



Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks

Parnian Ghasemi^{a,*}, Mohamad Aslani^b, Derrick K. Rollins^c, R. Christopher Williams^d,
Vernon R. Schaefer^d

^a Iowa State University, Civil Construction and Environmental Engineering Department, 427 Town Engineering Building, Ames, IA 50011, United States

^b Florida State University, Department of Mathematics, Tallahassee, FL 32304, United States

^c Iowa State University, Department of Chemical and Biological Engineering, 1033 Sweeney Hall, Ames, IA 50011, United States

^d Iowa State University, Civil Construction and Environmental Engineering Department, 488 Town Engineering Building, Ames, IA 50011, United States

Received 9 October 2017; received in revised form 23 December 2017; accepted 16 January 2018

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.

© 2018 Chinese Society of Pavement Engineering. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Principal component analysis; Data mining; Machine learning; Artificial neural networks; Rutting; Flow number; Asphalt mixture

1. 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 [1]. 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

* Corresponding author.

E-mail addresses: pghasemi@iastate.edu (P. Ghasemi), maslani@iastate.edu (M. Aslani), drollins@iastate.edu (D.K. Rollins), rwilliam@iastate.edu (R. Christopher Williams), vern@iastate.edu (V.R. Schaefer).
Peer review under responsibility of Chinese Society of Pavement Engineering.

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 [1,2]. 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 [3]. 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 [4]. 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 [5]. The FN has indicated a good correlation with field rutting of asphalt mixtures exposed to different levels of traffic [6].

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 [7]. Asphalt mixture properties, which affect rutting behavior (simulated by FN), were identified more precisely by Kvasnak et al. [8]. 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. [9] 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 [10,11,12,13], Apeageyi [14] 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 [15]. 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 [16]. 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 Fig. 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 [17].

Although ANN can be trained to approximate a non-linear, complicated relationship [18,19,20,21], 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 [22]. 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 [23]. 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 [24,25].

The present study focuses on developing a machine learning-based framework to reduce a large set of corre-

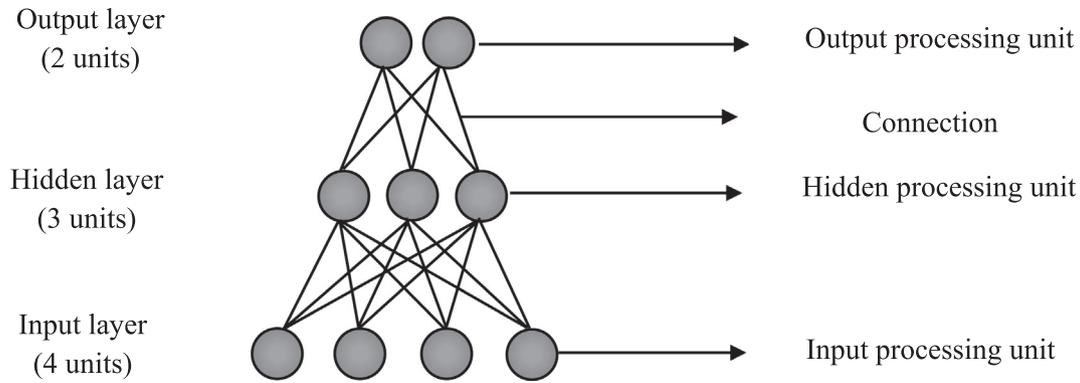


Fig. 1. Example of an architecture for ANN.

lated 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. Section 3 covers the pre-processing step for input variables evaluation. PCA is described in detail in Section 4. The proposed modeling methods, PCR and PCNN, are presented in Sections 5, followed by results, discussion, and conclusions in Sections 6 and 7 respectively.

2. Experimental materials and methods

2.1. 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 [26] and D3665 [27]. 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 [28] 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 Gyrotory 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 [29].

2.2. 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 [5] the test was conducted from higher to lower frequencies to mitigate the amount of deformation that is induced upon specimens.

2.3. 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 s and a dwell period of 0.9 s. 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.

2.4. 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 [30].

