Data Analysis in Pavement Engineering: An Overview

Qiao Dong, *Member, IEEE*, Xueqin Chen, Shi Dong, and Fujian Ni

Abstract—**Extensive studies on data analysis have been conducted to address pavement engineering problems including material and structure design, performance evaluation, maintenance, and preservation. This paper summarized and discussed more than 40 types of data analysis methods including statistical tests, experimental design, regressions, count data model, survival analysis, stochastic process models, supervised learnings, unsupervised learnings, reinforcement learnings, and Bayesian analysis applied in pavement engineering. Generally, traditional statistical regression models are proper for significant factors quantification and pavement performance predictions with explicit model equations and meanings of parameters. The supervised machine learnings are powerful in prediction, dealing with large data volume or unstructured data such as pavement distress images, sounds, and other unprocessed signals. The unsupervised machine learnings are usually used to pre-process data by reducing the dimensionality, extracting common factors of variables, and clustering the data samples. Selecting proper models and their combinations will be the key for the increasing accumulation of historical pavement performance data, as well as the big data from automatic pavement evaluations and pavement instrumentation in future practices and studies.**

Index Terms—**Pavement, data analysis, machine learning, unsupervised learning, supervised learning.**

I. Introduction

DATA analyses have been used in pavement material design, structure design, and maintenance planning since the beginning of modern pavement engineering. Data in pavement engineering are available from laboratory material or structure tests, numerical simulations of pavement mechanics, field pavement performance and distress evaluations, and the Pavement Management Systems (PMS). Pavement data can be classified into structured data and unstructured data. The structured data, which mainly include material test results, historical pavement performance data, etc. can be displayed in tables and relational databases. The unstructured data such as the pavement distress images need specific feature extraction or signal process to interpret.

Manuscript received 14 May 2021; revised 18 July 2021 and 14 August 2021; accepted 23 September 2021. Date of publication 26 October 2021; date of current version 7 November 2022. This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20200468, in part by the National Natural Science Foundation of China under Grant 51978163, in part by the Fundamental Research Funds for the Central Universities, in part by Chang'an University under Grant 300102341508, and in part by the Science and Technology Project of Zhejiang Provincial Department of Transport under Grant 2020045 and Grant 2020053. The Associate Editor for this article was X. Luo. *(Corresponding author: Qiao Dong.)*

Qiao Dong and Fujian Ni are with the School of Transportation, National Demonstration Center for Experimental Road and Traffic Engineering Education, Southeast University, Nanjing 211189, China (e-mail: qiaodong@seu.edu.cn).

Xueqin Chen is with the Department of Civil Engineering, Nanjing University of Science and Technology, Nanjing 210094, China.

Shi Dong is with the China Engineering Research Center of Highway Infrastructure Digitalization, College of Transportation Engineering, Chang'an University, Xi'an 710064, China.

Digital Object Identifier 10.1109/TITS.2021.3115792

As the completion of road network construction, pavement evaluation and preservation have become the focus of pavement engineering and raised more needs for data analysis [1]. Many statistical tests, regression, machine learning, and artificial intelligence methods and algorithms have been adopted to identify significant variables, determine optimal designs, quantify influencing factors, extract key features, evaluate performance, and predict future deterioration. In addition, emerging techniques including pavement instrumentation, crowdsourcing monitoring, cloud calculation, and the internet of things will add a huge amount of data into pavement engineering [2], [4]. For example, various types of sensors including electronic sensors [5], optic fiber sensors [6], [7], distributed fiber optic sensors [8], self-powered wireless sensors [9], time-domain reflectometry [10], vibration sensors, etc. have been installed for full scale accelerated loading tests or in-situ pavement structural health monitoring [11]. Those sensor data are extracted and fused to either directly evaluate the internal static or dynamic responses of pavement structure or to include external environmental conditions for pavement performance evaluation and prediction. The data collection, transmission, fusion, cleaning, mining, and training will be the keys to the "smart pavement" of the next era. However, as more resourceful as those data are, as many more challenges remain to be realized for pavement researchers. This review article summarizes current applications and achievements of data analysis in pavement engineering. As shown in Fig. 1, in addition to the traditional statistical tests and design of experiments, the majority of data analysis methods are the supervised learnings with labeled data including various statistical regression models and the neural networks, SVM etc., followed by unsupervised learning with unlabeled data and reinforcement learning which only has one reported study.

II. Pavement Performance Indices

Most studies on data analysis in pavement engineering are on the data from pavement performance modeling, followed by pavement nondestructive tests, pavement material tests, and numerical simulations. The data for pavement performance modeling include pavement condition data as well as related traffic, structure, material, and climatic data. Pavement condition data include pavement functional, structural, and distress conditions. Usually, an overall pavement performance index is calculated based on multiple pavement condition indicators.

1558-0016 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

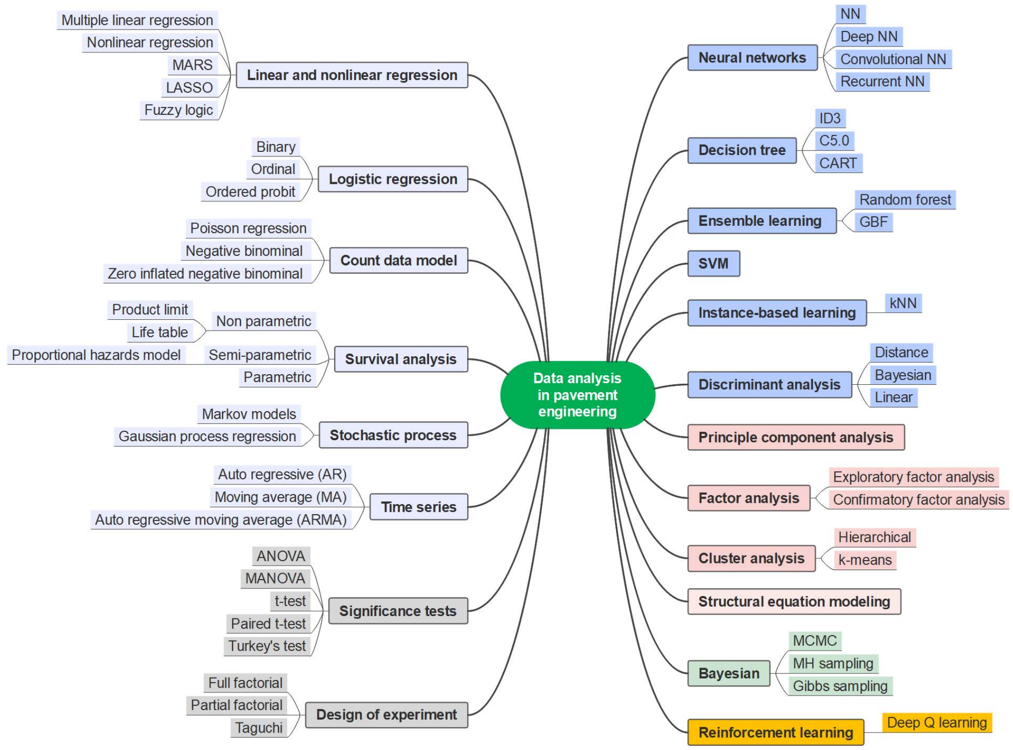


Fig. 1. Summary of data analysis methods in pavement engineering.

