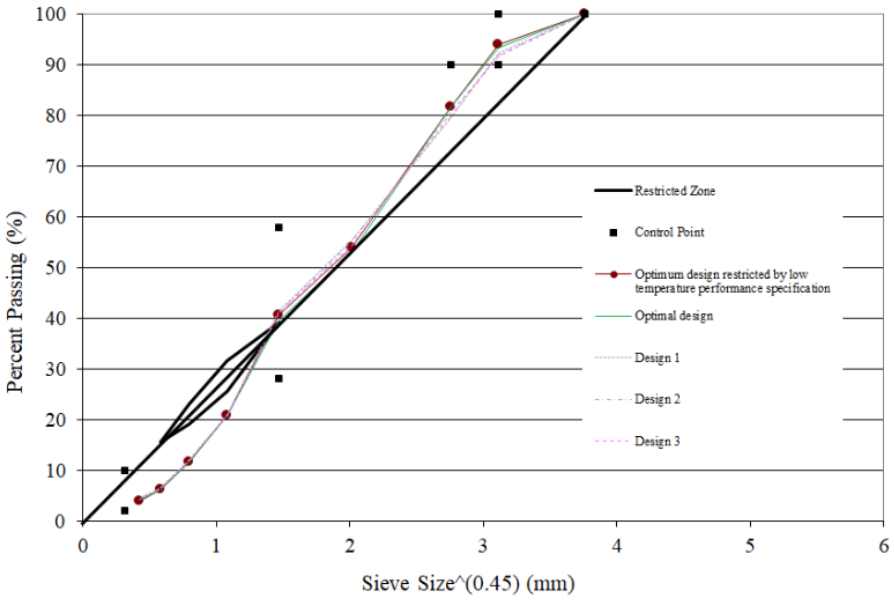

Figure 5.15: Aggregate gradation graphs.

not only eliminates the existing pairwise correlation between the original inputs but also it reduces the dimensionality of the data set and increases the accuracy of the prediction. The framework uses pseudo-inputs (PCs) and develops two predictive models using multivariate regression and ANN (models are called PCR and PCNN respectively.)

According to the size of the available data set, the framework implements a cross-validation technique to prevent developing biased or overfitted models. In case of limited amount of data points (e.g. rutting data set) the network performs k-fold cross validation technique to assure that the developed model is stable against an unseen data set. Empirical predictive models can lead to inaccurate results under extrapolation. A simple method is implemented to define the effective variable space in which both predictive models can be used. The defined hyperspace is added as a constraint to the modeling problem.

To illustrate the authority of the proposed framework, two separate performance prediction problems are defined and solved. In the first problem, the rutting behavior of asphalt mixture is predicted as a function of asphalt binder, aggregate, and mixture properties using experimental data of the flow number test. In the second problem, the dynamic modulus of asphalt mixture is predicted using asphalt binder, aggregate, and mixture volumetric properties using experimental data of the dynamic modulus test. In both of the problems, the developed models, PCR and PCNN, indicated satisfactory performance in terms of modeling the amount of permanent deformation and dynamic modulus value, with PCNN being significantly better in fitting the test data than the conventional performance predictive models.

To indicate one application of the proposed framework in pavement performance prediction, the problem of finding the optimal design parameters is solved using mean-variance mapping optimization algorithm for both of the performance prediction problems over their effective variable spaces. The value of 1772 micro-strain is obtained as the minimum accumulated strain. For the problem of predicting dynamic modulus, the inverse design problem of finding the design parameters corresponding to any pre-specified value of dynamic modulus is also solved over its effective variable space. In all the optimization problems the design parameters corresponding to optimum

or inverse designs are obtained for PCNN. The obtained optimal design parameters satisfy the current asphalt pavement design specifications and could be used as an appropriate starting point in the design procedure.

It is also worth pointing out that selection of study materials was based on the availability of asphalt mixtures for laboratory testing and like every empirical model the obtained results are based on the available empirical database. Eventually, for creating a more reliable predictive model, a larger data base is required. However, what makes machine learning-based models special is that the model will be retrained and modified when a new data set is fed into the framework. Thus, unlike other empirical predictive models there is no need to calibrate the model for each specific location based on its climate condition and available materials. Moreover, the proposed framework can be used to predict any performance-related characteristics of asphalt mixture including rutting, fatigue, low temperature behavior, etc. In other words, having several performance predictive models for asphalt mixture is no longer necessary. Besides, due to the high accuracy of the developed models in predicting pavement performance characteristics, the framework can be implemented to improve level 2 and level 3 inputs in MEPDG design procedure. The proposed framework can make a stand-alone software for predicting pavement performance which is highly beneficial for asphalt agencies when a large number of performance data is available.

Future research would address the capability of the model in handling larger data sets. Considering the high accuracy of the developed predictive models, these PCA-based approaches are strongly recommended as an appropriate modeling approaches in this application. Moreover, these methodologies appear to be capable of modeling asphalt binder chemical properties as well as finding performance volumetric relationships and such investigations are recommended as future studies.

5.9 References

Al-Khateeb, G., Shenoy, A., Gibson, N., and Harman, T. (2006). A new simplistic model for dynamic modulus predictions of asphalt paving mixtures. *J. Assoc. Asph. Paving Technol.*, 75.

Andrei, D., Witczak, M., and Mirza, M. (1999). Development of a revised predictive model for the dynamic (complex) modulus of asphalt mixtures. *Dev. 2002 Guide Des. New Rehabil. Pavement Struct. NCHRP*.

Apeagyei, A. K. (2011). Rutting as a Function of Dynamic Modulus and Gradation. *Am. Soc. Civ. Eng.*, 23 (9):1302–1310.

Aslani, M., Ghasemi, P., and Gandomi, A. H. (2018). Constrained mean-variance mapping optimization for truss optimization problems. *Struct. Des. Tall Spec. Build.*, 27(6):e1449.

Bari, J. and Witczak, M. (2007). New predictive models for viscosity and complex shear modulus of asphalt binders: For use with mechanistic-empirical pavement design guide. *Transp. Res. Rec. : J. Transp. Res. Board*, 1(2001):9–19.

Bashin, A., Masad, E., Kutay, M. E., Buttlar, W., Kim, Y.-R., Marasteanu, M., Kim, Y. R., Schwartz, C. W., and Carvalho, R. L. (2012). Applications of advanced models to understand behavior and performance of asphalt mixtures. *Transp. Res. E-Circ.*, 1(E-C161).

Birgisson, B., Sholar, G., and Roque, R. (2005). Evaluation of a predicted dynamic modulus for florida mixtures. *Transp. Res. Rec.*, 1929(1):200–207.

Cai, Z., Cai, Z., and Wang, Y. (2006). A Multiobjective Optimization-Based Evolutionary Algorithm for Constrained Optimization. *Evol. Comput. IEEE Trans.*, 10(6):658–675.

Cheng, B. and Titterington, D. M. (1994). Neural networks: A review from a statistical perspective. *Statistical science*, pages 2–30.

Christensen Jr, D., Pellinen, T., and Bonaquist, R. (2003). Hirsch model for estimating the modulus of asphalt concrete. *J. Assoc. Asph. Paving Technol.*, 72.

Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences*. Cengage learning.

Erlich, I., Venayagamoorthy, G. K., and Worawat, N. (2010). A mean-variance optimization algorithm. *IEEE Congr. Evol. Comput.*, pages 1–6.

Fodor, I. K. (2002). A Survey of Dimension Reduction Techniques. Technical report, U.S. Department of Energy, U.S. Department of Energy.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. (2018a). Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: Elimination of correlated inputs and extrapolation in modeling. *Struct. Multidiscip. Optim.*, pages 1–19.

Ghasemi, P., Aslani, M., Rollins, D. K., Williams, R. C., and Schaefer, V. R. (2018b). Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks. *Int. J. Pavement Res. Technol.*

He, J. and Yao, X. (2002). From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. *IEEE Trans. Evol. Comput.*, 6(5):495–511.

Hosseini, A., Faheem, A., Titi, H., and Schwandt, S. (2020). Evaluation of the long-term performance of flexible pavements with respect to production and construction quality control indicators. *Construction and Building Materials*, 230:116998.

Jalali, F., Vargas-Nordcbeck, A., and Nakhaei, M. (2019). Role of preventive treatments in low-volume road maintenance program: Full-scale case study. *Transportation Research Record*, page 0361198119863025.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.*, 30(3):487.

K Rollins Sr, D. (2015). A One-Dimensional PCA Approach for Classifying Imbalanced Data. *J. Comput. Sci. & Systems Biology*, 8(1):5–11.

Kaloush, K. E., Witczak, M., and Sullivan, B. W. (2003). Simple performance test for permanent deformation evaluation of asphalt mixtures. *Sixth Int. RILEM Symp. Perform. Test. Eval. Bitum. Mater.*, pages 498–505.

Kutner, M. H., Nachtsheim, C., and Neter, J. (2004). *Applied linear regression models*. McGraw-Hill/Irwin.

Kvasnak, A., Robinette, C., and Williams, R. (2007). A Statistical Development of a Flow Number Predictive Equation for the Mechanistic-Empirical Pavement Design Guide.CD-ROM . In Transportation Research Board, Washington, D., editor, *Transportation Research Board, Washington, DC.*, volume 3, pages 1–18, Transportation Research Board, Washington, DC. DC.Transportation Research Board, Washington.

Nobakht, M. and Sakhaeifar, M. S. (2018). Dynamic modulus and phase angle prediction of laboratory aged asphalt mixtures. *Constr. Build. Mater.*, 190:740–751.

Notani, M. A., Moghadas Nejad, F., Khodaii, A., and Hajikarimi, P. (2019). Evaluating fatigue resistance of toner-modified asphalt binders using the linear amplitude sweep test. *Road Materials and Pavement Design*, 20(8):1927–1940.

Pellinen, T. and Witczak, M. (2002). Use of Stiffness of Hot-Mix Asphalt as a Simple Performance Test. *Transp. Res. Rec. : J. Transp. Res. Board*, 1789(02):80–90.

Peng, C., Feng, J., Feiting, S., Changjun, Z., and Decheng, F. (2019). Modified two-phase micromechanical model and generalized self-consistent model for predicting dynamic modulus of asphalt concrete. *Constr. Build. Mater.*, 201:33–41.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer.

Rodezno, M., Kaloush, K., and Corrigan, M. (2010). Development of a Flow Number Predictive Model. *Transp. Res. Rec. : J. Transp. Res. Board*, 2181(2181):79–87.

Rollins, D. K., Zhai, D., Joe, A. L., Guidarelli, J. W., Murarka, A., and Gonzalez, R. (2006). A novel data mining method to identify assay-specific signatures in functional genomic studies. *BMC bioinformatics*, 7:377.

Sakhaeifar, M. S., Richard Kim, Y., and Garcia Montano, B. E. (2017). Individual temperature based models for nondestructive evaluation of complex moduli in asphalt concrete. *Constr. Build. Mater.*, 137:117–127.

Shu, X. and Huang, B. (2008). Dynamic modulus prediction of hma mixtures based on the viscoelastic micromechanical model. *J. Mater. Civ. Eng.*, 20(8):530–538.

Sobol', I. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(1):112–118.

Sun, P. and Freund, R. M. (2004). Computation of minimum-volume covering ellipsoids. *Operations Research*, 52(5):690–706.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association.

Timm, D., West, R., Priest, A., Powell, B., Selvaraj, I., Zhang, J., and Brown, R. (2006). Phase ii ncat test track results. *NCAT Report*, 6(05).

Todd, M. J. and Yıldırım, E. A. (2007). On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids. *Discret. Appl. Math.*, 155(13):1731–1744.

Wei, P., Lu, Z., and Song, J. (2015). Variable importance analysis: a comprehensive review. *Reliability Engineering & System Safety*, 142:399–432.

Witzcak, M. W. (2002). *Simple performance test for superpave mix design*, volume 465. Transportation Research Board.

Ziari, H., Amini, A., Goli, A., and Mirzaiyan, D. (2018). Predicting rutting performance of carbon nano tube (cnt) asphalt binders using regression models and neural networks. *Construction and Building Materials*, 160:415–426.

Zitzler, E. and Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.*, 3(4):257–271.

CHAPTER 6. AN INVERSE APPROACH FOR EVALUATING GANTRY CRANE-WAY PAVEMENT PERFORMANCE

A paper Submitted to the Journal of Transportation Geotechnics

6.1 Abstract

Gantry-crane pavements and foundations are significant assets within intermodal facilities. Because they are subjected to highly-variable loads and are critical to safe operations, traffic interruptions and costs associated with maintaining and rehabilitating distressed or failed pavements in associated areas are of particular importance. Design of gantry crane-way pavements and investigation of the contribution to overall performance of each of the pavement layers are challenging tasks for design engineers, and failure to seriously address this contribution can result in premature pavement failure or high maintenance requirements. The purpose of this study was to evaluate structural behavior and improve design procedures for gantry crane way pavement used at intermodal facilities by assessing interactions among pavements, subgrades, and operational loading conditions. The performance of the gantry crane pavements and foundations was assessed using a finite-element model, while pavement structural response to a crane load was measured using strain gages installed in the field. An inverse design approach was implemented to verify consistent behavior of the developed finite-element model with repect to the field measurements. Because of the significant cost of FEM simulation, the gradient-based solver used in this study was based on an interior point algorithm, following which the verified model was used to predict the critical responses of PCC layer, base course, and subgrade soil. These parameters were then used to conduct a pavement fatigue damage analysis, parametric analyses of material strength and slab geometry were carried out based on Model Code, and resulting fatigue-life recommendations for improved designs were made.

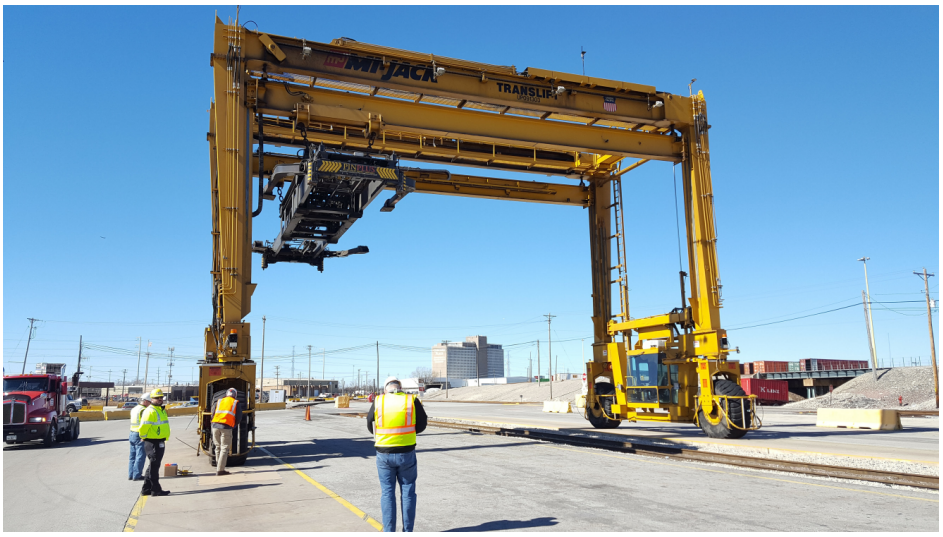

Figure 6.1: Gantry crane and crane way pavement.