3. Pre-processing step: input variable 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 [31].

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 $\eta_i = f(X_i; \theta)$, where η_i is the expected value of the response (i.e., output) at the i^{th} sampling time, $i = 1, \dots, n$; X_i is the vector of input values at the i^{th} sampling time; and θ is the vector of unknown model parameters with $\theta = [\theta_1 \dots \theta_q]^T$. Let the element of its Jacobian Matrix, $J^{n \times q}$, in the i^{th} row and j^{th} column be $\frac{\partial \eta_i}{\partial \theta_j}$ i.e., $J = \left\{ \frac{\partial \eta_i}{\partial \theta_j} \right\}$. 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 response space, this work seeks to minimize pairwise correlation of the inputs that are used to model the response behavior.

According to the literature [7,9], 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 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 [32], the pairwise correlation matrix of input variables is calculated and presented by Table 2. Results with absolute value of 0.5 and higher are in bold text.

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 [31].

4. 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

Table 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*	210,801	1,163,560	612,180	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		✓

Table 2
Correlation matrix for the input variables. Results with absolute value of 0.5 and higher are in bold text.

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17
x1	1	0.33	-0.73	0.60	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
x2	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
x3	-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
x4	0.60	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
x5	0.71	-0.03	-0.76	0.88	1	0.95	0.77	0.71	0.61	0.49	0.43	0.39	0.36	0.56	-0.33	0.44	-0.51
x6	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
x7	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
x8	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
x9	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
x10	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
x11	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
x12	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
x13	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
x14	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
x15	-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
x16	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
x17	-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

that axis is also small, and by omitting that axis 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 [24]. A schematic of this transformation for three inputs is presented in Fig. 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 X , PCA reduces the dimension of X by expressing the p ($p = 17$) original variables (x_1, \dots, 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 [34]. PCA can be considered as a data pre-processing methodology that determines an optimal rotational transformation of the dataset, X , and maximizes the amount of variance of the output γ that is explained by the PCs [31].

Considering the given dataset X , PCA is performed in the following steps:

1. Standardizing X by transforming it to Z using the following equations.

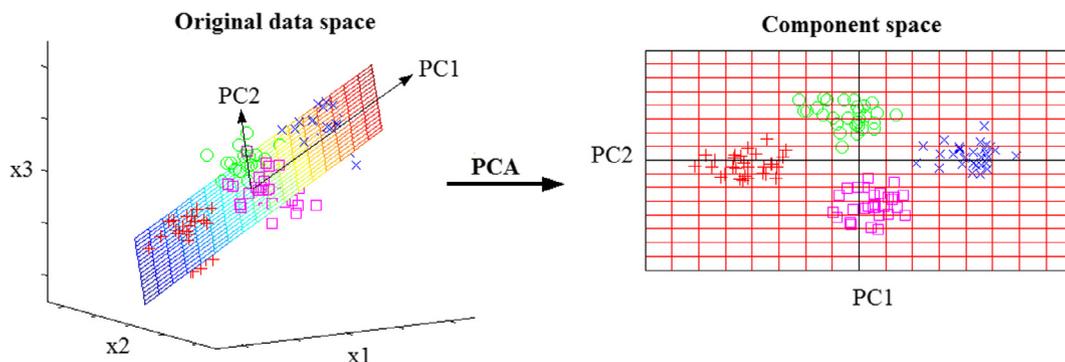


Fig. 2. Schematic of the PCA transformation taken from Scholz 2006 [33]. Original data space presented on the left with 3 input variables transformed to a 2-dimensional component space with PC1 and PC2 as the axes of the coordinate.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (1)$$

$$Z = [z_1 \quad z_2 \quad \dots \quad z_p] = \begin{bmatrix} \frac{x_{11}-\bar{x}_1}{s_1} & \frac{x_{12}-\bar{x}_2}{s_2} & \dots & \frac{x_{1p}-\bar{x}_p}{s_p} \\ \frac{x_{21}-\bar{x}_1}{s_1} & \frac{x_{22}-\bar{x}_2}{s_2} & \dots & \frac{x_{2p}-\bar{x}_p}{s_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x_{n1}-\bar{x}_1}{s_1} & \frac{x_{n2}-\bar{x}_2}{s_2} & \dots & \frac{x_{np}-\bar{x}_p}{s_p} \end{bmatrix} \quad (2)$$

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

2. Determine the unit eigenvectors e_1, \dots, e_p of Z .
3. Determine the corresponding eigenvalues $\lambda_1, \dots, \lambda_p$.
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^{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. [25].