In the 1950s, the American Association of State Highways Officials (AASHO) developed the first overall pavement performance index, the empirical Pavement Serviceability Rating (PSR) using a 1-5 rating scale and then the Pavement Serviceability Index (PSI). As shown in Equation (1), PSI is a regression function of roughness, cracking length, patching area, and rutting depth [12], [13]. The coefficients in those regressions have been modified to enhance the effectiveness of the PSI [14], [15].

Regarding pavement distresses, the United States Army Corps of Engineers (USACE) developed the first Pavement Condition Index (PCI) using a 1-100 rating scale. As shown in Equation (2), PCI equals 100 minus the cumulative deduct value calculated based on the severity levels and extent of different distresses [16]. The weights for calculating the deduct values are mainly determined based on experience. Many studies have been conducted to modify coefficients [17], [22]. Recently, fuzzy logic was adopted to determine the coefficients [19], [20], [23], [26]. Based on the PSI and PCI, highway agencies developed various pavement performance indices, including the Distress Score (DS) and Condition Score (CS) used by Texas, the Pavement Quality Index (PQI) used in China, the Maintenance Condition Index (MCI) used by Japan [27]

$$PSI = 5.03 - 1.9\log(1+SV) - 0.01\sqrt{C+P} - 1.38RD^2 \quad (1)$$

$$PCI = 100 - CDV \quad (2)$$

III. Design of Experiment and Significance Tests

For laboratory material test data, the Design of Experiment (DOE) and significance tests are the basic data analysis techniques. These techniques could also be used for pavement performance data analysis using the data from PMSs of highway agencies. It is noted that the Long-Term Pavement Program (LTPP) database is extensively used in many pavement performance data analysis studies. The LTPP has been monitoring more than 2400 pavement sections in North American since 1987 and started reporting valuable findings in the 1990s [28], [29].

The DOE methods including partial and full factorial design have been adopted for experiment planning to analyze the effects of factors and levels with a limited number of experiments [30], [31]. Taguchi method also called the robust design method or orthogonal design is a type of partial factorial design with a minimum number of experiments. It could use 16 experiments to analyze the effects of 6 factors and 4 levels for the mixture's shear stiffness [32], or 25 experiments for 5 factors and 5 levels for pavement stress intensity [32], [33].

Based on test results or field observations, significance tests including t-test, paired t-test, Turkey's test, etc. have been widely adopted to test the difference between groups or pairs. To identify key factors for material properties and pavement performances, the Analysis of Variance (ANOVA) was usually adopted to examine the significance of a predictor on a target. ANOVA and t-test are usually used with linear regression to identify significant influencing factors and interactions. ANOVA has been used to analyze the effects of materials type, temperature, pavement structure, traffic, and pavement surface texture on the shear stress in asphalt mixture [34], initial shear stress in a mixture [35], compound strain rate [36], [37], dynamic modulus of asphalt mixture, pavement modulus and deformation [38], the density of roller-compacted concrete pavement [39], pavement alligator cracking [40], pavement

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

skid number [41], pavement fatigue cracking [42]. Further, the Multivariate Analysis of Variance (MANOVA) capable of testing for two or more targets were adopted for rutting resistance of asphalt mixture [36]. In addition to traditional significant tests, regression, factor analysis, and discriminant analysis have also been integrated with ANOVA for pavement data analysis [40], [41].

IV. Linear and Nonlinear Regression

Linear and nonlinear regression models are simple while the most widely used statistical models for pavement data analysis. The coefficients can be estimated by either the least square method or the maximum likelihood method. The benefits of regression models include that the relationship between the target and predictors is explicit, the meanings of coefficients are easy to interpret, and the significance of each predictor can be tested, etc. However, it also has some limitations such as the normal distribution of the target variable, and the noncollinearity of the predictors for linear regression, etc. In addition to the traditional Multiple Linear Regression (MLR) and nonlinear regression, clusterwise regression, Multivariate Adaptive Regression Splines (MARS), Least Absolute Shrinkage and Selection Operator (LASSO), etc. have also been used to estimate mixture properties, to calculate pavement performance indices, and to predict pavement performance.

A. MLR

MLR is the most widely used regression model for both pavement performance index calculation and prediction. Interactions or the product of multiple predictors indicate the effect of one variable is dependent on other variables [43]. It can be transformed with power, logarithm, or exponential functions to describe nonlinear relationships [44]. A stepwise procedure is an iterative variable-selection procedure to select significant predictors for model fitting [45], [46]. In an MLR model as shown in Equation (3), the parameter estimate β_i is magnitude and direction change in response with each one-unit increase in predictor X_i while holding others constant.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \tag{3}$$

To calculate the performance index, MLR has been used to model International Roughness Index (IRI) [44], [47], surface deflection [48], pavement sustainability index [49], condition rating for continuously reinforced concrete [50], pavement flushing distress of thin-sprayed seal pavements [51], and pavement surface friction [45] based on a variety of factors including vehicle vertical acceleration, collected pavement distress, cracking, texture, rutting, and temperature, etc.

To predict pavement performance, MLR has been adopted to predict IRI [52], [53], IRI-drop, maintenance treatment effectiveness [43], rutting, riding quality [46], etc. based on pavement age, frost heave, pavement structure, pre-treatment condition, maintenance treatments, using the data collected from PMS in Canada, Spain, USA, etc. The R^2 of those models ranged from 0.47-0.86. Based on the MLR model, an incremental post-treatment pavement performance model can be developed to determine the optimized treatment application time [54].

TABLE I
Traditional Pavement Performance Prediction Models

Model		Equation
AASHO[57]		$PSI = PSI_0 - (PSI_0 - PSI_1)(\frac{w}{\rho})^{\beta}$
Paver [56]		$PCI = 100 - y\left(\frac{1.117}{a_x} + 0.143y_c + \frac{0.656}{T_c} - 1.23a_m\right)$
HDM[58]	Roughness	$\Delta RI = K_{gp}(\Delta RI_s + \Delta RI_c + \Delta RI_v + \Delta RI_t) + \Delta RI_e$
	Cracking	$CA = K_{cia}\left(CDS^2 a_0 e^{a_1 SNP + a_2\left(\frac{YE4}{SN^2}\right)} + CRT\right)$
	Rutting	$\Delta RDPD = K_{rpd} CDS^3 a_0 YE4 Sh^{a_1} HS^{a_2}$
Exponential[59]		$y = y_0 + be^{ct}$
Power[57]		$y = y_0 + bt^c$
Sigmoid	Garcia[60]	$IRI = IRI_0\left(1 - e^{a-bc^t}\right)$
	Texas[61]	$L_i = \rho e^{-(\frac{\chi\varepsilon\sigma\alpha}{t-\delta})^{\beta}}$
	Wu[62]	$PCI = a + \frac{b}{1 + e^{c(t+f)+d}}$

B. Nonlinear

Although MLR is simple and easy to interpret, nonlinear regression models are more preferred since they imply a specific relationship between predictors and targets based on engineering practices or mechanic theories. As summarized in TABLE I, the classic pavement performance model developed by the American Association of State Highway Officials (AASHO) used the power form based on the test roads in Illinois, and the model parameters have been modified since then [55]. The Paver's model and the HDM model use nonlinear polynomial equations [56]. Most performance models in PMS use exponential, power, sigmoid, or combinations of those. The most widely used is the sigmoid model capable of considering the change of performance deterioration rate over time.