6.2 Introduction

At intermodal facilities, the pavement, including paved entrances/exits, loading areas, truck aisles, and container parking areas, usually represents the largest asset in the facility. The United States economy significantly depends on successful and efficient operation of intermodal facilities in transporting goods across the nation, and operational safety, reduced operational variability , and satisfactory customer interactions all depend on satisfactorily-performing pavements under a variety of loadings and environmental conditions. Gantry-crane pavements and their foundations, depicted in Figure 6.1, comprise a key area subject to wide variation in loading and critical to the safe operation of intermodal facilities. When the cross-country train arrives at a halting point at an intermodal facility, heavy containers will be typically be unloaded using a Rubber Tire Gantry Crane (RTGC), and severe distresses (and failure) can occur to the crane-way pavement when it is subjected to numerous repeated heavy loads in a daily operation. Since traffic disruptions and costs associated with maintaining and rehabilitating distressed or failed pavements in these areas are of particular importance in achieving safe and efficient operation of gantry crane areas and associated intermodal facilities, crane-way pavement structural performance requires a unique assessment for development of appropriate design and maintenance strategies.

Concrete pavement deterioration is a function of several factors, including slab thickness, material properties (concrete and foundation beneath the pavement), boundary conditions, interactions between concrete slab and base course, subgrade soil characterization, environmental conditions, and magnitude and position of applied load Maitra et al. (2010). In the case of heavy-duty pavements, i.e., pavements designed to carry heavy loads, subgrade soil characteristics play an even more significant role in pavement performance Kermani et al. (2018); Shoukry et al. (2007); Uddin et al. (1995); Sadeghi and Hesami (2018). One of the most common cause of failure of such pavements is weak soil support (low bearing capacity) beneath the pavement, suggesting that failure to rigorously test and accurately characterize the native foundation soils can lead to premature pavement failure and high maintenance requirements Knapton (1989); Huang (1993); Miller et al. (2014), major issues when examining crane-way pavement performance at intermodal facilities.

Finite-element analysis has serves as a powerful tool for capturing pavement responses when analyzing rigid-pavement structure Huang (1993); Tabatabaie and Barenberg (1978); Shoukry et al. (2007); Uddin et al. (1995); Sadeghi and Hesami (2018). The methodology is not new to the pavement design and research community and has been used over the past several decades. Traditional pavement-subgrade analysis, based on static load and multilayer linear elastic formulation of infinite dimension in the horizontal plane and a semi-infinite subgrade, does not account for dynamic behavior or pavement discontinuities. The finite-element method, however, does support dynamic analysis of pavements and consideration of either finite or infinite dimensions of the physical pavement structure. Several finite-element programs have been developed exclusively for pavement analysis, e.g., ILLIPAVE and ILLISLAB for flexible and rigid pavements, respectively, but these packages are capable of performing static analysis only. The finite-element package ABAQUS is available for comprehensive structural pavement response analysis in both static and dynamic procedures. ABAQUS also has a variety of material models, i.e., linear elastic, nonlinear elastic, viscoelastic etc. Since surface-to-surface contact problems can also be defined and solved using this package, in the present study ABAQUS was implemented for simulating a crane-way pavement section under crane load. However, while finite-element analysis is an established method providing

sufficient efficiency and precision, it can simulate the actual performance of the pavement only if the input information is sufficiently precise. Limited time and resources may make it impossible to measure every single property in the field, and in such cases the problem under study may have both known and unknown properties (e.g. due to the time limitations the only field measurement is strain values in concrete pavements).

There are two major approaches to tackling the problem of finding unknown properties of a system from some experimental measurements. In the first approach (also called the forward problem), one can run the FEM simulation of the system using some ranges of given unknowns, then build a network using the resulting batch of simulations. Given some error threshold. Once the network is built one can use it inside an optimization solver to directly determine the unknown properties. While this approach has been used in a number of studies Ghasemi et al. (2019a), it may not be applicable if computational resources are limited, because running FEM simulation of even a medium-sized model is computationally cumbersome. Another approach would be using the inverse optimization approach (the forward model problem) in which the value of an unknown parameter under investigation is guessed and the solution obtained is validated. Once the forward model is solved for an assumed value of input, the next step would be to compute the error between simulated and experimental measurements. If the error is found to lie within the limit, the initial guess is taken as the solution to the inverse problem; otherwise, the process is repeated until the error condition is satisfied. It should be noted that, once the numerical simulation (in this case the FEM) has been verified, one can use this same model to study properties or measurements not available from experiments. While a step forward would be to use the current model to optimize the design of the system of interest, in this study we focus only on the feasibility of the inverse approach to finding a high-precision numerical simulation. The purpose of this study was to evaluate structural behavior and improve design procedures used for determining gantry-crane way pavement used at intermodal facilities. To achieve this purpose, existing facilities of the intermodal facilities were investigated, instrumented, and assessed to seek understanding of interactions between pavements, subgrade, operational loading conditions, and performance of such pavements. Finite-element analysis was

used to assess performance related to design and maintenance procedures . An inverse design methodology was implemented to verify the finite-element model against the field measurements, and pavement critical responses and structural life predictors were obtained from the verified model, leading to recommendations for modifying crane-way pavement design procedures . The rest of this paper is organized as follows: The field test setup is presented in the next section, followed by discussion of the modeling approach implemented and inverse-design strategy. Results, discussion, and conclusions are presented in the remaining sections.

6.3 Field test setup

Strain measurements were collected from a gantry crane way at intermodal facilities located in Chicago, Illinois. The pavement was a 21 inch (53.34 cm) thick Portland cement concrete (PCC) slab poured over a 6 inch (15.24 cm) thick crushed limestone base course, with the PCC layer saw-cut every 10 ft. (3.048 m). A total number of 18 vibrating-wire strain gages had been installed at top, middle and bottom of the middle, and the edge of the PCC slab before pouring the concrete, as shown in Figure 6.2. A schematic of the strain gages installation plan is shown in Figure 6.3. The Model 4200 (standard)strain gages used in this study,designed for direct embedment in concrete, had a 153 mm gage lengths and 1 $\mu\epsilon$ sensitivity, are commonly used for strain measurements in foundations, piles, bridges, dams, tunnel linings, etc. They were oriented horizontally to measure longitudinal strain developed as the crane slowly passed over the slab, applying a load for a period of 40 seconds before moving to the next loading situation. Consistent with the strain-gage orientation and the defined coordinate system, the strain in the x-direction ,ϵ_{xx}, was measured as the response of the pavement to the applied monotonic load. The footprint of the RTGC tire, also recorded during the site visit, was used to estimate the net contact area of the tire. Due to the costs associated to traffic interruptions, the field test setup must have been done in less than two days. Moreover, taking cores and bore holes from the field was not allowed to avoid traffic interruptions. Therefore, direct measurement/back-calculation of the materials properties using in-situ/laboratory tests was not possible in this study.

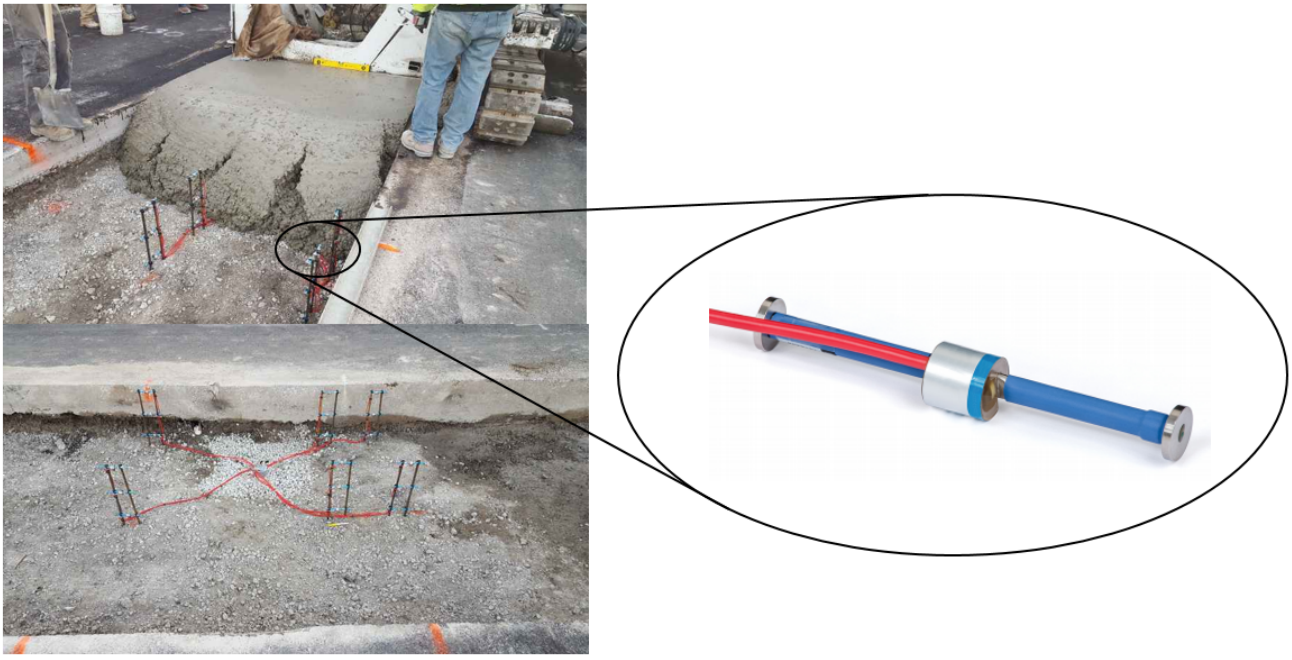

Figure 6.2: Vibrating wire strain gages

6.4 Modeling

A three-dimensional PCC pavement with properties similar to those of the experimental pavement was modeled in this study. The modeled pavement section consisted of five full-length slabs with untied asphalt-concrete shoulders. The concrete slabs along with a base course were placed on a saturated-clay subgrade soil. A schematic of the developed model is presented by Figure 6.4. The slabs are tied together every 30 ft (9.144 m) using dowel bars. Dowel bars are assumed to be embedded into the concrete slabs. The boundary conditions were considered fixed for the bottom of the subgrade layer and the joints at the edges of the soil and the base layer. The lateral sides of the concrete slabs were assumed to be free to simulate field conditions under which concrete shrinkage and expansion can happen. A schematic of the applied boundary conditions is presented in Figure 6.5.

Since the dowel bars are placed every three slabs in the field, in developing the model, aggregate interlock was assumed to be the main source of load transfer between the adjacent slabs, and a

2 Strain Instrumentation Plan.png

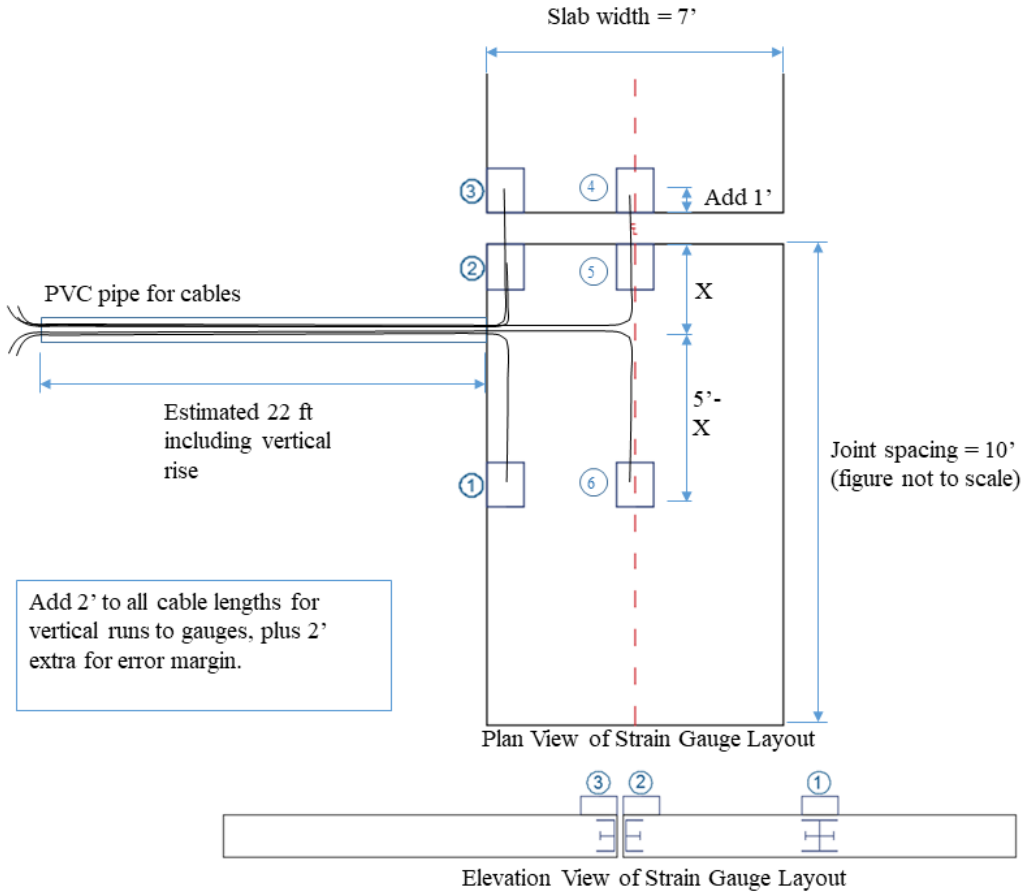

Figure 6.3: Strain gages installation.

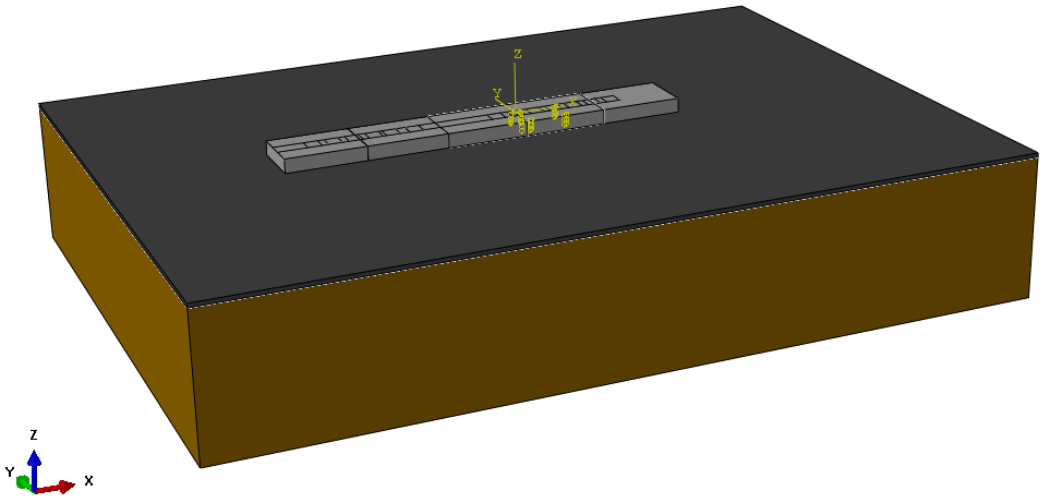

Figure 6.4: A schematic of the developed model.