Using the JMP statistical software package [32], the eigenvectors and eigenvalues of the correlation matrix are calculated and presented in Tables 3 and 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 Fig. 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. (3):

$$PC_i = \sum_{j=1}^{17} \alpha_{ij} X_j + \beta_i \quad (3)$$

Table 3
Eigenvectors of the Z matrix.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17
z1	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
z2	-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
z3	-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
z4	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
z5	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
z6	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
z7	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
z8	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
z9	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
z10	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
z11	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
z12	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
z13	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
z14	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
z15	-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
z16	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
z17	-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 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

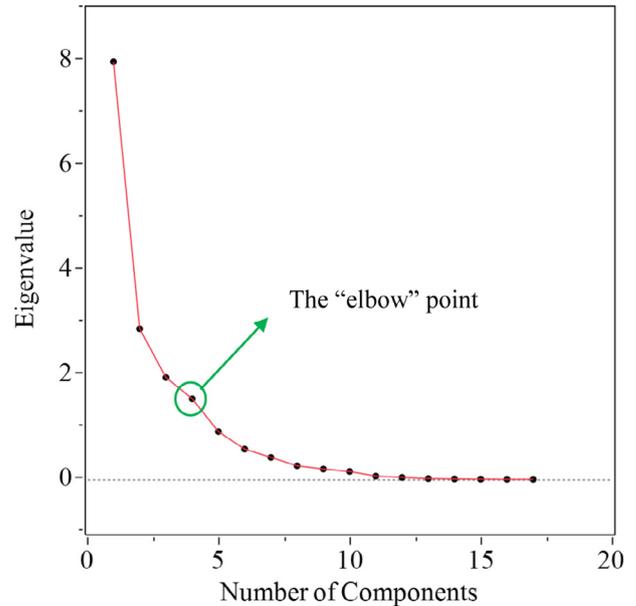


Fig. 3. Scree plot.

where, α is the corresponding coefficient, β is constant, and X_j 's are the original input variables.

Eq. (3) is presented in matrix notation by Eq. (4):

$$PC = AZ^T + B \tag{4}$$

where,

$$A^T = \begin{bmatrix} 3.41E-01 & -2.04E-01 & 2.89E-01 & -1.94E-01 & 3.71E-01 \\ -7.60E-08 & -1.17E-06 & 1.31E-06 & 1.05E-07 & 2.47E-06 \\ -6.59E-02 & 3.22E-02 & -4.53E-02 & 9.10E-02 & -7.97E-03 \\ 6.34E-02 & -1.31E-02 & 1.86E-02 & -9.38E-02 & -7.13E-03 \\ 2.11E-02 & -2.33E-03 & 5.26E-03 & -1.63E-02 & -9.03E-03 \\ 2.17E-02 & 7.23E-03 & 1.55E-03 & -1.23E-02 & -1.13E-02 \\ 2.13E-02 & 2.01E-02 & -2.82E-03 & -7.27E-03 & -7.11E-03 \\ 2.79E-02 & 2.89E-02 & -2.98E-03 & 3.79E-03 & 1.41E-03 \\ 3.28E-02 & 3.22E-02 & 4.43E-03 & 2.35E-02 & 1.71E-02 \\ 3.73E-02 & 3.39E-02 & 1.53E-02 & 4.98E-02 & 3.14E-02 \\ 7.08E-02 & 3.30E-02 & 7.38E-02 & 1.54E-01 & -9.05E-03 \\ 1.08E-01 & -2.03E-01 & 2.33E-01 & 2.30E-01 & -2.72E-01 \\ 9.18E-02 & -4.06E-01 & 2.29E-01 & 6.69E-02 & -3.18E-01 \\ 9.98E-02 & -8.77E-02 & -1.12E-01 & -1.14E-02 & 8.85E-02 \\ -1.73E-02 & 3.21E-02 & 4.57E-02 & -2.69E-02 & -9.43E-04 \\ 9.75E-02 & -1.42E-01 & -1.95E-01 & 7.31E-02 & 4.70E-02 \\ -5.50E-04 & 2.73E-04 & 8.20E-04 & -8.92E-06 & 1.36E-04 \end{bmatrix}$$

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

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

5. Proposed modeling approaches

5.1. 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 [35]. 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. 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.