To include pavement treatments in the performance models, many PMSs use the "family models", in which a group of models is defined for different treatments applied at different scenarios [63], [68]. For example, South Africa calibrated the HDM models for different combinations of structural capacity, traffic volume, base type, and climatic regions [64]. Washington State calibrated 24 performance models based on the data collected from 3000 pavement sections [65]. Tennessee State calibrated 81 models for 6 maintenance treatments at different traffic levels and pre-treatment pavement performance levels based on the data collected from 675 pavement sections [59].

Generally, the accuracy of nonlinear models is expected to be better than the MLR but not as good as Artificial Neural Networks (ANN) or Markov Chain (MC) models. In a study predicting rutting test results of asphalt mixture, the R2 of nonlinear regression and ANN were 0.92 and 0.99 respectively [69]. In another study predicting faulting distress of concrete pavement based on pavement age, pavement structural details, drainage features, traffic, and climate data, the MC performed the best, followed by ANN and nonlinear regression [70].

C. Clusterwise

The clusterwise MLR uses several regression equations called clusters for a dataset with a large variation. Each cluster indicates a portion of a dataset that follows a uniform tendency.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

A weighted regression function consisting of all clusters can be used for prediction. The clusterwise regression has been adopted to predict PSI and distress [71], [72], and obtained higher accuracy than the Markov model [72]. The clusterwise regression model can be modified by considering the membership of pavement to each cluster based on fuzzy logic and further reduce the prediction error [73]. A generalized algorithm can be added to the clusterwise regression to select the best linear or nonlinear model to predict pavement performance by exploring all possible combinations of potential significant predictors [74], [75]. The clusterwise MLR can also be improved to identify and address potential multiple collinearity issues [76].

D. MARS

The Multivariate Adaptive Regression Splines (MARS) is a non-parametric MLR including multiple basic functions. It is an extended linear model capable of modeling nonlinearities and interactions between variables. The MARS was firstly adopted to predict pavement IRI based on pavement age, cracking, environment, rutting, and patching, using the data generated by the HDM model [77]. In a study predicting pavement performance using the data from Turkey, and the R^2 of polynomial regression, MARS and ANN were 0.70, 0.71, and 0.75, respectively [78]. The MARS was also used to calculate pavement IRI based on pavement distress data including rutting, cracking, bleeding, corrugation, depression, patching, potholes, raveling, etc., and obtained an R^2 of 0.74 [79].

E. LASSO

The Least Absolute Shrinkage and Selection Operator (LASSO) is a regularized regression including both variable selection and regularization to enhance the prediction accuracy and interpretability and avoid overfitting. The LASSO was used to calculate pavement deflections based on cracking, structural number, climatic, layer thickness, and the modulus of pavement layers and subgrade soil [80], to predict the voids for curled concrete pavements based on pavement deflection data [81], and to determine a comprehensive performance indicator based on pavement comfort, safety and structural indicators [82].

F. Fuzzy Logic

One critical concern in pavement engineering is the large variation and uncertainty of data. Fuzzy logic in which membership functions are used to define the truth of degree of a value has been integrated with regression models for pavement performance evaluation and prediction. Fuzzy logic can be used with linear regression to predict pavement IRI based on pavement distresses [83], to calculate pavement performance based on roughness, transverse cracking, longitudinal cracking, pothole, and rutting [26], and to evaluate pavement condition based on roughness, pavement deflection, rutting, friction, and surface deterioration ratio [22].

V. Logistic Regression

Logistic Regression (LR) is a type of Generalized Linear Model (GLM) allowing the linear model to be related to the target by a link function including logarithm, exponential, logit, sigmoid, square root, etc. The LR uses a logarithm link function for a binary target. The “S” shaped logit function predicts two values (0 or 1), indicating the likelihood of the two events. As shown in Equation (4), the logit function of the probability P_i, defined as the natural logarithm transformation of the odds ratio, is expressed as a linear combination of predictors X_i [84]. LR is one of the simple while most extensively used machine learning algorithms for classification. The binary LR model predicts a binary categorical variable such as yes/no, while the ordinal and multinomial LR allows for more than two targets. LR was also used for the signal process for pavement evaluation. Hoang employed the Stochastic Gradient Descent Logistic Regression (SGD-LR) to identify pavement raveling based on extracted features from pavement images [85].

$$\begin{aligned}\operatorname{logit}(P_i) &= \operatorname{Ln}\left(\frac{P_i}{1-P_i}\right)\\ &= \beta_0 + \beta_1 X_1 + \beta_2 X_2\\ &\quad + \cdots + \beta_k X_k \end{aligned} \tag{4}$$

A. Binary LR

Binary LR models have been used to analyze the influence of mixture properties, traffic, climatic condition, pavement structural designs, and capacity, etc. on pothole patching serviceability [86], cracking initiation in both mixture and pavement [87], [89], pavement fatigue cracking [42], and pavement distress [90]. A mixed-effects binary LR has been developed to identify the relationship between the maintenance decisions and relevant factors based on the historical projects and to develop a maintenance decision-making prediction model [91]. One study reported that the multiple binary LR models were poor than the MC model in predicting flexible pavement distresses [90].

B. Ordered LR

The ordered LR model can model ordered multiple categories and has been adopted to analyze the severity levels for alligator cracking [92], pavement cracks intensity [93], [94], pavement crack progression [88], pavement treatment effectiveness [42]. Similar R^2 were reported using nonlinear regression, ordered and multinomial LR, and MC to predict pavement performance of 5 groups of pavement maintenance treatments using the data in Melbourne, Australia [95].

C. Ordered Probit Models

The ordered probit model is similar to the ordered LR and is also a type of GLM with different link functions. The link function for the ordered probit model is the inverse of the standard normal cumulative distribution shown in Equation (5). The ordered probit models have been used to predict the discrete condition of pavement performance [96], evaluate pavement maintenance effectiveness [97], and conduct pavement maintenance decision-making.

$$\begin{aligned}\operatorname{Probit}(P_i) = \Phi^{-1}(P_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2\\ + \cdots + \beta_k X_k\end{aligned} \tag{5}$$

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

VI. Count Data Model

Some pavement distress data such as the number of pavement cracks, potholes, or patches are count data, in which the observations are only non-negative integer values. The count data models including Poisson regression, negative binomial, and zero-inflated models be adopted for this type of modeling.