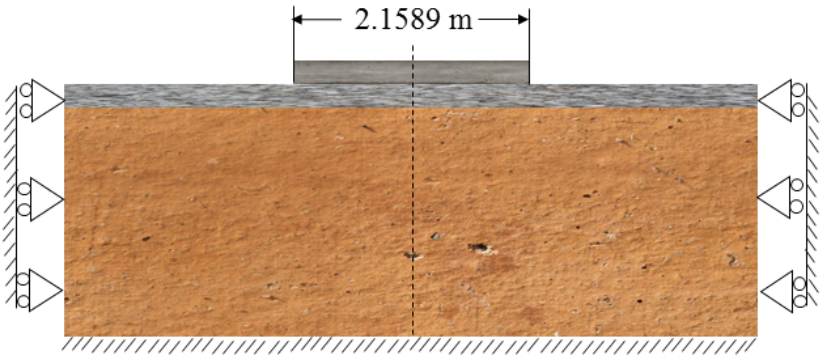

Figure 6.5: Applied boundary conditions in finite element.

sliding interface with a coefficient of friction value of 1.5 and a possibility of slab-base separation was provided between the slabs and the base layer. As mentioned previously, during the site visit a tire footprint was taken, so to enhance image quality and extract useful information image-processing analysis was implemented in MATLAB. After enhancing the contrast, removing noise, and providing some filtering, the net contact area of the tire footprint was calculated, and the original tire footprint and final image-processing results are presented in Figure 6.6. According to information provided about the crane model (mj MI-Jack 1200) and the obtained contact area a 689,851 Pa (690 MPa) pressure is applied on the pavement by each wheel of the crane. To achieve compatibility with the field setup, during the analysis the same amount of pressure was applied by the two wheels for a duration of 40 seconds. Linear elastic materials were used for concrete slabs as well as the base course, and to examine the inelastic behavior of the concrete, a Concrete Damage Plasticity (CDP) model was implemented. That model assumed main concrete failure mechanisms were tensile cracking and compressive crushing. Equation 6.1 known as the Kent and Park model, was used to describe concrete uniaxial compression behavior in the absence of laboratory test results.

$$\sigma_c = f'_{c0}[2(\frac{\epsilon_c}{\epsilon'_c}) - (\frac{\epsilon_c}{\epsilon'_c})^2] \tag{6.1}$$

where σ_c is compressive stress, ϵ_c is compressive strain, f'_{c0} is compressive strength, and ϵ'_c is strain at compressive strength. The concrete compression damage behavior is defined using the following equations:

$$\sigma_c = (1 - d_c)E_0(\epsilon_c - \widetilde{\epsilon_c^{pl}}) \tag{6.2}$$

$$\widetilde{\epsilon_c^{pl}} = \widetilde{\epsilon_c^{in}} d - \frac{1}{(1 - d_c)}\sigma_c / E_0 \tag{6.3}$$

$$d_c = 1 - \frac{\sigma_c}{f'_c} \tag{6.4}$$

where $\widetilde{\epsilon_c^{pl}}$ is the plastic compressive strain, $\widetilde{\epsilon_c^{in}}$ is the inelastic compressive strain,E_0 is the initial elastic modulus, and d_c is the compression damage coefficient. Concrete tensile strength is assumed

to be 10% of its compressive strength. The tensile behavior of concrete is defined using the following equations.

$$\sigma_t = (1 - d_t)E_0(\epsilon_t - \widetilde{\epsilon_t^{pl}}) \tag{6.5}$$

$$\widetilde{\epsilon_t^{pl}} = \widetilde{\epsilon_t^{ck}} - \frac{1}{(1 - d_t)}\sigma_t / E_0 \tag{6.6}$$

where σ_t is tensile stress, d_t is tensile damage coefficient, ϵ_t is tensile strain, $\widetilde{\epsilon_t^{pl}}$ is plastic tensile strain, and $\widetilde{\epsilon_t^{ck}}$ is cracking strain. In this way the cycling behavior of concrete were defined Kent and Park (1971); Grassl and Jirásek (2006) and the elastic and plastic properties (with kinematic hardening) of the steel dowel bars were defined. According to field observations, the ground water table level was as high as the base course height, so the subgrade soil was 100% saturated, and the Mohr-Coulomb model was implemented to describe the elastic-perfectly-plastic behavior of saturated clay. A geostatic procedure was used as the first step of the analysis to ensure that gravity load and the initial geostatic stress were in equilibrium and produced zero deformation. An 8-node linear brick with reduced integration and hourglass control (C3D8R) was used for concrete slabs and base course, and for the subgrade soil a 20-node brick with pore pressure, quadratic displacement, linear pore pressure, and reduced integration (C3D20RP) was used. Two-node linear beam elements (B31) were used for the dowel bars. A convergence study was also performed to ensure that the finite-element analysis results were not affected by changing the mesh size. The finite-element method enables evaluation of the state of stress and strain in a continuum by transforming the continuum into an assemblage of finite elements interconnected at their mutual nodes. Dynamic finite-element analysis involves solving the differential equation of motion:

$$M\ddot{U} + C\dot{U} + KU = F \tag{6.7}$$

where M is mass matrix, C is damping matrix, K is stiffness matrix, $\ddot{U}, \dot{U}, U$ are vector of acceleration, velocity, and displacement, and F is vector of nodal forces.

Figure 6.6: Tire foot print taken in the field.

As previously discussed, in this study an ABAQUS finite-element package was utilized to solve the aforementioned differential equation, and 500 CPU-hours on the High Performance Computer (HPC) at Iowa State University was used to complete each simulation.

6.4.1 An inverse approach for material property estimation

As previously discussed, during field measurements the amount of strain in the x-direction, ϵ_{xx}, at top, middle, and bottom of the slabs were measured as pavement responses to the applied load, and the values of material properties, including layer modulus or Mohr-Coulomb parameters, were not known a *priori*. [1]. To verify the developed model against field measurements one should find the set of the material properties for use in the numerical model that results in the closest ϵ_{xx} value to the field measurements. Performing this verification is the same as solving the problem inversely that would require solving an optimization problem. The main objective of using the inverse algorithm was to obtain the solution in the least possible time, thereby imposing minimal computational penalties. The inverse problem is usually associated with estimation of unknown parameters based on the measurements from sensors located at known accessible locations in the domain. The current optimization problem has a number of functions calculated based on the FEM simulation, and such simulation-based optimization often involves derivatives that are unavailable and, in this particular case, prohibitively expensive in terms of calculation or numerical approxi-

[1]Note that one might have access to this information given their available facility

mation. In an inverse problem we typically require two solvers: a forward solver (in this case the FEM model) used to evaluate the objective function, and an inverse solver that iteratively locates the optimum point. The forward solver (FEM simulation) was described in the previous section, and the outputs from the forward problem (in this case the strain values) were used as inputs to the inverse solver. The inverse solver uses an iterative optimization solver to locate the optimum point where the following error function is minimized:

$$F_{error}(\mathbf{M}) = \sum_{i=1}^{N} \left(\epsilon_{i,FEM} - \epsilon_{i,EXP}\right)^2 \tag{6.8}$$

where N is the number of points to be considered for the comparison. Since the strain obtained from the experiments is constant, the error function is dependent only on the strain rates obtained from the FEM simulation. The vector $\mathbf{M}$ contains the set of material properties that remain constant when the forward problem (FEM simulation) is called.

There are a number of naturally-occurring constraints in this case, including field measurement errors. After initial trial-and-error finite-element analysis (to calibrate the material properties) it was found that not all of the strain gages achieved the same measurement accuracy, so the six most accurate gages were selected for use in the objective function while the others would be used as constraints for the inverse problem. These constraints were associated with strain values not appearing in the error function that should be constrained to a prescribed specific values. These strains were obtained directly from the forward solver and entered into the inverse problem as a nonlinear constraint. If $\mathcal{E}_c = \mathcal{E}_c(\mathbf{M})$ (with size N_c),) is taken as the vector of constraints not contained in the error function, $\mathcal{E}_c < \mathcal{E}_0$ is the desired set of constraints for the inverse problem. As can be seen, unlike regular bounds for variables that enter the optimization problem linearly, the constrains on the strains are nonlinear and must be treated appropriately.

If $\mathcal{E}_c(\mathbf{M}) - \mathcal{E}_0 = \mathbf{c(M)} \leq 0$, are the constraints expressed in functional form we can define a Lagrangian function for the general problem in the following form:

$$L(\mathbf{M}, \lambda) = F_{error}(\mathbf{M}) - \sum_i \lambda_i c_i(\mathbf{M}), \tag{6.9}$$

where λ_i is the Lagrangian multiplier for constraint i. Our goal is to find M^* and λ^* such that the following conditions (also known as KKT conditions) are simultaneously satisfied

$$\nabla_M L(\mathbf{M}^*, \lambda^*) = 0, \quad c(\mathbf{M}^*) \leq 0, \quad \lambda_i^* c_i(\mathbf{M}^*) = 0 \quad for\ all \quad i \tag{6.10}$$

where the last complementary condition implies that a constraint is either active or $\lambda_i = 0$ or both. In this way, values of λ_i can be interpreted as parameters for study of sensitivity analysis of the constraints of the problem. We also note that the first condition can be rewritten in the following form:

$$\nabla_M L(M^*, \lambda^*) = \nabla F_{error}(\mathbf{M}^*) - \sum_i \lambda_i^* \nabla c_i(\mathbf{M}^*) = 0 \tag{6.11}$$

in which we are looking for a stationary point when first-order necessary conditions are met. As can obviously be seen, the above function involves calculation of the gradient function for the error function and the constraints, but since we had no direct access to those functions (The Abaqus solver is basically a black box here!), we used an approach based on a gradient-based optimization solver and instead attempted to approximate the derivatives of our error and constraint functions. We also note that the application of evolutionary heuristic or meta-heuristic algorithms are not appropriate because they require many function evaluations to converge (if we can precisely define convergence for them). Also, since we seeded a local convergence rate that was near quadratic, a fast termination of the algorithm was expected.

In this work we employed a trust-region-based algorithm used in the context of sequential quadratic programming (or SQP) because of the presence of the constraints, and we were trying to minimize sub-problems defined by fully-linear or fully-quadratic models. These problems are typically formed from linear/polynomial interpolations, and our modified function written in the form of an SQP was:

$$\begin{aligned} \text{minimize} \quad & f_k + \nabla f_k^T p + \frac{1}{2} p^T \nabla^2_{xx} L_k p \\ \text{with respect to} \quad & p \\ \text{subject to} \quad & \nabla c_i(x_k)^T p + c_i(x_k) > 0 \quad for\ all \quad i \end{aligned} \tag{6.12}$$

where calculation of the gradient of the function is based on quadratic approximation in the vicinity of the current points (x_k) and the second derivative is approximated based on the BFGS algorithm Nocedal and Wright (2006). The major interest in this algorithm stems from the fact that it is able to converge quickly given a proper initial condition, and also that, in all iteratipons of the algorithm, the inequality constraint was strictly satisfied. This was an important advantage. especially because the cost of performing the FEM simulation is excessively high. There are many details of the above algorithm that have been omitted in this discussion, and the reader is referred to the Nocedal and Wright (2006) for further details.

Our experimentation with the current problem has shown that the this algorithm converges to a local minimum in a number of epochs (O(10)) thanks to locally-descending Newton steps at each iteration. Figure 6.7 shows convergence history for the normalized gradient and objective functions on the left and right axis, respectively, demonstrating the rapid convergence of the aforementioned methodology. As can be seen, the simulations were terminated after approximately 12 epochs.

A schematic summery of the implemented inverse approach is presented in Figure 6.8. The inverse solver starts from a guess (in this case a set of prescribed values obtained from the literature Shoukry et al. (2007)) aand the forward FEM problem is then solved. In the next step, a gradient-based optimization algorithm determines a new set of properties, then the forward problem is solved once again and the error function is reassessed. This procedure continues until the target function has been minimized. The inverse solver is coded in MATLAB and combined with ABAQUS software to solve the forward problem.

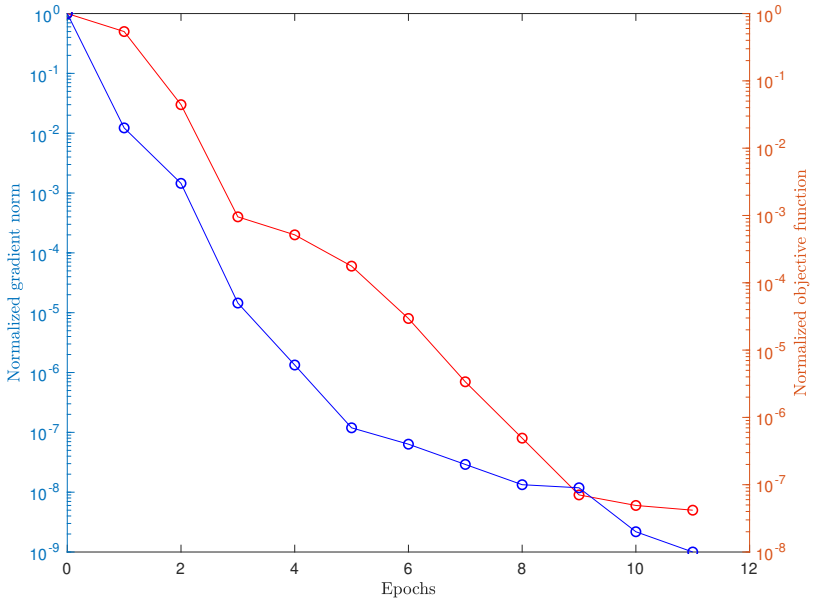

Figure 6.7: Inverse solver convergence history

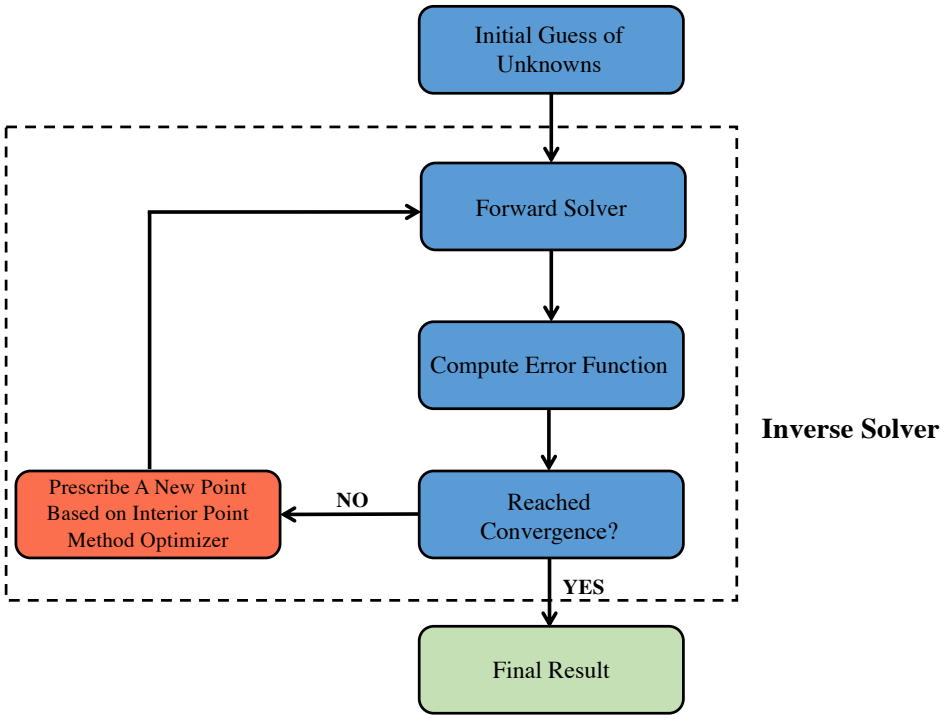

Figure 6.8: Flowchart of the forward and inverse solver

Table 6.1: Material properties obtained from the inverse approach

	Material			
Property	Concrete	Crushed limeston base	Clay subgrade	Steel dowel bars
Density (kg/m^3)	2350	2240	1800	7850
Elastic Modulus (Pa)	3.445×10^{10}	1.76×10^{8}	3×10^{7}	2.10×10^{11}
Poisson's ratio	0.18	0.4	0.48	0.3
Permeability (m/s)	-	-	1×10^{-7}	-
Void ratio	-	-	0.889	-
friction angle ($degree$)	-	-	0	-

6.5 Results and Discussions

Results from both the finite-element analysis and the field measurements are presented in this section. The inverse design approach resulted in a set of material properties that led to the closest ϵ_{xx} (or minimum difference between the measured and the predicted response). The obtained material properties for concrete, crushed limestone base course, subgrade soil, and steel dowel bars are presented in Table 6.1. After verifying the finite-element model against the field-produced data using obtained material properties, six of the best-performing strain gages were selected and their corresponding measured and estimated strain values are presented in Figure 6.9 that indicates the amount of strain captured by the strain gages as the crane passes over the pavement. It should be mentioned that since these six strain gages are those used in the inverse design procedure, they indicate the strain values closest to the field measurements, while the measurements obtained from the other gages were used as constraint (to stay within a prescribed range) in the inverse solver.