5.2. 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 they did not significantly improve fit to the training data and gave worse fit to the test data.

5.3. 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 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 program [36]. 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.

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 conver-

gence is obtained). The Bayesian Regularization algorithm is implemented for the training efficiency of the network.

6. Results and discussion

6.1. 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. (5) as \hat{y} .

$$\hat{y} = c_0 + c_1 * PC_1 + c_2 * PC_2 + c_3 * PC_3 + c_4 * PC_4 + c_5 * PC_5 + c_6 * PC_1 * PC_2 + c_7 * PC_2 * PC_4 + c_8 * PC_1 * PC_3 + c_9 * PC_2 * PC_3 + c_{10} * PC_3 * PC_5 \tag{5}$$

where, *c_i* for *i* = 0, . . . , 10 are presented in Table 5:

The successfully trained ANN models can be presented by Eq. (6) for ease of use and wider reproduction. Each ANN is presented by the connection weights and biases in a three-layer topology [37].

$$\hat{y} = f_2 \left\{ B_0 + \sum_{j=1}^n [W_j \cdot f_1(B_{Hj} + \sum_{i=1}^m W_{ij} P_i)] \right\} \tag{6}$$

where, *B₀* = 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*; *f₁*(*t*) = transfer function of the hidden layer, and *f₂*(*t*) = transfer function of the output layer.

Both transfer functions *f₁*(*t*) and *f₂*(*t*) are sigmoid functions defined in Eq. (7) [37].

$$f_k(t) = \frac{1}{1 + e^{-t}} \text{ for } k = 1, 2 \tag{7}$$

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

Table 5
Linear Regression Equation's Coefficients.

Coefficient	value
<i>c</i> ₀	1.64 e + 4
<i>c</i> ₁	-8.89 e + 2
<i>c</i> ₂	-1.24 e + 3
<i>c</i> ₃	-1.24 e + 3
<i>c</i> ₄	-92.41
<i>c</i> ₅	6.55 e + 2
<i>c</i> ₆	1.58 e + 2
<i>c</i> ₇	-4.35 e + 2
<i>c</i> ₈	2.35 e + 2
<i>c</i> ₉	3.74 e + 2
<i>c</i> ₁₀	-8.28 e + 2

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

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

Performance results of the PCR and PCNN models are given based on the following statistics, and presented in Table 5. The first statistic is the “average error (AE)” defined as

$$AE = 1/n \sum_{i=1}^n (y_i - \hat{y}_i) \tag{8}$$

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 statistic is the “average absolute error (AAE)” and defined as

$$AAE = 1/n \sum_{i=1}^n |y_i - \hat{y}_i| \tag{9}$$

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 \hat{y}_i and defined as

$$r_{fit} = \frac{n \sum_{i=1}^n y_i \hat{y}_i - (\sum_{i=1}^n y_i)(\sum_{i=1}^n \hat{y}_i)}{\sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2} \sqrt{n \sum_{i=1}^n \hat{y}_i^2 - (\sum_{i=1}^n \hat{y}_i)^2}} \tag{10}$$

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.

According to the values of r_{fit} presented in Table 6, the predicted values of accumulated strain by PCR and PCNN models have a high correlation 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. (9) below:

$$\frac{\epsilon_p}{\epsilon_r} = aT^bN^c \tag{11}$$

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 [2]. 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 [2].

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.

Table 6
Statistical analysis of PCR and PCNN modeling. Highest values are in bold text.

Subset	Statistics	PCR			PCNN		
		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	0.80	0.73	0.97	0.95	0.92
	R^2	na*	na*	na*	na*	na*	na*

* Not applicable.