A. Poisson Regression

The Poisson process is a counting process, describing the number of events happening within a certain time interval. Poisson regression model or the log-linear model assumes the target has a Poisson distribution, and the logarithm of its expected value is a linear combination of predictors. It is often used to model distress occurrence in pavement engineering. Poisson regression is a type of GLM. It is noted that the GLM is not a simple transformation of the linear model. The link function is determined by the specific distribution of the target variable. Equation (6) shows the Poisson model, in which the logarithm of the mean of the time interval is a linear combination of predictors.

$$\ln E(Y \mid \boldsymbol{X}) = \ln \lambda = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad (6)$$

To evaluate the errors in pavement distress automatic acquisition, the number of occurrences of cracks with width intervals of 2.5 mm could be defined as a Poisson event [98]. The Poisson GLM has been used to simulate pavement degradation [99], and to predict pavement transverse cracking considering pavement age, traffic, climatic, etc. [100]. A Generalized Additive Model (GAM) can be used to extend the GLM to predict pavement fatigue cracking based on age, traffic, and climatic data and the R^2 ranged from 0.42 to 0.58 [101].

B. Zero-Inflated Models

However, Poisson distribution means the variance equals the mean. When this assumption is not valid, we can use the Negative Binomial (NB) regression model for those over-dispersed count data. When there are too many zeros in the observation which is the case for pavement cracks or potholes, we can use the Zero-Inflated Poisson (ZIP) or Zero-Inflated Negative Binomial (ZINB) models. A piecewise model consisting of a probit model and a logarithm generalized model was developed to describe the occurrence and propagation of pavement cracking, respectively [102]. Then, the NB and ZINB models were adopted to evaluate the initiation and propagation of pavement transverse cracking considering pavement age, traffic, materials, overlay thickness, and specific treatments [103]. The ZINB model included a logistic model for crack initiation and an NB model for crack propagation and outperformed the NB model. The NB model was used to predict pavement condition index with improved predictions by adding a Linear Empirical Bayesian (LEB) approach [104].

VII. Survival Analysis

The uncensored data, in which we only know the pavement service life is longer than a specific time but don't know its exact service time, has been suggested to be included in the pavement performance empirical models since the 1960s [105]. Survival analysis is to investigate the time of an event such as the occurrence of pavement distress or pavement failure. In 1986, the World Bank adopted the survival analysis model in the HDM-III [106] and some researchers believed that the survival model is better than the original AASHO pavement performance model [107], [108]. The survival models used in pavement engineering include three types: non-parametric, semi-parametric, and parametric models. As shown in Equation (7) and (8), the two key descriptive functions for survival analysis are the survival function $S(t)$ describing the probability that the event will not fail at time t, and the hazard function $h(t)$ describing the risk that the event will fail at time t. Hazard function can be defined as a function of predictors X_i to consider the effects of predictors.

$$S(t) = P(T \geq t) = 1 - \int_0^t f(u)du \quad (7)$$

$$h(t) = \lim_{\Delta t \to 0}\left(\frac{P(t \leq T \leq t + \Delta t)}{\Delta t}\right) = \frac{f(t)}{S(t)} \quad (8)$$

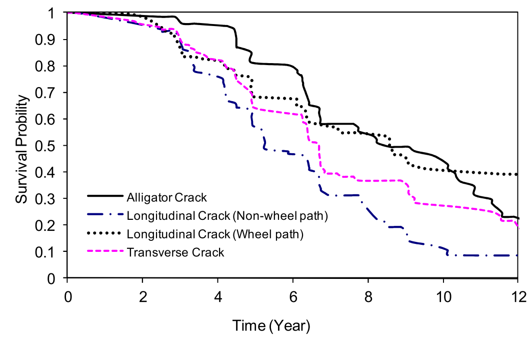


Fig. 2. Survival curves of four types of cracks [113].

A. Non-Parametric

The non-parametric models include the Kaplan-Meier (KM) product-limit method and the life table method, which can be used to test the significance of factors on survival time and compare different survival curves. Fig. 2 shows the survival curves based on the occurrence of pavement cracks. It has been used to evaluate the effect of RAP on pavement overlays [39], to analyze pavement deterioration subjected to hurricanes [109], compared survival curves of flexural and rigid pavements [110], to determine pavement rutting failure probability based on the full-scale accelerated pavement test in Louisiana [111], and to compare warranty and no warranty pavements in Mississippi [112].

B. Semi-Parametric

The semi-parametric model such as the proportional hazards model or the Cox model includes a model describing the relationship between survival time and influencing factors. It assumes the hazard rate of two individuals does not change

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

with time. It has been used to evaluate the stiffness deterioration of asphalt concrete under fatigue damage [114], the service life of asphalt surfacing in Norway [115], the pavement failure in Ohio [116], the effects of cracking sealing and filling on pavement performance [117], the performance of different treatments in US and Sweden [118], [119]. Mixed proportional hazards models were better than the Cox model and could incorporate the random effects caused by the traffic load, pavement type, climatic factors [120].

C. Parametric

The parametric model assumes the survival time meets a specific distribution such as exponential, Weibull, Logistic, Gamma, and Lognormal depending on the hazard function and includes a model describing the relationship between survival time and influencing factors. It usually requires the distribution test of survival time before building the model. It has been used to evaluate the occurrence of pavement cracking [113], the failure of pavement [121], the failure of pothole repairs [86], pavement failure indicated as extensive fatigue cracking [122], friction degradation in Pennsylvania [123], and the failure of pavement [124], [125]. Recent studies include incorporating Markov chain Monte Carlo (MCMC) sampling of Bayesian analysis into survival models to consider the effect of unobserved heterogeneity [126], and the correlations between different types of failures in the survival model [127].

VIII. Stochastic Process

A stochastic process is a process to describe the family of random variables indexed against some other variables, usually against time. It is the result of random experiments over time. Through observing the random phenomena, the random variables changing with time can be studied. In pavement engineering, the performance prediction, deterioration model, service life estimation, maintenance optimization, pavement strength prediction, and pavement design can all be analyzed through the stochastic process. In many cases, discrete ordinal variables are used for grading infrastructure conditions by setting threshold values for performance indices, such as the 5 levels based on pavement PSI [128], and the 8 levels based on pavement PCI [129], [130]. The ordinal variables are sufficiently accurate for network-level decision-making since the minor variation of the continuous pavement condition indices does not change the grading or the maintenance necessity.