Now that it has been indicated that the developed model is capable of accurately predicting known (measured) pavement structural responses, it can be used to predict critical pavement responses considered to be pavement structural-life predictors. Since the main cause of slab deterioration are cracks that can be considered as indicators of tensile failure in concrete slabs that will happen when tensile stresses exceed the concrete flexural strength, the maximum horizontal tensile stress, $\sigma_{xx,max}$, obtained from the verified model at the bottom edge of the PCC slab, is considered to be a predictor of pavement structural life Brill (1998). The maximum horizontal tensile stress is

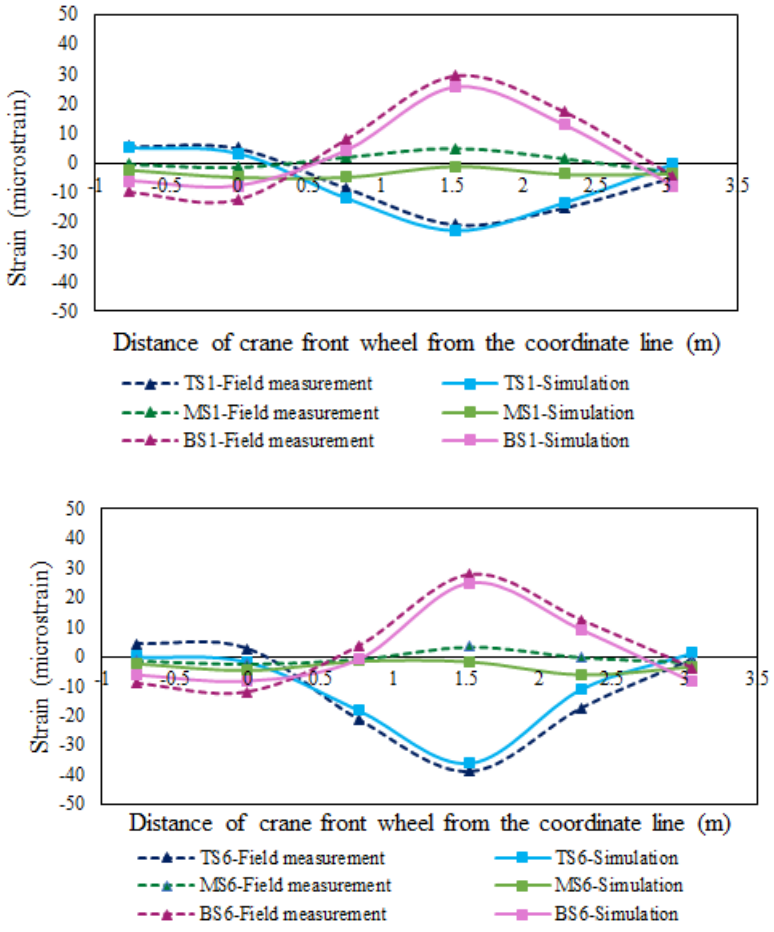

Figure 6.9: Field measurements versus estimated strain.

obtained for different load locations to emphasize the importance of the load location in the concrete fatigue life. In the remainder of this section we first introduce the various fatigue predictive models based on the existing literature. These models will be implemented to predict the fatigue life of the crane-way pavement based on the subsequent FE analysis outputs.

6.5.1 Crane way pavement fatigue analysis

Rigid pavement structures deterioration occurs in the form of fatigue cracks over the pavement service life. The deterioration process can be accelerated with the introduction of heavier loads that can change the boundary conditions through subgrade support loss and slab curling. Predicting fatigue life of a rigid pavement can be performed by calculating a slabs critical tensile stress and implementing fatigue equations (S-N curves) to predict the allowable number of load repetitions until material failure based on the applied stress ratio (tensile bending stress divided by flexural strength of the concrete slab). Smith, et al., (Smith and Roesler 2004) have conducted a review of fatigue models for concrete airfield pavement design, and Table 6.2 presents a summary of the fatigue models they discussed in their study.

Table 6.2 summarizes key characteristics for fatigue models in terms of some of the key factors of importance in fatigue-model development, and the type of data used in the development of fatigue models can be seen to be a primary factor affecting their results. The fatigue models mentioned in Table 6.2 are either based on field data or on concrete beam data, and the primary advantage of a model based on field data is that it can include many influential real-world factors, including direct consideration of traffic and foundation support conditions and indirect consideration of environmental forces, to be considered in evaluating fatigue life. However, the fact that the entire range of possible design and stress conditions are not included, that the number of available field sections for evaluation is generally limited, that the available field studies were conducted over short periods of time, and that the effect of material and construction variability may not be fully reflected, are some of disadvantages in the use of field data. While laboratory testing of concrete beams is a relatively simple, inexpensive, and rapid way of developing relationships between load applications

Table 6.2: A summary of the existing models

Fatigue Model	Data Sources	Failure Definition	Aspects of Stress Computation Used in Development of Fatigue Model			Concrete Strength Characterization
			Location	Stress Types	Computation Method	
COE-Westergaard (Rollings 1981; Rollings 1989)	COE Field Data	50% of slabs cracked	Edge*0.75	Load Only	Westergaard	Flexural (various ages)
COE-LEA (Parker 1979; Rollings and Witczak 1990)	COE Field Data	50% of slabs cracked	Interior	Load Only	Layered Elastic Analysis	Flexural (various ages)
FAA (FAA 1995a)	COE Field Data	50% of slabs cracked	Edge*0.75	Load Only	Westergard	Flexural (various ages)
LEDF AA (FAA 1995b)	COE Field Data	50% of slabs cracked	Interior or Edge*0.75 (times scaling factor)	Load Only	Layered Elastic Analysis	Flexural (various ages)
Darter (Darter 1990)	COE Field Data	50% of slabs cracked	Edge*0.75	Load Only	H-51	Flexural (various ages)
Foxworthy (Foxworthy 1985)	COE Field Data	50% of slabs cracked	Edge	Load Only	Finite Element (ILLI-SLAB)	Flexural (various ages)
NCHRP 1-26 (Thompson and Barenberg 1992)	COE Field Data & AASHO Road Test Data	50% of slabs cracked	Edge	Load and Temperature Curling	Finite Element (ILLI-SLAB)	28-Day MR (third point loading)
Portland Cement Association (Packard 1973)	Concrete Beams	Complete Beam Fracture	Bottom of Concrete	Load Only	Beam Bending Equation	Flexural (various ages)
Zero-Maintenance (Darter 1977)	Concrete Beams	Complete Beam Fracture	Bottom of Concrete	Load Only	Beam Bending Equation	Flexural (various ages)

and failure, there are several factors affecting the results of laboratory testing, including loading rate, loading sequence, stress ratio (ratio of minimum to maximum stress levels), etc. In addition to using different data sources, fatigue models can also differ based on the failure criterion (e.g., 50 percent of the slabs cracked versus failure of concrete beam) used to develop the model. The particular stress computation process used in developing the fatigue model is also important. There are several stress computation methods, including Westergaard edge stress equations, elastic-layered theory, and finite-element methods. The location of the critical response (stress) also varyies in different equations. Because of such differences in the way fatigue models are developed, they cannot be simply substituted for one another in a given design procedure. That being said, based on stress computation methods, Foxworthy, NCHRP 1-26, PCA, and Zero-maintenance methods were selected for use in predicting fatigue life of the pavement. Roesler et al., also conducted a series of large-scale concrete fatigue tests and suggested a comprehensive model that considers temperature stresses, built-in curling stresses, stress range, peak stresses, and loading rate Roesler and Barenberg (1999); Hiller and Roesler (2005); Roesler et al. (2005). Other than the aforementioned equations for predicting fatigue life of a concrete beam or slab, the Model Code Taerwe et al. (2013) also provides a predictive model for concrete structures fatigue life taking into account both pure compression and a combination of compression and tension. The remainder of this section describes prediction of the fatigue life. the number of load cycles to failure (the failure criterion might be different for each fatigue predictive model), using all the aforementioned equations, and compares the results for two different load locations, mid and edge slab. Design recommendations will be subsequently provided. It is worth note that, in order to perform valid fatigue analysis, the most unfavorable node (for each load location) should be considered. These analyses are based on the FE model results.

6.5.1.1 Load location 1: Mid-slab

A schematic diagram of the concrete slabs under the mid-slab load along with the base course and subgrade soil layers and their corresponding critical responses is presented in Figure 6.10. The

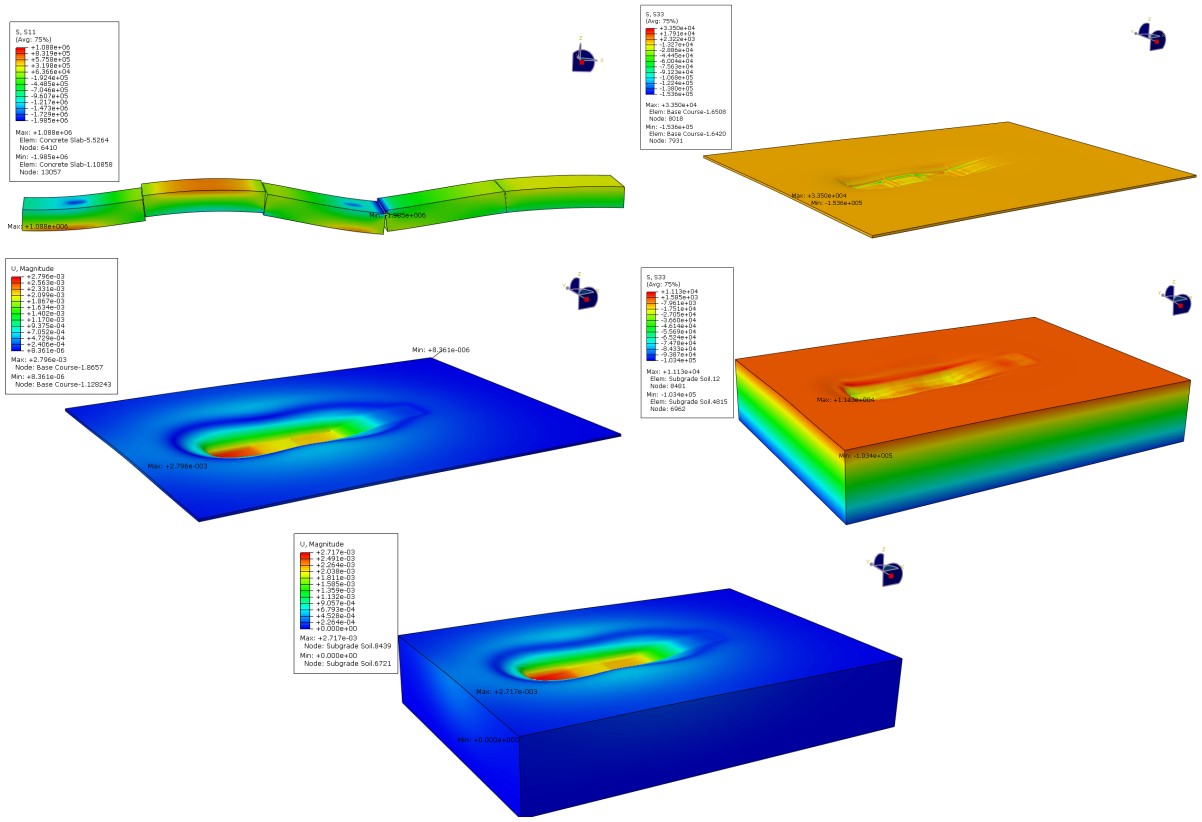

Figure 6.10: Deformed shape ($scale = 500$) of one of the slabs.

load was applied on the slabs by two crane wheels, obtaining $\sigma_{xx,max}$ is 1.088×10^6 Pa that occurs directly beneath the load at the bottom edge of the concrete slab; it was smaller than the concrete slab flexural strength (f_f, 6.24×10^6 Pa, obtained from the following equations from ACI 318

$$E = 15000\sqrt{f'_c} \tag{6.13}$$

$$f_f = 0.44 f_c^{2/3} \tag{6.14}$$

where E is concrete elastic modulus, f'_c is concrete compressive strength,and f_f is concrete flexural strength.

Table 6.3: Pavement layers critical response parameters

Layer	Critical Response	Value (Pa),(m)
Concrete	$\sigma_{xx,max}$	1.088×10^{6}
Base Course	$\sigma_{zz,max}$	3.35×10^{4}
	$\delta_{zz,max}$	2.795×10^{-3}
Subgrade Soil	$\sigma_{zz,max}$	1.13×10^{4}
	$\delta_{zz,max}$	2.717×10^{-3}

Table 6.4: The number of cycles to failure and fatigue

Fatigue Model	Cycles to failure	Fatigue life (year)
Model Code pure compression	9.10^{25}	4.16×10^{21}
Model Code pure tension and tension-compression with $\sigma_{ct,max}$>0.026$\lvert\sigma_{c,max}\rvert$	2.92×10^{7}	1.33×10^{3}
FoxWorthy	-	-
NCHRP 1-26	-	-
PCA	4.72×10^{9}	2.15×10^{5}
Zero maintenance	3.29×10^{14}	1.50×10^{10}
MEPDG	4.98×10^{16}	2.27×10^{12}
FAA Compression	2.27×10^{14}	1.04×10^{10}
FAA Tension	1.12×10^{12}	5.11×10^{7}
Roesler	1.21×10^{5}	5.52

The number of cycles to failure and the pavement fatigue life were calculated for each fatigue model, with an assumed average daily operation of 60 load repetitions. The number of cycles to failure for a mid-slab load is presented in Table 6.4.