6.2. Model validation

Equation (10) is a general regression model for this study:

$$y_i = f(Z_i, \theta) + \varepsilon_i \tag{12}$$

where f is the expectation function, θ is the vector of parameters and ε_i is random error term assumed to be

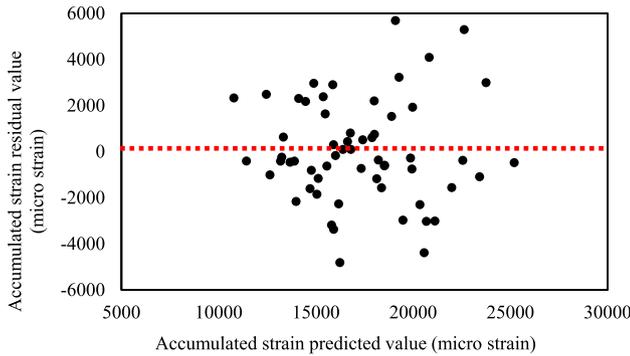


Fig. 4a. Plot of the residuals for best fold of PCR model.

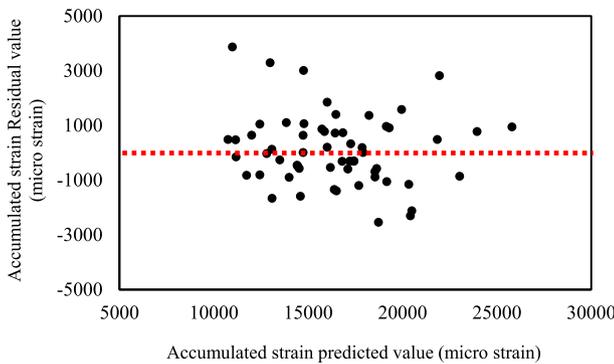


Fig. 4b. Plot of the residuals for best fold of PCNN model.

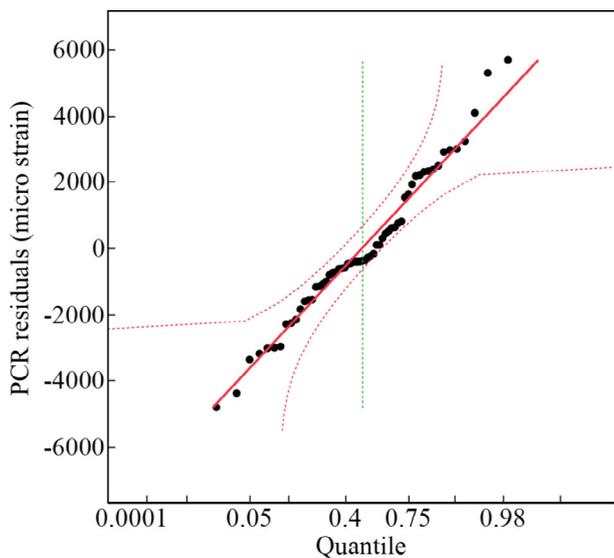


Fig. 5a. Normal probability plot of the residuals for best fold of PCR model.

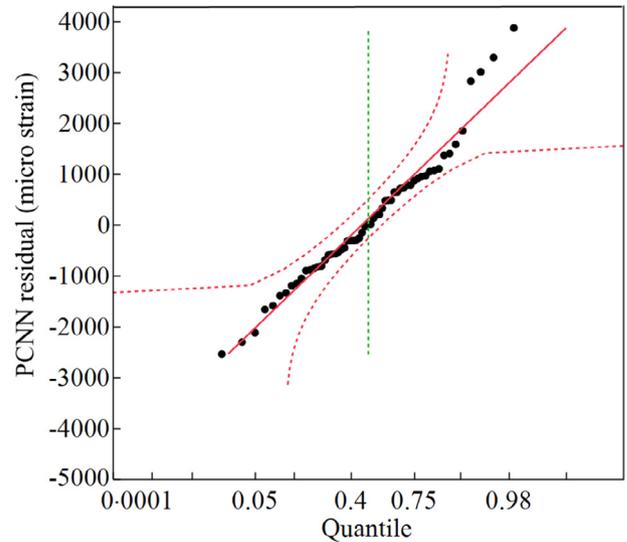


Fig. 5b. Normal probability plot of the residuals for best fold of PCNN model.

normally distributed with mean zero and unknown variance σ^2 for $i = 1, \dots, n$, 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. (11):

$$e_i = y_i - \hat{y}_i \tag{13}$$

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 [38]. These assumptions were checked for the best fold in each method. The plots of residuals for the best folds are presented in Fig. 4. 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 Fig. 5. The data points are not too far away from a straight line. Therefore, the normality assumption does not appear to be violated.