A. Markov Models

Among various stochastic models, Markov Chain (MC) is the most widely used, mainly for network-level analysis. The key of MC is the Transition Probability Matrix (TPM), describing the probability of the transition between different states of pavement condition. It can be defined as the function of influencing factors. The MC has been used to predict pavement remaining strength and pavement design thickness [131], pavement deterioration [90], [132], [134], pavement distresses such as cracking [135], [138], pavement IRI for both flexural and rigid pavements [139], [141], pervious pavement performance [142], airfield pavement deterioration in Canada [143], the service life of pavement thermoplastic markings [144], and pavement effectiveness [145]. The semi-Markov model in which a state is defined for every given time in the process was developed to simulate the crack deterioration, which was proved to be superior to the traditional Markov chain model [141]. The Markov-based model can also be used for multi-objective optimization [146]. It has been integrated with the reinforced learning process to find the optimal pavement maintenance strategy [147].

B. Gaussian Process Regression

A Gaussian process is a stochastic process in which any finite sub-collection of random variables has a multivariate Gaussian or normal distribution. The mean and covariance functions can be obtained to model the probability distributions over functions determine or as prior knowledge. Gaussian process regression is a nonparametric Bayesian approach for regression. The Bayesian approach specifies a prior distribution and calculates the posterior distribution based on the training data. The Gaussian process regression could calculate the probability distribution over all admissible functions that fit the data and has been used to analyze the uncertainty in the Mechanistic-Empirical Pavement Design Guide (MEPDG) [148], to estimate pavement structural capacity based on surface deflections and surface temperature [149], and to predict the viscoelastic behavior of modified asphalt binders [150].

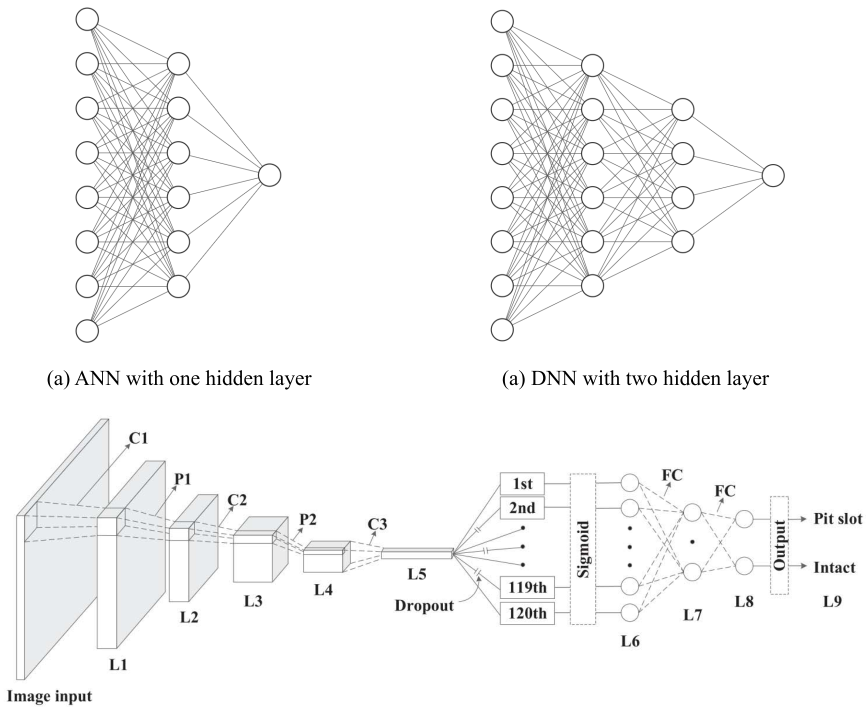
IX. Time Series

Time series data is a series of observations equally spaced in time. What to be noted is that time series is also a type of stochastic process. The time series model assumes the value at time t is composed of the trend, seasonal and random components in an additive or a multiplicative manner. Time series models can be classified into regression models including linear regression, moving averages, or exponential smoothing for prediction; and analytic models including Auto Regressive (AR), Moving Average (MA), Auto Regressive Moving Average (ARMA) models, and the more general Autoregressive Integrated Moving Average (ARIMA) model to describe a variable using its past values. In a typical $ARIMA\,(p, d, q)$ model, p is the models' autoregressive order, q is the moving average order, and d is the degree of differencing needed to achieve stationarity. Equation (9) shows an $ARMA\,(p, q)$ model, in which the p previous values and q previous errors were included for prediction.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{9}$$

Time series is a straightforward model to predict future pavement performance based on previous pavement performance and has been used in many practices. The unweighted moving average model has been used to smooth pavement condition data first to calculate a composite health index [151]. Autoregressive models with varying lags were also developed to predict pavement performance [152]. An ARMA(2,2) model was used in one study to smooth and predict pavement rutting

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.



(a) ANN with one hidden layer

(a) DNN with two hidden layer

(c) CNN (C1 is the first convolution, P1 is the first pooling[166])

Fig. 3. Structures of different ANNs.

data [153]. The ARMA models were reported to have good data-fitting capabilities, while structural time series models can provide a framework to identify the trend, seasonality, and random errors [154].

X. Artificial Neural Networks

Artificial Neural Networks (ANN) is now the most popular supervised machine learning and Artificial Intelligence (AI) algorithm. In the ANN, the weights of nodes that minimize the predictive error are determined during training. An activation function is applied to the sum of weighted input signals to determine its output. Backpropagation (BP) is the most common training method computing the gradient of the case-wise error function for the weights of a feed-forward network. A key benefit of ANN is that different layers can perform different transformations on their inputs, enabling complicated non-linear classification and regression. In the last decade, there has been an incrementing interest in using ANN to solve problems in pavement engineering. Not only are many pavement material or performance models built based on ANN, ANN-based algorithms including Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), etc. have been proved to be an effective technique to deal with unstructured data such as pavement image or vehicle acceleration signals. Fig. 3 shows the structure of an ANN with one hidden layer, a DNN with multiple hidden layers, and a CNN with 3 convolutional and 2 pooling calculations.

A. ANN

ANN has already been extensively used in pavement material properties prediction and pavement performance modeling, capable of handling a large number of input variables with high accuracy than most of the traditional regression models. It has been utilized to predict rutting test results of asphalt mixture [69], IRI [53], [155], [157], PCI, pavement cracking [158], [159], pavement roughness based on distress level [160], and overall concrete pavement condition index [161], and geogrid reinforced flexible pavement performance based on numerical simulations with different model parameters and scenarios [162].