It should be mentioned that concrete fatigue equations are dependent on four main factors that explain why so many different fatigue algorithms exist even for the same data sets:

- Specimen size and geometry, loading configuration, and boundary conditions (generally called the size effect)
- Theoretical model for stress calculation (plate theory, layered elastic, or finite-element analysis)

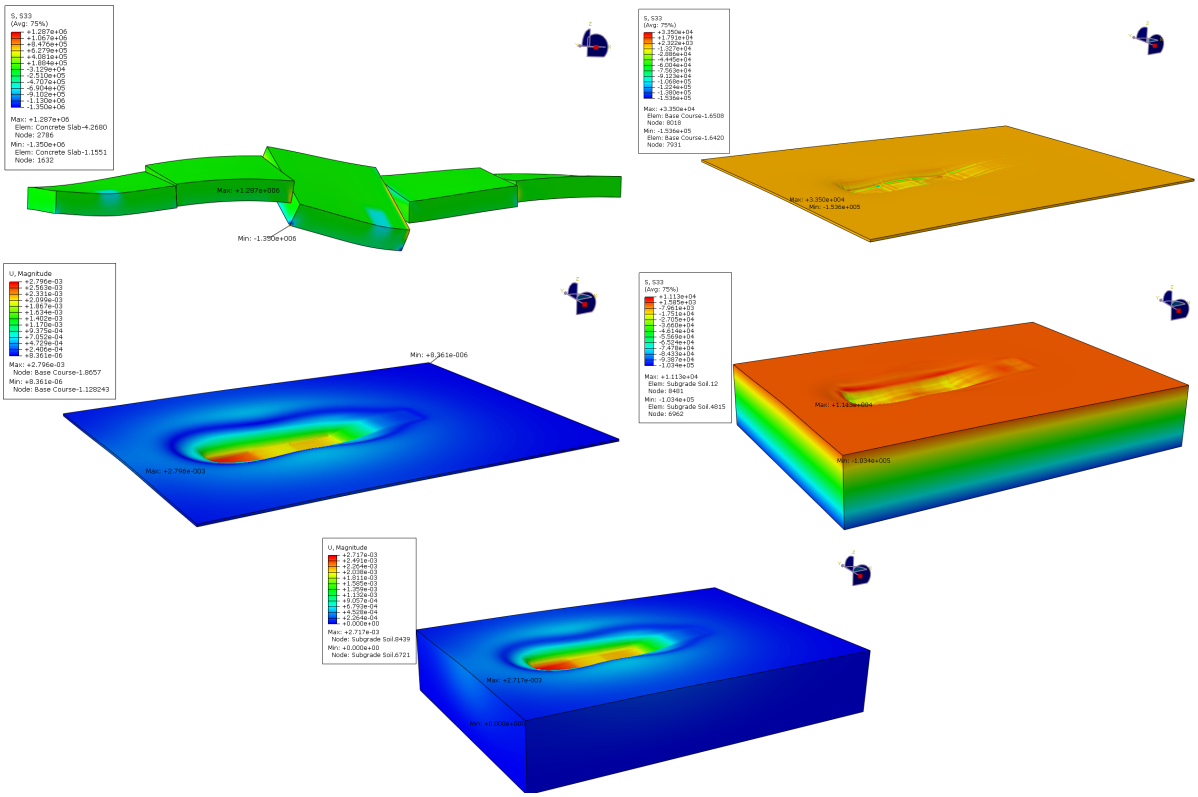

Figure 6.11: Deformed shape ($scale = 500$) of one of the slabs.

- Bending stresses included (temperature and moisture curling and mechanical loading)
- Method of counting stress repetitions

6.5.1.2 Load location 2: Slab edge

A schematic of the concrete slabs under the mid slab load, along with the base course and subgrade soil layers and their corresponding critical responses, is presented in Figure 6.11. The load was applied on the slabs by two of the crane wheels. The pavement critical responses and their corresponding fatigue life calculations are presented in Tables 6.5 and 6.6 ,respectively.

By comparing the fatigue life results obtained for two different load locations, one can see that the edge load is more critical in terms of structural critical response of the pavement. Monitoring the exact load positions and load repetitions would result in a cumulative damage study of the crane-way pavement. According to the obtained results, the most conservative fatigue models with 3, 7, and 12

Table 6.5: Pavement layers critical response parameters

Layer	Critical Response	Value (Pa),(m)
Concrete	$\sigma_{xx,max}$	1.626×10^{6}
Base Course	$\sigma_{zz,max}$	4.195×10^{4}
	$\delta_{zz,max}$	3.331×10^{-3}
Subgrade Soil	$\sigma_{zz,max}$	2.134×10^{4}
	$\delta_{zz,max}$	3.010×10^{-3}

Table 6.6: The number of cycles to failure and fatigue life

Fatigue Model	Cycles to failure	Fatigue life (year)
Model Code pure compression	2.10×10^{23}	9.10×10^{17}
Model Code pure tension and tension-compression with $\sigma_{ct,max}$>0.026$\lvert\sigma_{c,max}\rvert$	1.68×10^{5}	7.69
FoxWorthy	4.28×10^{5}	1.95×10^{1}
NCHRP 1-26	2.63×10^{14}	1.20×10^{10}
PCA	4.16×10^{8}	1.90×10^{4}
Zero maintenance	9.78×10^{12}	4.47×10^{8}
MEPDG	1.71×10^{10}	7.82×10^{5}
FAA Compression	1.72×10^{14}	7.87×10^{9}
FAA Tension	6.00×10^{10}	2.74×10^{6}
Roesler	6.58×10^{4}	3.01

years predicted fatigue life are Roesler, Model Code, and FortWorthy, respectively. As previously mentioned, intermodal facilities often must deal with premature-cracking in the first one or two years of the crane-way pavement service life. According to the obtained critical response parameter ($\sigma_{xx,max}$), while concrete slabs have a high flexural strength value, 100% saturated clay does not provide sufficient support for pavement for large applied loads. Previous studies have shown, that during a wet season, cyclic loading of heavy traffic can result in pumping of fine particles from the subgrade into the granular layers, redistribution of materials beneath slabs, and progressive ejection of materials through pavement joints Kermani et al. (2018). During this process pavement will lose its subgrade support and may face several distresses (premature cracking, faulting, etc.) resulting to pavement failure. Moreover, at the end of wet season subgrade settlements will result in another set of premature cracks that are extended continuously to several slabs. Therefore, to design specific heavy-duty pavements, it is crucial to consider the contribution of pavement foundation to the overall pavement performance. A damage evaluation along with a parametric analyses on material strength and slab geometry are conducted and discussed in the next section.

6.5.1.3 Fatigue damage evaluation

As discussed previously, since the slabs were not made of reinforced concrete, the fatigue damage due to tension in concrete is the main concern. For a general stress state, while the stress at a specific direction should be selected for cycle counting and damage evaluation, it is known that a stress gradient can cause redistribution of concrete stress, and if this effect is ignored the obtained fatigue damage will increase slightly. For damage evaluation purposes, a load was applied at two different load locations, slab middle and slab edge, for 20 seconds and the analysis considered horizontal tensile stress, (σ_{xx}), at the most unfavorable node. Fatigue damage evaluation was then conducted following the Mode Code 2010 formulation. For the tensile stress, σ_{xx}, cycle counting was carried out using the rain-flow algorithm developed by Downing and Socie Downing and Socie (1982). A subroutine was developed in MATLAB for performing this task on stress time histories and to present the counted cycles in the matrix shown in Figure 6.12 and Figure 6.13 for each

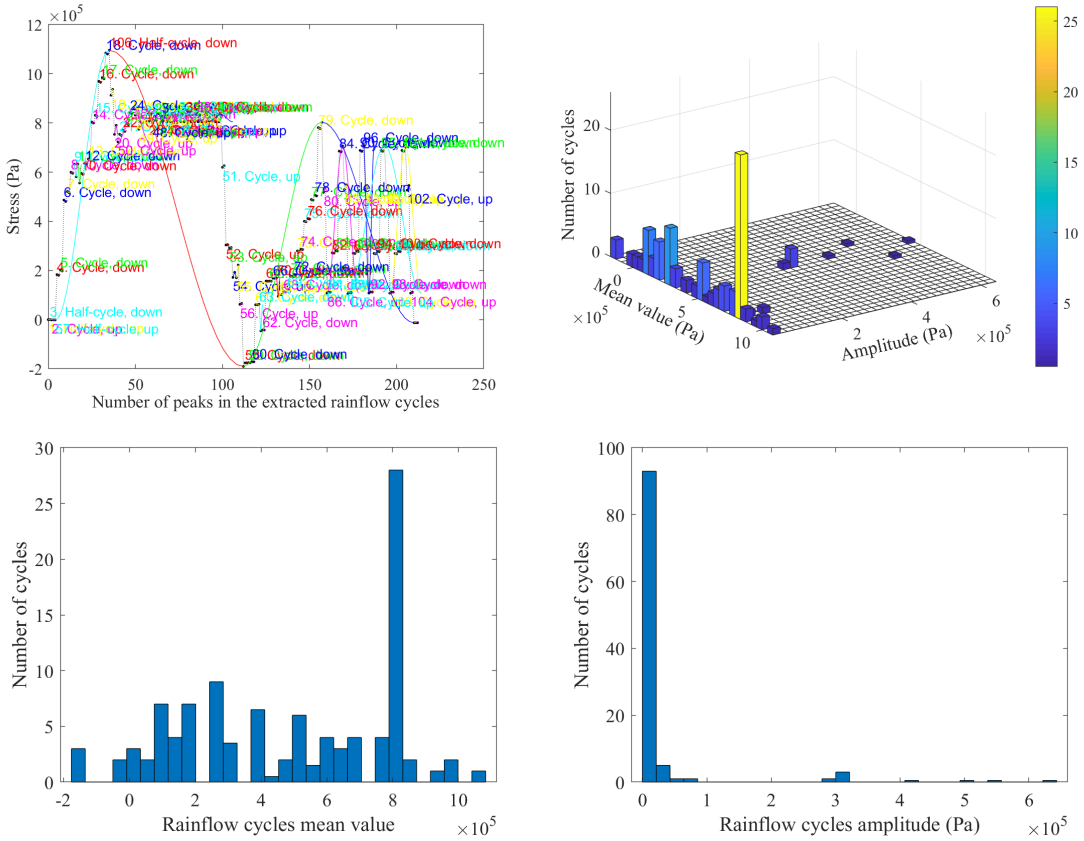

Figure 6.12: The total number of 106 cycles counted.

stress range and mean. For each stress range, the fatigue life N was calculated according to the Model Code using the following equations both for pure tension and for tension-compression with $\sigma_{ct,max} > 0.026|\sigma_{c,max}|$:

$$LogN = 12(1 - S_{ct,max}) \tag{6.15}$$

$$S_{ct,max} = \frac{\sigma_{ct,max}}{f_{ctk,min}} \tag{6.16}$$

where N is the number of cycles to failure, $S_{ct,max}$ is the maximum tensile stress level, $\sigma_{ct,max}$ is the maximum tensile stress in MPa, and $f_{ctk,min}$ is the minimum characteristic tensile strength.

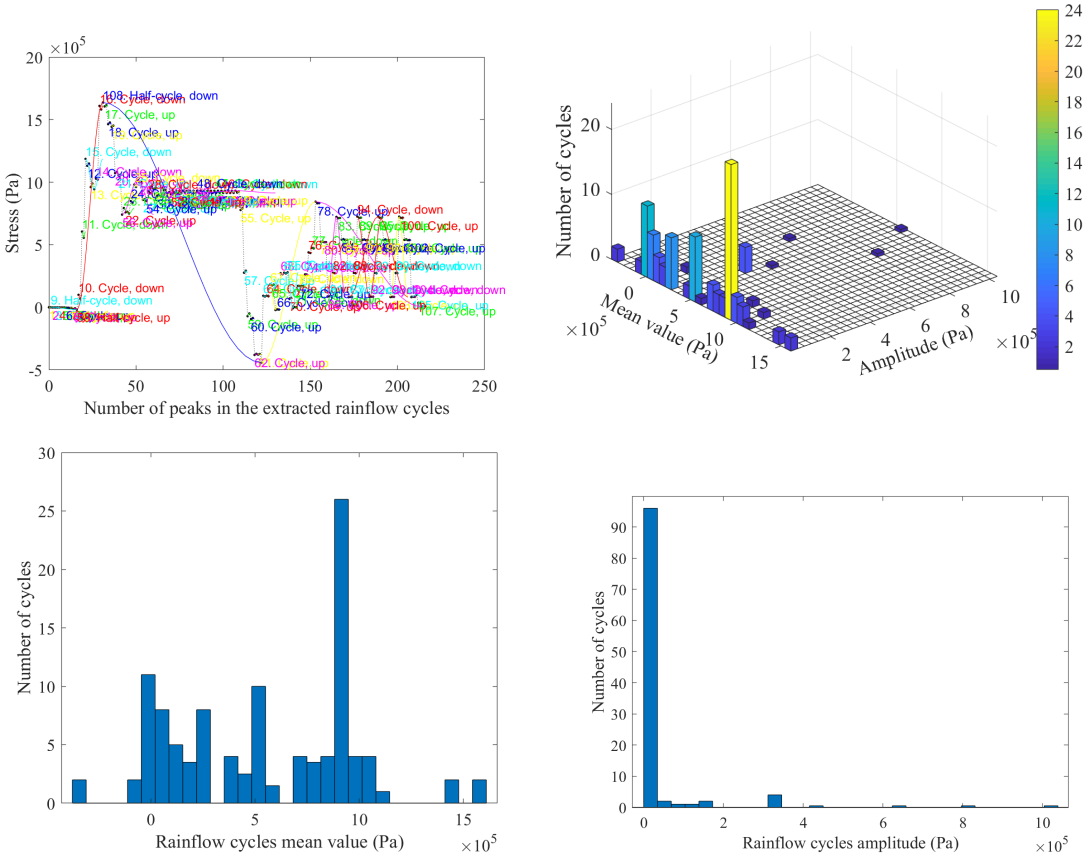

Figure 6.13: The total number of 108 cycles counted

Based on Miner's rule, damage resulting from a given stress range is a linear function of the number of cycles of that stress range. The Miner's rule formula can be expressed as:

$$D_i = \frac{n_i}{N_i} \tag{6.17}$$

where N_i is the number of cycles needed to cause failure at the i^{th} stress range S_i, and n_i is the number of cycles of the stress range S_i. For both load locations, the damage, D_i for each stress range is determined and presented in Figure 6.14. The total damage D can be calculated using the following equation:

$$D = \sum_{i}^{N} D_i \tag{6.18}$$

According to Miner's rule, the failure occurs if the total damage, D, is equal to 1. The total estimated damage for 10 seconds of loading (one crane passage) for mid-slab load was 0.0000176% for edge-slab load it was 0.00149%.