7. 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, these 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 input on the response. Considering the obtained regression model (Eq. (5)), 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 3. The one with the largest absolute value is associated variable contributes the most to that pseudo-variable.

Acknowledgements

The data for this work were provided by the Wisconsin Department of Transportation and the Wisconsin Highway Research Program under project number 0092-04-07. The assistance from these organizations and Christopher J. Robinette, Jason Bausano, Tamer Breakah, and Dr. Joseph Podolsky is highly appreciated.

References

- [1] J.B. Sousa, J. Craus, C.L. Monismith, Summary Report SHRP-A/IR-91-104: Permanent Deformation in Asphalt Concrete, Transportation Research Board, Washington, D.C., 1991.
- [2] M.W. Witczak, M.M. El-Basyouny, NCHRP Report 1-37A- Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures. Appendix A: Calibration of Permanent Deformation Models For Flexible Pavements, Transportation Research Board, Washington, DC., 2004.
- [3] A. Bashin, E. Masad, M.E. Kutay, W. Buttlar, Y.-R. Kim, M. Marasteanu, Y.R. Kim, C.W. Schwartz, R.L. Carvalho, Applications of Advanced Models to Understand Behavior and Performance of Asphalt Mixtures, Transportation Research Board, Washington, DC, 2012.
- [4] J. Zhang, J. Yang, Y.R. Kim, Characterization of mechanical behavior of asphalt mixtures under partial triaxial compression test, Construct. Build. Mater. 79 (2015) 136–144, <https://doi.org/10.1016/j.conbuildmat.2014.12.085>.
- [5] AASHTO TP 79-13. AASHTO TP 79-13 Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Asphalt Mixtures Performance Tester. American Association of State Highway and Transportation Officials, 2013, pp. 1–19.
- [6] M.W. Witczak, NCHRP Report 580: Specification Criteria for Simple Performance Tests for Rutting, Volume I: Dynamic Modulus (E*) and Volume II: Flow Number and Flow Time, Transportation Research Board, Washington, DC., 2007.
- [7] K.E. Kaloush, M.W. Witczak, B.W. Sullivan, Simple performance test for permanent deformation evaluation of asphalt mixtures, in: Sixth International RILEM Symposium on Performance Testing and Evaluation of Bituminous Materials, 2003, pp. 498–505, <https://doi.org/10.1617/2912143772.062>.
- [8] A. Kvasnak, C. Robinette, R.C. Williams, A Statistical Development of a Flow Number Predictive Equation for the Mechanistic-Empirical Pavement Design Guide.CD-ROM. No. 3, D. C., Transportation Research Board, Washington, 2007, pp 1–18.
- [9] M. Rodezno, K. Kaloush, M. Corrigan, Development of a flow number predictive model, Transport. Res. Rec. 2181 (2181) (2010) 79–87, <https://doi.org/10.3141/2181-09>.
- [10] B. Birgisson, R. Roque, J. Kim, L.V. Pham, The Use of Complex Modulus To Characterize Performance of Asphalt Mixtures and Pavements in Florida, Florida Department of Transportation, 2004.
- [11] T. Pellinen, M. Witczak, Use of stiffness of hot-mix asphalt as a simple performance test, Transport. Res. Rec. 1789 (2) (2002) 80–90, <https://doi.org/10.3141/1789-09>.
- [12] E.R. Brown, L.A. Cooley Jr., D. Hanson, C. Lynn, B. Powell, B. Prowell, D. Watson, NCAT Report 02-12: NCAT Test Track Design, Construction, and Performance, 2002.
- [13] D. Timm, I. Selvaraj, R. Brown, R.C. West, A. Priest, NCAT Report 06-05: PHASE II NCAT TEST TRACK Results, 2006.
- [14] A.K. Apeageyi, Rutting as a function of dynamic modulus and gradation, Am. Soc. Civil Eng. 23 (9) (2011) 1302–1310, [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0000309](https://doi.org/10.1061/(ASCE)MT.1943-5533.0000309).
- [15] B. Cheng, D. Titterton. Neural networks: a review from a statistical perspective. Stat. Sci. 9(1), 2–54.
- [16] J. Devore, Probability and Statistics for Engineering and the Sciences, 2012.
- [17] N. Kartam, Neural networks in civil engineering: systems and application, J. Comput. Civil Eng. 8 (2) (1994) 149–162.
- [18] Nabil Kartam, I. Flood, Neural networks in civil engineering: principal and understanding, J. Comput. Civil Eng. 8 (2) (1994) 131–148.
- [19] H. Rahami, A. Kaveh, M. Aslani, R.N. Asl, A hybrid modified genetic-nelder mead simplex algorithm for large-scale truss optimization, Int. J. Opt. Civil Eng. 1 (2011) 29–46.
- [20] M. Aslani, R. Asla, A novel hybrid simplex-genetic algorithm for the optimum design of truss structures, Lecture Notes Comput. Sci. II (2010).
- [21] M. Saltan, H. Sezgin, Hybrid neural network and finite element modeling of sub-base layer material properties in flexible pavements, Mater Design 28 (5) (2007) 1725–1730, <https://doi.org/10.1016/j.matdes.2006.02.017>.
- [22] R. May, G. Dandy, H. Maier, Review of input variable selection methods for artificial neural networks, Artif. Neural Netw. Method. Adv. Biomed. Appl. (2011) 19–44, <https://doi.org/10.5772/644>.
- [23] D.K. Rollins, A. Roggendorf, Y. Khor, Y. Mei, P. Lee, S. Loveland, Dynamic modeling with correlated inputs: theory, method, and experimental demonstration, Indust. Eng. Chem. Res. 54 (7) (2015) 2136–2144, <https://doi.org/10.1021/ie5045956>.
- [24] I.T. Jolliffe, Principal component analysis, Encycl. Stat. Behav. Sci. 30 (3) (2002) 487, <https://doi.org/10.2307/1270093>.
- [25] D.K. Rollins, D. Zhai, A.L. Joe, J.W. Guidarelli, A. Murarka, R. Gonzalez, A novel data mining method to identify assay-specific signatures in functional genomic studies, BMC Bioinform. 7 (2006) 377, <https://doi.org/10.1186/1471-2105-7-377>.
- [26] Standard Practice for Sampling Bituminous Paving Mixtures, Am. Soc. Testing Mater. (2016) 1–4, <https://doi.org/10.1520/D0979>.
- [27] Standard Practice for Random Sampling of Construction Materials, Am. Soc. Testing Mater. (2016) 1–13, <https://doi.org/10.1520/D3665-12.2>.
- [28] Standard test method for maximum specific gravity and density of bituminous paving mixtures, Am. Soc. Testing Mater., 14(C) 2003, pp. 2–6. <https://doi.org/10.1520/D6857-09.2>.