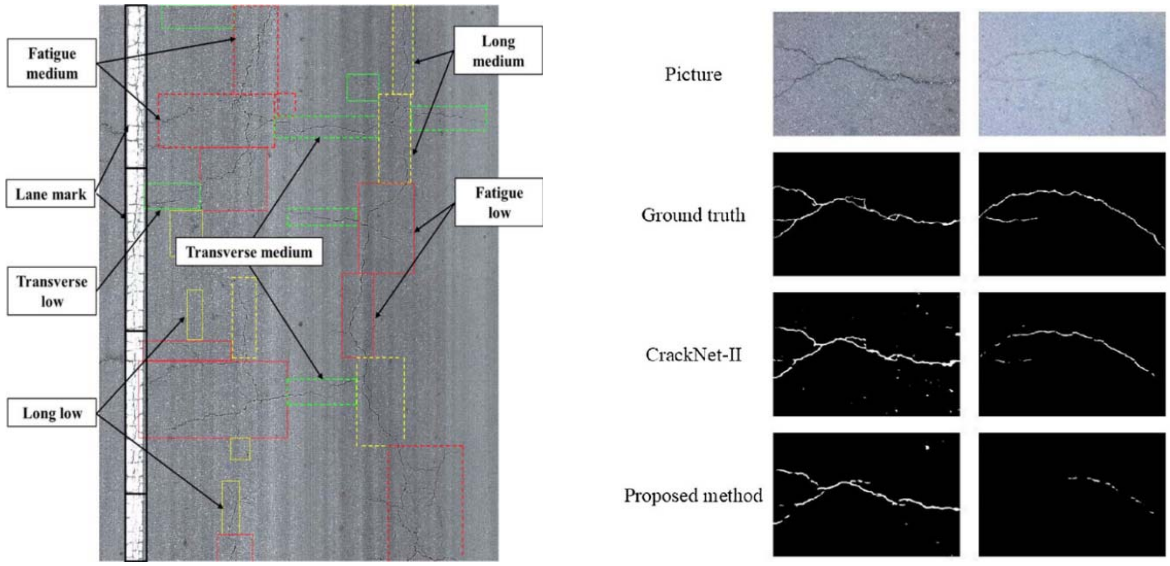
B. DNN

DNN is an ANN with multiple hidden layers between the input and output layers and therefore can model very complex non-linear relationships. DNN has been adopted to predict J-Integral of top-down cracking in asphalt pavement [163], to predict pavement rutting using 21 inputs with up to 3 hidden layers and 200 nodes [164], and to predict pavement roughness, rutting, cracking, and friction using 39 inputs [165].

C. CNN

CNN is a high efficient Deep Learning (DL) algorithm for image classification with multiple convolutional layers, pooling layers, activation layer, and the fully connected layer. The special structure of CNN enables it a proper technique for feature extraction for unstructured data. It assigns learnable weights and biases to various aspects/objects in the image to differentiate them. CNN had been widely used for distress recognition, location, and feature extraction and there is a fast increasing trend in this topic. CNN has been utilized to identify pavement cracking [167]–[169], potholes and texture

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.



(a) Object detection[177]

(b) Semantic segmentation[179]

Fig. 4. Pavement crack detection based on object detection and semantic segmentation.

from surface images [170], [171], and subgrade defects, moisture damages, and concealed cracks from Ground Penetrating Radar (GPR) images [172], [173]. An adaptive lightweight CNN model "Microcrack" has been developed for fast object classification on asphalt pavement crack images [174].

To classify, detect and extract pavement distress from pavement surface images or Ground Penetration Radar (GPR) graphs, the traditional image processing methods include histogram, threshold processing, morphology, edge detection, etc. [175], which are generally based on the concept that crack pixels are darker than the background. As the development of deep learning, many CNN-based computer vision algorithms are developed for image classifications and segmentations, as shown in Fig. 4. Image classification determines whether an image contains a specific type of object, exp. pavement distress. Object detection takes image classification one step further and provides the location of multiple objects, exp. different types of pavement distress. Frequently adopted object detection algorithms for pavement distress detection include YOLO [176], updated R-CNN [177], and the Faster R-CNN [178]. Image segmentation partitions an image into multiple segments or sets of pixels. Semantic segmentation specifies the object class, exp. distress or not distress, of each pixel in an image. Frequently adopted semantic segmentation algorithms include the two-step CNN [179], the feature pyramid and hierarchical boosting network [175], the Fully Convolutional Network (FCN) [180], U-net, and CrackU-net [181], [182]. Instance segmentation separates individual instances of each type of object, exp. every single distress in an image is segmented as an individual object. The Mask R-CNN which is an extension of Faster R-CNN [183], [184], was adopted to detect multiple pavement cracks in an image.

D. RNN

RNN uses the output from the previous step as input to the current step and is designed to handle sequential data. RNN is used in stock price modeling, speech recognition, natural language processing as well as pavement performance modeling since pavement performance data are time-series data. RNN has been used to predict PSI [185], cracking, rutting depth, and IRI [186]. RNN can also be used for pavement crack detection based on 3D asphalt surface data by treating pavement crack as a sequence of pixels that formulates a descended pattern [187].

XI. Decision Tree

Decision trees are supervised machine learning algorithms using tree-like models for predicting the class of the target from input variables. Decision trees do not require assumptions on the distribution of target variables, can handle a large number of factors, are tolerant of missing values, and are not sensitive to outliers. Therefore, it is one of the most effective and robust machine learning for classification and prediction. The general algorithm of a decision tree is to examine each input variable one at a time, create two or more groupings of the values of the input variable. After calculating all possible groupings for different input variables, it will select the single input variable that maximizes similarity within groupings and differences between groupings.

The Inductive Dichotomiser 3 (ID3) decision tree was firstly developed in the 1980s, based on which the modified C4.5 and C5.0 tree was developed. Then, the Classification and Regression Tree (CART) which generates two splits at each node was proposed as shown in Fig. 5. Decision trees have been used to investigate asphalt's adhesive behavior [188], the influence of material and traffic factors on pavement pothole patches [86], the influence of construction details in the effectiveness of slurry seals [189], and the influence of pavement design feature on roughness level [190]. Recent applications include the LR trees with Unbiased Selection (LOTUS) and Classification Rule with Unbiased Interaction Selection and Estimation (CRUISE) to identify critical

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

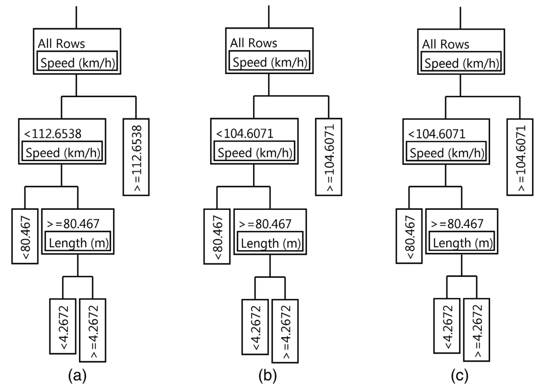


Fig. 5. Effects of factors on pothole patching performance using CART [86].

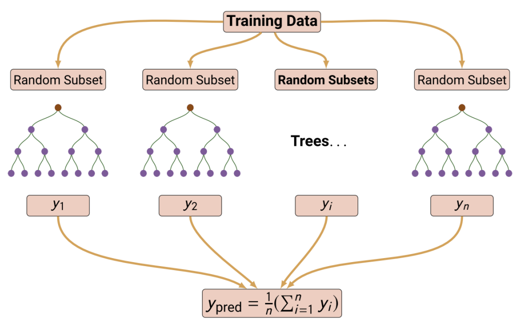


Fig. 6. Structure of RF algorithm [193].

pavement distresses for maintenance based on the maintenance history [191].