6.5.1.4 Parametric analysis and design recommendations

This section presents results from studies on the effect of concrete compressive strength, concrete slab geometry, elastic modulus of the base course material, and thickness of the base layer, on damage at the most unfavorable node. Each of these individual parametric studies were done while keeping constant the other material properties and model geometries, and the results are presented in Figures 6.15 to 6.19. According to Figure 6.15, damage decreases with increasing base course elastic modulus. However, concrete pavements do not require strong foundation support since the load is distributed over large areas of the subgrade soil and therefore, deflections are small. It is much more important that the support be reasonably uniform with no abrupt changes in degree of support. To provide an appropriate base, its materials should meet minimum requirements to prevent mud-pumping of subgrade soil. Moreover, proper gradation controls should be applied to ensure a reasonably unique support for the concrete pavement. Studying the effect of base-course thickness on damage at the most unfavorable node of the model indicated that, while increasing the base-course thickness would decrease damage, as presented in Figure 6.16, a suitable base-course thickness should be selected based on construction specifications and material availability.

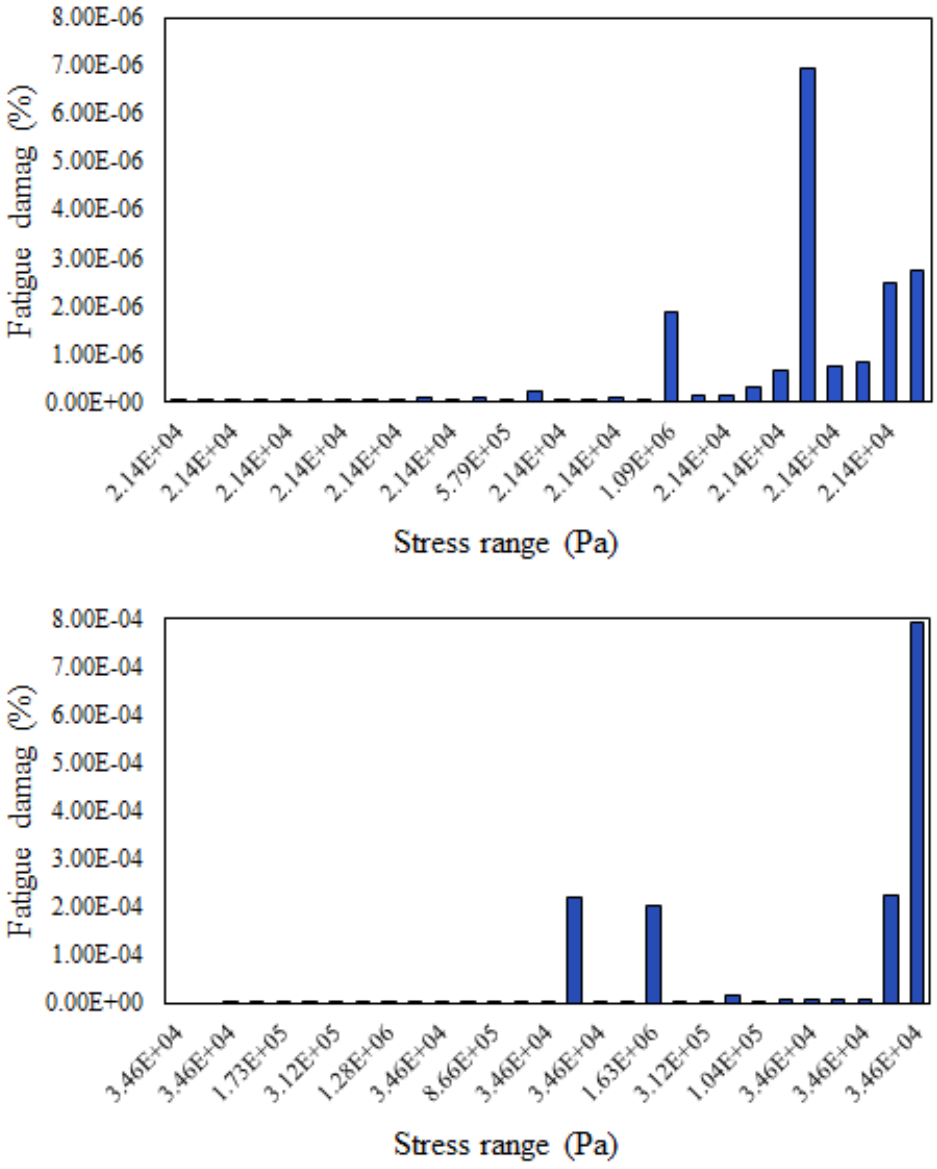

Figure 6.14: The fatigue damage estimated

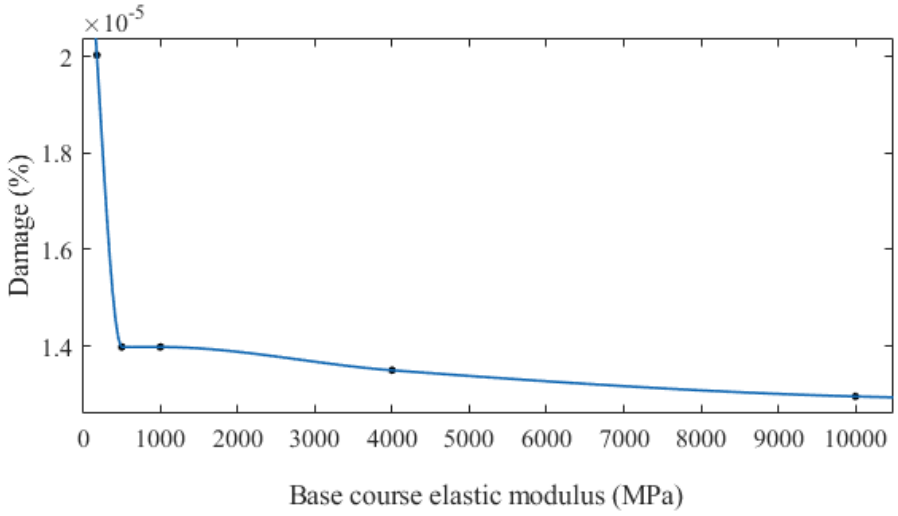

Figure 6.15: Damage of the most unfavorable node

According to Figure 6.17, while damage decreases rapidly with an increase in concrete compressive strength, for compressive strength values greater than 50 MPa the damage reduction rate decreases, so 50 MPa appears to be a reasonable value for the concrete compressive strength. According to Figure 6.18 increasing concrete slab thickness can drastically decrease the amount of damage; when it changed from 0.4 m to 0.533 m and the damage reduction rate decreases after this point. It is worth pointing out that the accumulated damage, D, for slab thickness of 0.2 m is one, meaning that the slab will be cracked under the applied load. Figure 6.19 describes the effect of concrete slab width-to-length ratio ($\frac{w}{l}$) (when $l = 10ft. = 3.048m$). When ($\frac{w}{l}$) ratio changes from 0.7 to 0.82, damage decreases by $10^{-5}\%$, after which the damage reduction rate will decrease.

6.6 Conclusions

The focus of this study was evaluating gantry crane-way pavement performance at intermodal facilities. First, 18 vibrating-wire strain gages were installed before concrete was poured. These gages were used to capture pavement structural response, ϵ_{xx}, when the crane travels over the slabs. Several loading locations at which the crane stopped for around 40 seconds each were selected, and the tire footprint was also captured in the field to be used for further image processing. The amount

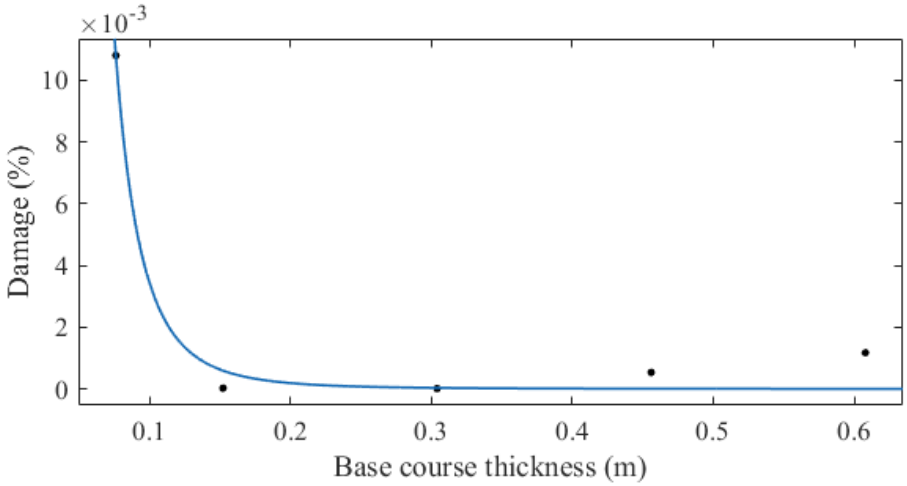

Figure 6.16: Damage of the most unfavorable node

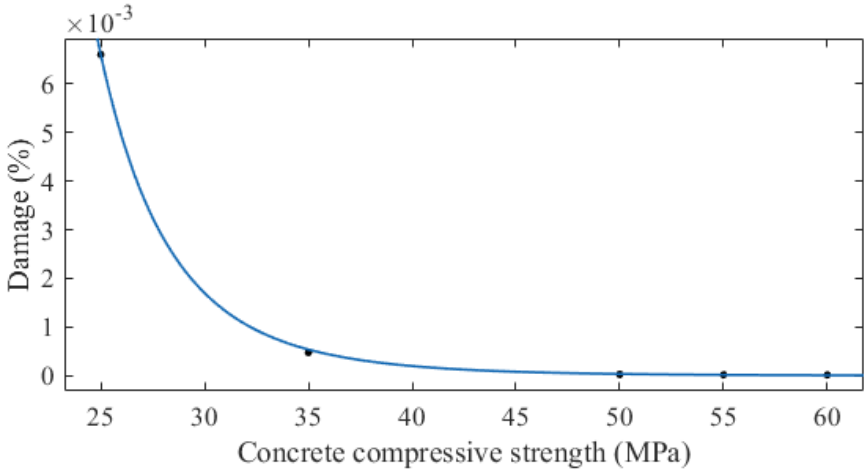

Figure 6.17: Damage of the most unfavorable node

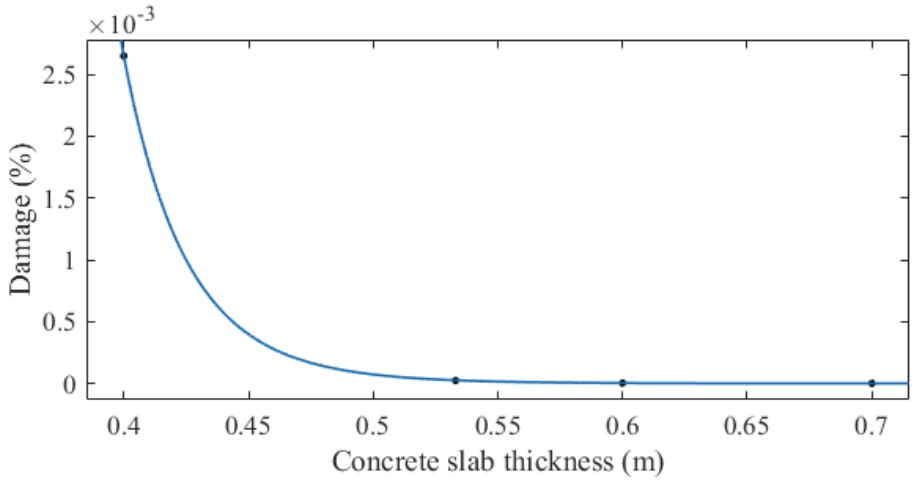

Figure 6.18: Damage of the most unfavorable node

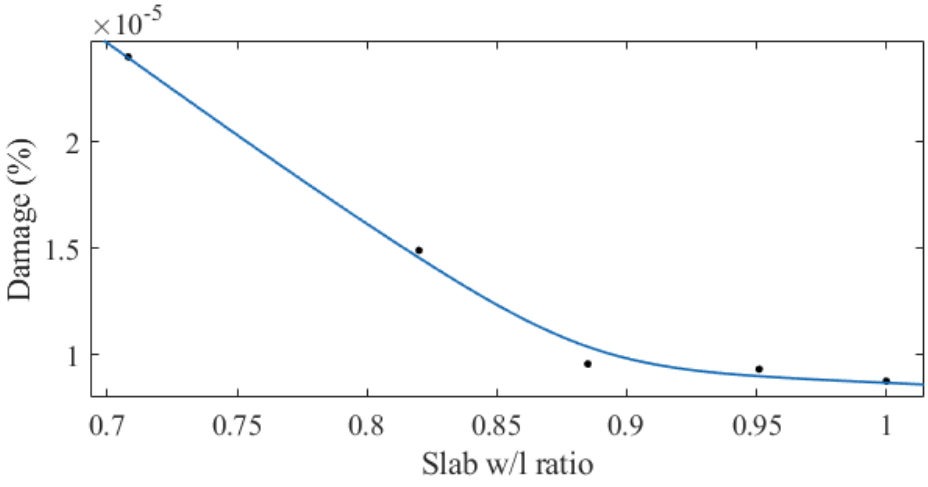

Figure 6.19: Damage of the most unfavorable node

of tire pressure was also obtained for use in further modeling. Second, to monitor the structural behavior of the pavement, a finite-element-based model was developed reflecting the same geometry, boundary conditions, and applied load as those in the field. Since the strain values were the only field measurements, an inverse approach was used to obtain a set of material properties to be used in the model for determining the ϵ_{xx} value closest to the field measurements. The inverse approach begins with an assume value for the unknown parameter and obtains a solution to the forward model problem. Once the forward model has been solved for the assumed value of input, the next step is to compute the error between the simulated and experimental measurements, and the process will iterate until the convergence goal (minimum error) has been achieved. Using the obtained material properties, the finite-element model is verified against field measurements (using the same inputs and same load and boundary conditions to give the closest ϵ_{xx} at the gage locations), and the verified finite-element model is then used to study pavement critical response parameters. The maximum vertical deflection and maximum vertical stress at the top of the base and the subgrade and the maximum tensile stress at the bottom of the concrete slab are obtained for mid-slab and edge-load locations. After evaluating the obtained critical response parameters, a fatigue damage analysis was conducted for a mid-slab load and an edge load. The amount of accumulated damage after one crane passage with a load applied on the slab edge was found to be approximately 85 times that for a load applied at the slab middle. Finally, a parametric analysis of the material properties and slab geometry was conducted, and based on the obtained results the following design recommendations are suggested:

- The concrete compressive strength should be no less than 50 MPa to decrease the amount of damage.
- The slab thickness should be no less than 0.5 m to decrease the amount of damage.
- The slab width-to-length ratio (with $l = 3.048m$) should increase from 0.8 to 0.9 to reduce the produced damage.

- The design procedure should consider the contribution of foundation behavior to overall pavement performance rather than focusing only on PCC layer failure.
- Due to the heavy applied load, while fine particles will be pumped up from the subgrade into granular base and ejected through the joints, there are several soil stabilization techniques that could be used to avoid material loss and premature cracking of the pavement.
- A permeable base material should be considered for this heavy duty pavement which experience indicated the potential for pavement faulting and pumping.
- Since edge-loading is more destructive than the mid-slab loading, the crane driver should avoid slab edges.