- [29] Standard test method for bulk specific gravity and density of non-absorptive compacted bituminous mixtures. Am. Soc. Testing Mater., i, 2005, pp. 5–8. <https://doi.org/10.1520/D2726>.
- [30] Standard Test Method for Determining the Complex Shear Modulus (G^*) Of Bituminous Mixtures Using Dynamic Shear Rheometer. Am. Soc. Testing Mater., No. Reapproved 2014, 2016, pp. 1–11. <https://doi.org/10.1520/D7552-09R14.2>.
- [31] I.K. Fodor, *A Survey of Dimension Reduction Techniques*, U.S. Department of Energy, 2002.
- [32] JMP, Pro 12. SAS Institute Inc., Cary, NC, 1989–2007.
- [33] M. Scholz, *Approaches to Analyse and Interpret Biological Profile Data*, 2006.
- [34] R.A. Johnson, D.W. Wichern. *Applied Multivariate Statistical Analysis*, New Jersey, 1992.
- [35] P. Refaeilzadeh, L. Tang, H. Liu, Cross-validation, *Encycl. Database Syst.* (2009) 532–538, https://doi.org/10.1007/978-0-387-39940-9_565.
- [36] MATLAB. Matheworks, T., 2012. Matlab 2012b.
- [37] C.H. Juang, C.J. Chen, CPT-based liquefaction evaluation using artificial neural networks, *Comput Aided Civil Infrastruct. Eng.* 14 (3) (1999) 221–229, <https://doi.org/10.1111/0885-9507.00143>.
- [38] D.C. Montgomery, *Design and Analysis of Experiments*, Tempe, Arizona, 2012.