XII. Ensemble Learning

The recent development of decision trees is to use an ensemble of trees instead of a single tree, which could greatly improve the accuracy. Ensemble learning is machine learning in which multiple learners are trained to solve the same problem. It combines the predictions from multiple machine learnings such as decision trees, NN, etc. The ensemble learning includes two stages. The first stage is to generate a population (exp. 80%) of base learners from the training set, and the second stage is to combine them to create a stronger predictive model. Two important categories of ensemble learning are bagging and boosting. As shown in Fig. 6, Random Forest (RF) is a "bagging" algorithm combining results at the end of the process based on averaging or majority rules. Gradient Boosted Tree (GBT) is a boosting algorithm that builds each new tree to the residuals from the previous steps to improve the model.

The accuracy of the two ensemble trees is significantly higher than traditional decision trees. The RF has been used to predict pavement roughness [192], pavement distress based on mixture properties [193], pavement roughness based on vehicle responses [194], the strength of roller-compacted concrete pavement [195], and to identify pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery [196]. The GBT was also adopted to predict dynamic modulus of asphalt concrete [197], pavement rutting [198], cracking [199], roughness, etc. and was found to obtain the highest R^2 in predicting pavement deterioration, followed by RF, ANN, quadratic regression and linear regression [200].

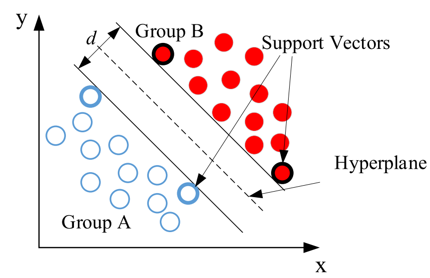


Fig. 7. The hyperplane and support vectors.

XIII. Support Vector Machine

Support Vector Machine (SVM) is a kernel-based non-probabilistic binary classifier and is one of the most popular and robust supervised machine learning algorithms for classification. As shown in Fig. 7, the support vectors are the data points closest to the decision surface or hyperplane and the algorithm is to find the optimum hyperplane maximizing the margin between the hyperplane and support vector. It can transform linear classification to nonlinear separation by mapping the data to a higher-dimensional space.

SVM has been used to predict the dynamic modulus of asphalt mixtures [201], pavement IRI [202], [204], and pavement remaining service life [205]. SVM could obtain comparative R^2 as NN due to its powerful capability of nonlinear fitting [201], [202]. In pavement nondestructive tests, SVM can be adopted to evaluate pavement roughness based on vehicle responses data such as accelerometer and wheel speed [194], and to identify transverse, longitudinal, and fatigue cracks based on pavement surface images [206], [207]. It is noted that SVM performed better than RF, ANN, CART, and discriminant analysis in the two studies on dealing with large volume datasets of pavement test data interpretation [194], [206], [207].

XIV. k-Nearest Neighbor

k-nearest neighbor (kNN) is an instance-based learning or lazy learning, which does not contain a training phase or build a model. The new samples are classified by comparing them against the entire training set. It is a non-parameter algorithm since there is no assumption for underlying data distribution. The kNN algorithm calculates the metrics distance of a new data point to the training data points, selects k nearest data points, and then classifies the data point to the class to which the majority of the k data points belong. As shown in Fig. 8, the unknown shape of the center point is estimated based on which shape, square or triangle, accounts for the majority in its k neighbors. The kNN has been used to classify mixtures with different moisture susceptibility [208] and to classify pavement PCI using the LTPP data [209]. The number of neighbors could be identified with optimum performance. Recently, kNN showed promising capability in

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

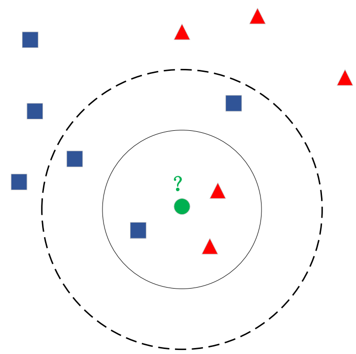


Fig. 8. kNN algorithm.

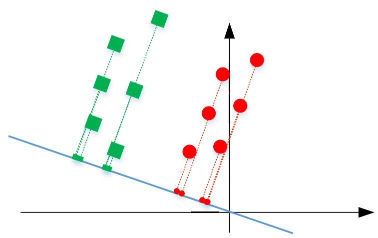


Fig. 9. Linear discriminant analysis.

pavement cracking classification and identification based on pavement surface images [210], [212], compared with RF, ANN, GBT, etc.

XV. Discriminant Analysis

The discriminant analysis was developed in the 1930s to classify observed data or samples into one of two or more groups based on their multiple characteristics. Different from cluster analysis which is unsupervised machine learning, the discriminant analysis is supervised machine learning and needs the training process with a sample of known classification. Frequently adopted discriminant algorithms include distance discriminant, Bayesian discriminant, linear discriminant, etc. The distance discriminant firstly determines the population of samples with known classification and then classifies a sample based on its distance to each classification. Frequently used distance metrics include the Euclidean distance, Manhattan distance, Minkowski distance, Hamming distance, etc. The Bayesian discriminant is capable of considering the prior probability of different populations. The linear discriminant is developed by Fisher in 1937 and is also called the Fisher discriminant. It uses a discriminant function maximizing the sum of squares between different groups and minimizing the sum of squares within a group, which is to find the optimal projection as shown in Fig. 9. In 1987, the discriminant model has been used to determine if the pavement section needs an overlay treatment based on a z value, to analyze the design and site factors on the performance of in-service flexible pavements from the LTPP [42], to analyze if different rest time has a significant influence on the fatigue life of asphalt mixture [213], and to classify pavement cracks based on images [207].

XVI. Cluster Analysis

Cluster analysis is to classify either samples or variables into different groups based on their similarity. It has been used in many fields including psychology, economics, bioinformatics, image analysis, etc., and is a type of exploratory data mining or unsupervised machine learning method. Distance metrics are usually used to measure the similarity between samples while cosine similarity is used to can classify variables. There are many algorithms for cluster analysis. The most frequently used clustering methods are hierarchical clustering and k-means clustering. Hierarchical clustering begins with treating each sample as one cluster and starts merging the two nearest clusters until a single all-encompassing cluster remains. It creates a hierarchical tree-like structure. k-means clustering partitions n observations into k clusters based on the distance. It uses selected k centroids as the beginning points and then performs iterative calculations to optimize the positions of the centroids by minimizing the distances within each cluster as shown in Equation (10).

$$d_{ij} = min\left(\left\|x_i - z_j\right\|\right), \quad x_i \in S,\ z_j \in Z \tag{10}$$

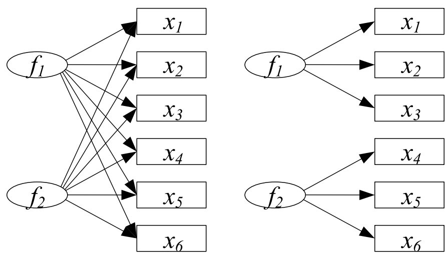
Hierarchical clustering has been used to generate axle loading distribution input for the MEPDG pavement design using the data obtained from the weight in motion system [214]. The normalized cuts cluster algorithm has been used to classify 35 pavement sections into 5 clusters based on 8 performance indicators or maintenance decision making [215]. Cluster analysis is also promising in the signal process with automatically collected data. It has been used to extract the smartphone sensor data for pavement potholes and pumps identification [216], to identify the potential dipping in the groove measurement with laser profiling data [217], to classify the sound measured inside a vehicle for pavement riding quality measurement [218], and to identify cracking modes in porous asphalt based on the acoustic emission data [219].