A longer structural-health monitoring analysis could be implemented for the pavement section, and the results could be used in future studies to develop a probability damage study and predict the remaining life of the pavement sections under each loading condition. Such analyses will be conducted in future studies.

6.7 References

Brill, D. (1998). Development of advanced computational models for airport pavement design. *Rep. No. DOT/FAA/AR-97*, 47.

Downing, S. D. and Socie, D. (1982). Simple rainflow counting algorithms. *International journal of fatigue*, 4(1):31–40.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. (2019). Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: elimination of correlated inputs and extrapolation in modeling. *Structural and Multidisciplinary Optimization*, 59(4):1335–1353.

Grassl, P. and Jirásek, M. (2006). Damage-plastic model for concrete failure. *International journal of solids and structures*, 43(22-23):7166–7196.

Hiller, J. E. and Roesler, J. R. (2005). Determination of critical concrete pavement fatigue damage locations using influence lines. *Journal of Transportation Engineering*, 131(8):599–607.

Huang, Y. H. (1993). *Pavement analysis and design*. Nptel.

Kent, D. C. and Park, R. (1971). Flexural members with confined concrete. *Journal of the Structural Division*.

Kermani, B., Stoffels, S. M., Xiao, M., and Qiu, T. (2018). Experimental simulation and quantification of migration of subgrade soil into subbase under rigid pavement using model mobile load simulator. *Journal of Transportation Engineering, Part B: Pavements*, 144(4):04018049.

Knapton, J. (1989). *The structural design of heavy duty pavements for ports and other industries*.

Maitra, S. R., Reddy, K., and Ramachandra, L. (2010). Load transfer characteristics of aggregate interlocking in concrete pavement. *Journal of Transportation Engineering*, 136(3):190–195.

Miller, J. S., Bellinger, W. Y., et al. (2014). Distress identification manual for the long-term pavement performance program. Technical report, United States. Federal Highway Administration. Office of Infrastructure.

Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Roesler, J. R. and Barenberg, E. J. (1999). Fatigue and static testing of concrete slabs. *Transportation Research Record*, 1684(1):71–80.

Roesler, J. R., Hiller, J. E., and Littleton, P. C. (2005). Large-scale airfield concrete slab fatigue tests. In *8th International Conference on Concrete Pavements: Innovations for Concrete Pavement: Technology Transfer for the Next Generation*, pages 1247–1268.

Sadeghi, V. and Hesami, S. (2018). Investigation of load transfer efficiency in jointed plain concrete pavements (jpcp) using fem. *International Journal of Pavement Research and Technology*, 11(3):245–252.

Shoukry, S. N., Fahmy, M., Prucz, J., and William, G. (2007). Validation of 3dfe analysis of rigid pavement dynamic response to moving traffic and nonlinear temperature gradient effects. *International Journal of Geomechanics*, 7(1):16–24.

Tabatabaie, A. M. and Barenberg, E. J. (1978). Finite-element analysis of jointed or cracked concrete pavements. *Transportation Research Record*, (671).

Taerwe, L., Matthys, S., et al. (2013). Fib model code for concrete structures 2010.

Uddin, W., Hackett, R. M., Joseph, A., Pan, Z., and Crawley, A. B. (1995). Three-dimensional finite-element analysis of jointed concrete pavement with discontinuities. *Transportation Research Record*, 1482:26–32.

CHAPTER 7. SUMMARY AND DISCUSSION

Maintenance, rehabilitation, and reconstruction of the highway system are the major expenses in a state general expenditure. Thus, predicting pavement performance and detecting the current state of pavement health are crucial for pavement agencies as well as state departments of transportation (DOT). In pavement engineering, performance models are mathematical expressions that relate pavement response to several factors including material components' properties, loading condition and frequency, and climatic condition. Over the past decades, researchers have developed numerous models to predict pavement performance. However, there are several issues in conventional modeling approaches. In this thesis, the existing issues are discussed and a new robust framework for predicting pavement performance is developed.

One of the performance characteristics of asphalt pavement is its resistance to rutting. Rutting or permanent deformation defines as surface depression in the wheelpath and can cause safety issues for drivers when it is filled with water. It has been demonstrated that rutting can be defined as a function of material properties, loading condition, and temperature. However, these predictors can be correlated and not appropriate to be used in modeling. Seventeen input variables are selected to be used in model development. This study focuses on evaluating the quality of input variables through pairwise correlation analysis. A multivariate statistical technique called Principal Component Analysis (PCA) is implemented to first reduce the dimensionality of the data set by using five pseudo-inputs instead of the 17 original input variables, and second to eliminate the pairwise correlation within the inputs by transforming the original data to a new coordinate system. Two separate model is produced to predict rutting using artificial neural networks (ANN) and multivariate regression analysis. The developed models are called PCNN and PCR, respectively, and both modeled the amount of permanent deformation satisfactorily with PCNN fitting the test data significantly better. Since the developed models are based on empirical data they are prone

to extrapolation. To prevent extrapolation from happening, a simple method is developed to characterize the effective variable space. The obtained n-dimensional enclosing ellipsoid is added as a constraint to the problem. To indicate one of the applications of the developed framework, an optimization problem of finding minimum amount of permanent strain and its corresponding design parameters is defined and solved.

The proposed framework is implemented to predict another performance related property of asphalt mixture called dynamic modulus. Dynamic modulus is the stiffness characteristic of asphalt mixture and defines the stress-strain relationship of the material due the compressive sinusoidal load. Two separate dynamic modulus predictive models are developed using artificial neural network as well as multivariate regression analysis. Performances of the proposed models were compared to conventional predictive models using both statistical analysis and Receiver Operating Characteristic (ROC) analysis. Both models predicted the amount of dynamic modulus remarkably better that the traditional models. Two optimization problems including optimal design and inverse design have been presented and solved using a mean-variance mapping optimization algorithm. The results are in a good agreement with current flexible pavement design specifications and thus can be a reasonable starting point in solving real-life design problems. Moreover, these methodologies appear to also have much promise in modeling other material properties at other effective temperatures and thus can be used to be a proper substitute to conventional performance models, i.e. fatigue, rutting, dynamic modulus, low temperature behavior. Besides, the ability of the framework to be retrained with more, or different data from different condition and location, (dynamic nature) of the developed framework will make local calibration of the traditional empirical models unnecessary.

Another way of tackle the problem of predicting pavement performance is also discussed in this study. The structural response of a concrete pavement to the applied load is studied by installing strain gages during construction. The strain measurements are used to validate a finite element model of the pavement section with similar geometry and boundary conditions as those in the field. In the absence of some field measurements, an inverse algorithm is implemented to find the

unknown material properties. The algorithm assumes an initial value for the unknown parameter that is under investigation and validates the solution to the forward model problem. Once the forward model is solved for the assumed value of input, it calculates the error between the simulated and experimental measurements. If the error belongs to a prespecified limit, the initial guess is the solution to the inverse problem; otherwise, the process is repeated until the condition is satisfied. After finding the unknown material properties, the finite element model is validated against the field measurements. The validated model is used to predict pavement critical response as well as predicting fatigue life of the pavement. Although, the developed predictive framework as well as obtained optimal design parameters are based on the empirical database created in this study, the suggested framework has the capability of being re-trained and adjusted to fit new data. For obtaining more reliable and applicable results, a larger empirical database would be required.

References

AASHTO, A. (2008). Mechanistic-empirical pavement design guide: A manual of practice. *Aaoshat Off. Ed.*

AASHTO TP 79-13 (2013). AASHTO TP 79-13 Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Aphalt Mixtures Performance Tester. *Am. Assoc. State Highw. Transp. Off.*, pages 1–19.

Al-Khateeb, G., Shenoy, A., Gibson, N., and Harman, T. (2006). A new simplistic model for dynamic modulus predictions of asphalt paving mixtures. *J. Assoc. Asph. Paving Technol.*, 75.

Andrei, D., Witczak, M., and Mirza, M. (1999). Development of a revised predictive model for the dynamic (complex) modulus of asphalt mixtures. *Dev. 2002 Guide Des. New Rehabil. Pavement Struct. NCHRP.*

Apeagyei, A. K. (2011). Rutting as a Function of Dynamic Modulus and Gradation. *Am. Soc. Civ. Eng.*, 23 (9):1302–1310.

ARA, Inc., E. C. D. (2004). Guide for mechanistic–empirical design of new and rehabilitated pavement structures. *Final Rep., Nchrp Proj. 1-37a.*

Arabali, P., Sakhaeifar, M. S., Freeman, T. J., Wilson, B. T., and Borowiec, J. D. (2017). Decision-making guideline for preservation of flexible pavements in general aviation airport management. *J. Transp. Eng. Part B: Pavements*, 143(2):04017006.

Aslani, M. and Asla, R. (2010). A Novel Hybrid Simplex-Genetic Algorithm For The Optimum Design Of Truss Structures. *Lect. Notes Comput. Sci.*, II.

Aslani, M., Asla, R. N., Oftadehb, R., and Panahic, M. S. (2010). A novel hybrid simplex-genetic algorithm for the optimum design of truss structures. In *Proceedings of the world congress on engineering*, volume 2.

Aslani, M., Ghasemi, P., and Gandomi, A. H. (2018). Constrained mean-variance mapping optimization for truss optimization problems. *Struct. Des. Tall Spec. Build.*, 27(6):e1449.

Bari, J. and Witczak, M. (2007). New predictive models for viscosity and complex shear modulus of asphalt binders: For use with mechanistic-empirical pavement design guide. *Transp. Res. Rec. : J. Transp. Res. Board*, 1(2001):9–19.

Bashin, A., Masad, E., Kutay, M. E., Buttlar, W., Kim, Y.-R., Marasteanu, M., Kim, Y. R., Schwartz, C. W., and Carvalho, R. L. (2012). Applications of advanced models to understand behavior and performance of asphalt mixtures. *Transp. Res. E-Circ.*, 1(E-C161).

Bi, J. and Bennett, K. P. (2003). Regression error characteristic curves. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 43–50.

Birgisson, B., Roque, R., Kim, J., and Pham, L. V. (2004a). The Use of Complex Modulus To Characterize Performance of Asphalt Mixtures and Pavements in Florida. Technical Report September 2004, Florida Department of Transportation, Florida Department of Transportation.

Birgisson, B., Roque, R., Kim, J., and Pham, L. V. (2004b). The use of complex modulus to characterize the performance of asphalt mixtures and pavements in florida. Technical report.

Birgisson, B., Sholar, G., and Roque, R. (2005). Evaluation of a predicted dynamic modulus for florida mixtures. *Transp. Res. Rec.*, 1929(1):200–207.

Bozorgzad, A. et al. (2017). Consistent distribution of air voids and asphalt and random orientation of aggregates by flipping specimens during gyratory compaction process. *Constr. Build. Mater.*, 132:376–382.

Brill, D. (1998). Development of advanced computational models for airport pavement design. *Rep. No. DOT/FAA/AR-97*, 47.

Brown, E., Cooley, L. A., Hanson, D., Lynn, C., Powell, B., Prowell, B., and Watson, D. (2002). Ncat test track design, construction, and performance.

Cai, Z., Cai, Z., and Wang, Y. (2006). A Multiobjective Optimization-Based Evolutionary Algorithm for Constrained Optimization. *Evol. Comput. IEEE Trans.*, 10(6):658–675.

Carpenter, W. and Barthelemy, J.-F. (1993). A comparison of polynomial approximations and artificial neural nets as response surfaces. *Struct. Multidiscip. Optim.*, 5(3):166–174.

Cheng, B. and Titterington, D. (1994a). Neural networks: A review from a statistical perspective.

Cheng, B. and Titterington, D. M. (1994b). Neural networks: A review from a statistical perspective. *Statistical science*, pages 2–30.

Christensen Jr, D., Pellinen, T., and Bonaquist, R. (2003). Hirsch model for estimating the modulus of asphalt concrete. *J. Assoc. Asph. Paving Technol.*, 72.

Cominsky, R. J., Huber, G. A., Kennedy, T. W., and Anderson, M. (1994). *The superpave mix design manual for new construction and overlays*. Number SHRP-A-407. Strategic Highway Research Program Washington, DC.

Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences*. Cengage learning.

Downing, S. D. and Socie, D. (1982). Simple rainflow counting algorithms. *International journal of fatigue*, 4(1):31–40.

Ebrahimi, M., ShafieiBavani, E., Wong, R. K., Fong, S., and Fiaidhi, J. (2017). An adaptive meta-heuristic search for the internet of things. *Future Generation Computer Systems*, 76:486–494.

El-Badawy, S., Abd El-Hakim, R., and Awed, A. (2018). Comparing artificial neural networks with regression models for hot-mix asphalt dynamic modulus prediction. *J. Mater. Civ. Eng.*, 30(7):04018128.

Erlich, I., Venayagamoorthy, G. K., and Worawat, N. (2010). A mean-variance optimization algorithm. *IEEE Congr. Evol. Comput.*, pages 1–6.

Fakhri, M., Azami, A., et al. (2017). Evaluation of warm mix asphalt mixtures containing reclaimed asphalt pavement and crumb rubber. *Journal of cleaner production*, 165:1125–1132.

Fathi, A., Mazari, M., Saghafi, M., Hosseini, A., and Kumar, S. (2019). Parametric study of pavement deterioration using machine learning algorithms. In *Airfield and Highway Pavements 2019: Innovation and Sustainability in Highway and Airfield Pavement Technology*, pages 31–41. American Society of Civil Engineers Reston, VA.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognit. Lett.*, 27(8):861–874.

Fodor, I. K. (2002). A Survey of Dimension Reduction Techniques. Technical report, U.S. Department of Energy, U.S. Department of Energy.

García-Segura, T., Yepes, V., and Frangopol, D. M. (2017). Multi-objective design of post-tensioned concrete road bridges using artificial neural networks. *Struct. Multidiscip. Optim.*, 56(1):139–150.

Ghasemi, P. (2018). Performance evaluation of coarse-graded field mixtures using dynamic modulus results gained from testing in indirect tension mode of testing.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. (2018a). Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: Elimination of correlated inputs and extrapolation in modeling. *Struct. Multidiscip. Optim.*, pages 1–19.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. (2019a). Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: elimination of correlated inputs and extrapolation in modeling. *Structural and Multidisciplinary Optimization*, 59(4):1335–1353.

Ghasemi, P., Aslani, M., Rollins, D. K., Williams, R. C., and Schaefer, V. R. (2018b). Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks. *Int. J. Pavement Res. Technol.*

Ghasemi, P., Lin, S., Rollins, D. K., and Williams, R. C. (2019b). Predicting dynamic modulus of asphalt mixture using data obtained from indirect tension mode of testing. *arXiv:1905.06810*.

Ghasemi, P., Podolsky, J., Christopher Williams, R., and Dave, E. (2016). Performance evaluation of coarse-graded field mixtures using dynamic modulus results gained from testing in the indirect tension mode. In *International Conference on Transportation and Development 2016*, pages 1111–1121.

Gong, H., Sun, Y., Mei, Z., and Huang, B. (2018). Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks. *Constr. Build. Mater.*, 190:710–718.

Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., and Maggio, R. M. (2017). Use of principal component analysis (pca) and hierarchical cluster analysis (hca) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends Food Sci. & Technol.*

Grassl, P. and Jirásek, M. (2006). Damage-plastic model for concrete failure. *International journal of solids and structures*, 43(22-23):7166–7196.

Guide, N. D. (2004). Guide 1-37a, guide for mechanistic-empirical design of new and rehabilitated pavement structures, national cooperative highway research program. *Transp. Res. Board Natl. Res. Counc.*

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J. Am. Stat. Assoc.*, pages 1–11.

Hargrove, L., Scheme, E., Englehart, K., and Hudgins, B. (2017). Principal components analysis tuning for improved myoelectric control. *CMBES Proceedings*, 30(1).

He, J. and Yao, X. (2002). From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. *IEEE Trans. Evol. Comput.*, 6(5):495–511.

Hiller, J. E. and Roesler, J. R. (2005). Determination of critical concrete pavement fatigue damage locations using influence lines. *Journal of Transportation Engineering*, 131(8):599–607.

Hosseini, A., Faheem, A., Titi, H., and Schwandt, S. (2020). Evaluation of the long-term performance of flexible pavements with respect to production and construction quality control indicators. *Construction and Building Materials*, 230:116998.

Hua, X., Ni, Y., Ko, J., and Wong, K. (2007). Modeling of temperature–frequency correlation using combined principal component analysis and support vector regression technique. *J. Comput. Civ. Eng.*, 21(2):122–135.

Huang, Y. H. (1993). *Pavement analysis and design*. Nptel.

Jalali, F., Vargas-Nordcbeck, A., and Nakhaei, M. (2019). Role of preventive treatments in low-volume road maintenance program: Full-scale case study. *Transportation Research Record*, page 0361198119863025.

Jamrah, A., Kutay, M. E., and Ozturk, H. I. (2014). Characterization of asphalt materials common to michigan in support of the implementation of the mechanistic-empirical pavement design guide. Technical report.

Johnson, R. and Wichern, D. (1992). Applied multivariate statistical analysis prentice hall new jersey 641.

Johnson, R. A., Wichern, D. W., et al. (2014). *Applied multivariate statistical analysis*, volume 4. Prentice-Hall New Jersey.

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.*, 30(3):487.

Juang, C. H. and Chen, C. J. (1999a). CPT-Based Liquefaction Evaluation Using Artificial Neural Networks. *Comput. -Aided Civ. Infrastruct. Eng.*, 14(3):221–229.

Juang, C. H. and Chen, C. J. (1999b). CPT-Based Liquefaction Evaluation Using Artificial Neural Networks. *Comput. -Aided Civ. Infrastruct. Eng.*, 14(3):221–229.

K Rollins Sr, D. (2015). A One-Dimensional PCA Approach for Classifying Imbalanced Data. *J. Comput. Sci. & Systems Biology*, 8(1):5–11.

Kaloush, K. E., Witczak, M., and Sullivan, B. W. (2003). Simple performance test for permanent deformation evaluation of asphalt mixtures. *Sixth Int. RILEM Symp. Perform. Test. Eval. Bitum. Mater.*, pages 498–505.

Kartam, N. (1994). Neural Netwroks in Civil Engineering: Systems and Application. *J. Comput. Civ. Eng.*, 8(2):149–162.

Kartam, Nabil. Flood, I. (1994). Neural Netwroks in Civil Engineering: Principal and Understanding. *J. Comput. Civ. Eng.*, 8(2):131–148.

Kent, D. C. and Park, R. (1971). Flexural members with confined concrete. *Journal of the Structural Division*.

Kermani, B., Stoffels, S. M., Xiao, M., and Qiu, T. (2018). Experimental simulation and quantification of migration of subgrade soil into subbase under rigid pavement using model mobile load simulator. *Journal of Transportation Engineering, Part B: Pavements*, 144(4):04018049.

Kim, Y. R., Underwood, B., Far, M. S., Jackson, N., and Puccinelli, J. (2011). Ltpp computed parameter: Dynamic modulus. Technical report.

Knapton, J. (1989). *The structural design of heavy duty pavements for ports and other industries*.

Kutner, M. H., Nachtsheim, C., and Neter, J. (2004). *Applied linear regression models*. McGraw-Hill/Irwin.

Kuźniar, K. and Waszczyszyn, Z. (2006). Neural networks and principal component analysis for identification of building natural periods. *J. Comput. Civ. Eng.*, 20(6):431–436.

Kvasnak, A., Robinette, C., and Williams, R. (2007). A Statistical Development of a Flow Number Predictive Equation for the Mechanistic-Empirical Pavement Design Guide.CD-ROM . In Transportation Research Board, Washington, D., editor, *Transportation Research Board, Washington, DC.*, volume 3, pages 1–18, Transportation Research Board, Washington, DC. DC.Transportation Research Board, Washington.

Lever, J., Krzywinski, M., and Altman, N. (2017). Points of significance: Principal component analysis.

Liew, K., Tan, H., Ray, T., and Tan, M. (2004). Optimal process design of sheet metal forming for minimum springback via an integrated neural network evolutionary algorithm. *Struct. Multidiscip. Optim.*, 26(3):284–294.

Lim, J., Jang, Y. S., Chang, H. S., Park, J. C., and Lee, J. (2018). Role of multi-response principal component analysis in reliability-based robust design optimization: An application to commercial vehicle design. *Struct. Multidiscip. Optim.*, pages 1–12.

LOMBARDERO, A. C., Piqueras, V. Y., and Vidosa, F. G. (2012). Automatic design of concrete vaults using iterated local search and extreme value estimation. In *Latin American Journal of Solids and Structures*, volume 9, pages 675–689. Argentinean Association of Computational Mechanics, Brazilian Association of Computational Mechanics, Mexican Association of Numerical Methods in Engineering and Applied Sciences.

Maitra, S. R., Reddy, K., and Ramachandra, L. (2010). Load transfer characteristics of aggregate interlocking in concrete pavement. *Journal of Transportation Engineering*, 136(3):190–195.

Malik, M. R., Isaac, B. J., Coussement, A., Smith, P. J., and Parente, A. (2018). Principal component analysis coupled with nonlinear regression for chemistry reduction. *Combust. Flame*, 187:30–41.

Marti-Vargas, J. R., Ferri, F. J., and Yepes, V. (2013). Prediction of the transfer length of prestressing strands with neural networks. *Comput. Concr.*, 12(2):187–209.

MATLAB (2012). Matheworks, T., 2012. Matlab 2012b. ,.

May, R., Dandy, G., and Maier, H. (2011). Review of input variable selection methods for artificial neural networks. In *Artificial neural networks-methodological advances and biomedical applications*. InTech.

Miller, J. S., Bellinger, W. Y., et al. (2014). Distress identification manual for the long-term pavement performance program. Technical report, United States. Federal Highway Administration. Office of Infrastructure.

Mirhosseini, A. F., Tahami, S. A., Hoff, I., Dessouky, S., and Ho, C.-H. (2019). Performance evaluation of asphalt mixtures containing high-rap binder content and bio-oil rejuvenator. *Construction and Building Materials*, 227:116465.

Nicholas, P. E., Padmanaban, K., Vasudevan, D., and Ramachandran, T. (2015). Stacking sequence optimization of horizontal axis wind turbine blade using fea, ann and ga. *Struct. Multidiscip. Optim.*, 52(4):791–801.

Nobakht, M. and Sakhaeifar, M. S. (2018). Dynamic modulus and phase angle prediction of laboratory aged asphalt mixtures. *Constr. Build. Mater.*, 190:740–751.

Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Noh, Y., Choi, K., and Du, L. (2009). Reliability-based design optimization of problems with correlated input variables using a gaussian copula. *Struct. Multidiscip. Optim.*, 38(1):1–16.

Notani, M. A., Moghadas Nejad, F., Khodaii, A., and Hajikarimi, P. (2019). Evaluating fatigue resistance of toner-modified asphalt binders using the linear amplitude sweep test. *Road Materials and Pavement Design*, 20(8):1927–1940.

Paya-Zaforteza, I., Yepes, V., González-Vidosa, F., and Hospitaler, A. (2010). On the weibull cost estimation of building frames designed by simulated annealing. *Meccanica*, 45(5):693–704.

Pellinen, T. and Witczak, M. (2002). Use of Stiffness of Hot-Mix Asphalt as a Simple Performance Test. *Transp. Res. Rec. : J. Transp. Res. Board*, 1789(02):80–90.

Peng, C., Feng, J., Feiting, S., Changjun, Z., and Decheng, F. (2019). Modified two-phase micromechanical model and generalized self-consistent model for predicting dynamic modulus of asphalt concrete. *Constr. Build. Mater.*, 201:33–41.

Rahami, H., Kaveh, A., Aslani, M., and Asl, R. N. (2011a). A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization. *Int. J. Optim. Civ. Eng.*, 1(January):29–46.

Rahami, H., Kaveh, A., Aslani, M., and Najian Asl, R. (2011b). A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization. *Iran University of Science & Technology*, 1(1):29–46.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer.

Ren, R., Han, K., Zhao, P., Shi, J., Zhao, L., Gao, D., Zhang, Z., and Yang, Z. (2019). Identification of asphalt fingerprints based on atr-ftir spectroscopy and principal component-linear discriminant analysis. *Constr. Build. Mater.*, 198:662–668.

Rodezno, M., Kaloush, K., and Corrigan, M. (2010). Development of a Flow Number Predictive Model. *Transp. Res. Rec. : J. Transp. Res. Board*, 2181(2181):79–87.

Roesler, J. R. and Barenberg, E. J. (1999). Fatigue and static testing of concrete slabs. *Transportation Research Record*, 1684(1):71–80.

Roesler, J. R., Hiller, J. E., and Littleton, P. C. (2005). Large-scale airfield concrete slab fatigue tests. In *8th International Conference on Concrete Pavements: Innovations for Concrete Pavement: Technology Transfer for the Next Generation*, pages 1247–1268.

Rollins, D. K., Roggendorf, A., Khor, Y., Mei, Y., Lee, P., and Loveland, S. (2015). Dynamic Modeling With Correlated Inputs: Theory, Method, and Experimental Demonstration. *Industrial & Engineering Chemistry Research*, 54(7):2136–2144.

Rollins, D. K., Zhai, D., Joe, A. L., Guidarelli, J. W., Murarka, A., and Gonzalez, R. (2006). A novel data mining method to identify assay-specific signatures in functional genomic studies. *BMC bioinformatics*, 7:377.

Rueda, J. L. and Erlich, I. (2015). Testing MVMO on learning-based real-parameter single objective benchmark optimization problems. *2015 IEEE Congr. Evol. Comput. CEC 2015 - Proc.*, pages 1025–1032.

Sadeghi, V. and Hesami, S. (2018). Investigation of load transfer efficiency in jointed plain concrete pavements (jpcp) using fem. *International Journal of Pavement Research and Technology*, 11(3):245–252.

Sakhaeifar, M. S., Richard Kim, Y., and Garcia Montano, B. E. (2017). Individual temperature based models for nondestructive evaluation of complex moduli in asphalt concrete. *Constr. Build. Mater.*, 137:117–127.

Sakhaeifar, M. S., Richard Kim, Y., and Kabir, P. (2015). New predictive models for the dynamic modulus of hot mix asphalt. *Constr. Build. Mater.*, 76:221–231.

Saltan, M. and Sezgin, H. (2007). Hybrid neural network and finite element modeling of sub-base layer material properties in flexible pavements. *Mater. Des.*, 28(5):1725–1730.

Sanabria, N., Valentin, V., Bogus, S., Zhang, G., and Kalhor, E. (2017). Comparing neural networks and ordered probit models for forecasting pavement condition in new mexico. Technical report.

Shoukry, S. N., Fahmy, M., Prucz, J., and William, G. (2007). Validation of 3dfe analysis of rigid pavement dynamic response to moving traffic and nonlinear temperature gradient effects. *International Journal of Geomechanics*, 7(1):16–24.

Shu, X. and Huang, B. (2008). Dynamic modulus prediction of hma mixtures based on the viscoelastic micromechanical model. *J. Mater. Civ. Eng.*, 20(8):530–538.

Sobol', I. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(1):112–118.

Sousa, J. B., Craus, J., and Monismith, C. L. (1991). Summary report on permanent deformation in asphalt concrete. Technical report.

Su, C.-T. and Tong, L.-I. (1997). Multi-response robust design by principal component analysis. *Total Qual. Manag.*, 8(6):409–416.

Sun, P. and Freund, R. M. (2004). Computation of minimum-volume covering ellipsoids. *Operations Research*, 52(5):690–706.

Tabatabaie, A. M. and Barenberg, E. J. (1978). Finite-element analysis of jointed or cracked concrete pavements. *Transportation Research Record*, (671).

Taerwe, L., Matthys, S., et al. (2013). Fib model code for concrete structures 2010.

Talatahari, S., Gandomi, A. H., Yang, X.-S., and Deb, S. (2015). Optimum design of frame structures using the eagle strategy with differential evolution. *Engineering Structures*, 91:16–25.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association.

Timm, D., West, R., Priest, A., Powell, B., Selvaraj, I., Zhang, J., and Brown, R. (2006). Phase ii ncat test track results. *NCAT Report*, 6(05).

Todd, M. J. and Yıldırım, E. A. (2007). On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids. *Discret. Appl. Math.*, 155(13):1731–1744.

Uddin, W., Hackett, R. M., Joseph, A., Pan, Z., and Crawley, A. B. (1995). Three-dimensional finite-element analysis of jointed concrete pavement with discontinuities. *Transportation Research Record*, 1482:26–32.

Wei, P., Lu, Z., and Song, J. (2015). Variable importance analysis: a comprehensive review. *Reliability Engineering & System Safety*, 142:399–432.

Witczak, M. W. (2002). *Simple performance test for superpave mix design*, volume 465. Transportation Research Board.

Witczak, M. W. (2007). *Specification criteria for simple performance tests for rutting*, volume 1. Transportation Research Board.

Witzcak, M. W. (2002). *Simple performance test for superpave mix design*, volume 465. Transportation Research Board.

Xiong, F., Wang, D., Zhang, S., Cai, K., Wang, S., and Lu, F. (2017). Lightweight optimization of the side structure of automobile body using combined grey relational and principal component analysis. *Struct. Multidiscip. Optim.*, 57:441–461.

Xiong, F., Wang, D., Zhang, S., Cai, K., Wang, S., and Lu, F. (2018). Lightweight optimization of the side structure of automobile body using combined grey relational and principal component analysis. *Struct. Multidiscip. Optim.*, 57(1):441–461.

Yi, S., Lai, Z., He, Z., Cheung, Y.-m., and Liu, Y. (2017). Joint sparse principal component analysis. *Pattern Recognition*, 61:524–536.

Zhang, J., Yang, J., and Kim, Y. R. (2015). Characterization of mechanical behavior of asphalt mixtures under partial triaxial compression test. *Constr. Build. Mater.*, 79:136–144.

Ziari, H., Amini, A., Goli, A., and Mirzaiyan, D. (2018). Predicting rutting performance of carbon nano tube (cnt) asphalt binders using regression models and neural networks. *Construction and Building Materials*, 160:415–426.

Ziari, H., Nakhaei, M., Akbari Nasrekani, A., and Moniri, A. (2016). Characterization of rutting resistance of ebs-modified asphalt mixtures. *Petroleum Science and Technology*, 34(13):1107–1112.

Zitzler, E. and Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.*, 3(4):257–271.