XVII. Principal Component Analysis

Principal Component Analysis (PCA) is to convert a set of possibly correlated variables into a set of linearly uncorrelated variables called principal components using an orthogonal transformation. As shown in Equation (11), each principal component F_i is a linear combination of original variables $x_1, x_2, \ldots, x_p$. a_{ij} is the loading coefficients of x_i on F_j. The first principal component F_1 contains the most variance, the second principal component F_2 is orthogonal to the first and contains the second greatest variance, the third principal component is orthogonal to all previous ones and also contains the third greatest variance, etc. Since the first several of the principal components can explain the major variation of the original dataset, PCA is usually used to reduce the dimensionality of a data set. The principal components can be calculated by the covariance or correlation matrix.

$$\begin{cases} F_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{p1}x_p \\ F_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{p2}x_p \\ \quad\cdots \\ F_p = a_{1p}x_1 + a_{2p}x_2 + \cdots + a_{pp}x_p \end{cases} \tag{11}$$

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.



(a) EFA (b) CFA

Fig. 10. Difference between EFA and CFA.

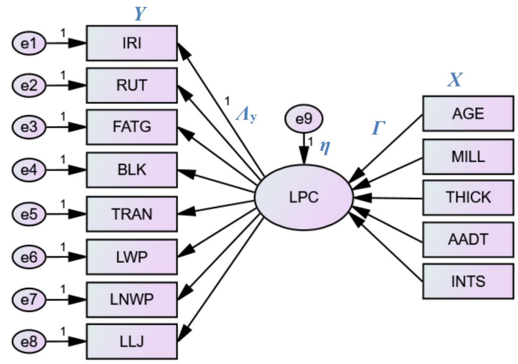


Fig. 11. An SEM model for latent pavement condition.

PCA has been used to determine 3 principal components representing structure, deformation, and interlayer bonding respectively to replace the original 7 pavement performance indicators [220], to reduce 21 traffic variables into 3 principle components for the pavement performance prediction [165], and to reduce 16 pavement performance variables to 3 principle components covering 98.7% of the total variance [221], and to reduce 17 asphalt mixture properties to 5 principal components explaining 89.72% of the total variance [222]. Based on the results of PCA analysis, the pavement sections could be classified into 4 clusters with similar performance levels [221]. To process the unstructured pavement test data, PCA can be used to analyze the sound recorded underneath a moving vehicle to estimate the mean texture depth of pavement [223].

XVIII. Factor Analysis

Factor analysis is to describe variability among correlated variables with a lower number of unobserved factors. Each variable is a linear combination of common factors and a unique factor or an error term. The coefficients in the linear function of each variable are also called the loadings, indicating the contribution of common factors on the variance of the variable. As shown in Equation (12), each variable X_i is a linear combination of uncorrelated common factors $f_1, f_2, \ldots, f_m$ and an error term. a_{ij} is the factor loadings. Similar to the PCA, factor analysis can also be used for dimensionality reduction. The traditional factor analysis is also called Exploratory Factor Analysis (EFA) aiming to identify the common factors for all variables, while the Confirmatory Factor Analysis (CFA) is to investigate how well the hypothesized factor structure fits with the variables. As shown in Fig. 10, each variable loads on all factors in EFA while each variable loads on only one factor in CFA.

$$X_i = \mu + a_{i1} f_1 + a_{i2} f_2 + \cdots + a_{im} f_m + \varepsilon_i \tag{12}$$

In pavement engineering, factor analysis is used to analyze the relationship between material properties or pavement distresses and the unobserved factors. One study reported that the 27 properties of asphalt mixture have three common factors: the permanent deformation factor highly correlated with voids in mineral aggregate and Marshall stability, the shear resistance factor highly correlated with voids and voids fill with asphalt, and the moisture susceptibility factor highly correlated with residual Marshall stability and the tensile strength ratio [224]. Pavement condition data include the surface distress factor, the surface roughness factor, and the pavement structural condition factor correlated solely with the rolling wheel deflectometer data [40]. Based on the CFA analysis, pavement condition data include three factors: the riding comfort factor highly correlated with roughness, the early age cracking factor highly correlated with longitudinal and transverse cracking, and the aged severe damage factor highly correlated with fatigue cracking, block cracking, longitudinal joint distress and rutting [225].

XIX. Structural Equation Modeling

The Structural Equation Modeling (SEM) is composed of a measurement model describing the relationship between unobservable variables and their observable measurements and a structural model describing the relationship between those unobservable variables. A CFA model describing how well variables load on several factors is a measurement model. The major contribution of SEM in pavement engineering is that it treats the real pavement performance as an unobservable variable and different pavement performance indicators are its observable measurements.

In pavement performance modeling, SEM was firstly adopted to estimate the latent PSI considering the traffic, pavement, and climatic factors [226], [228]. Then, a time series was integrated with the SEM to evaluate the pavement maintenance effectiveness [154], [229]. SEM was also used to determine the weights for calculating latent PCI considering various pavement distresses based on the data collected from the LTPP database [227], and to detail the effects of different pavement overlay treatments on specific pavement performance [230]. Fig. 11 shows an example of the SEM model. AGE, AADT, THICK, MILL, and INTS are the exogenous observed pavement age, traffic, structure, and grade variables influencing the latent pavement condition (LPC) variable, IRI, RUT, FATG, BLK, TRAN, LWP, LNWP, LLJ, and PATCH are the endogenous observed pavement performance and distress variables of the three latent pavement condition variables [227].

XX. Reinforcement Learning

Machine learning includes three basic paradigms. The supervised learnings include regression, LR, ANN, decision

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

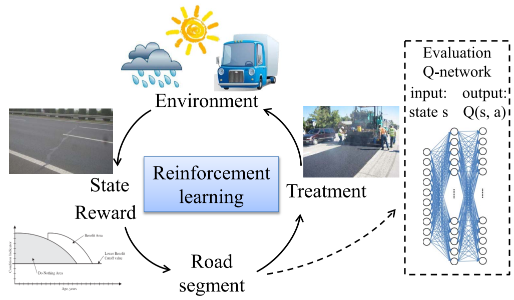


Fig. 12. The deep RL model for pavement maintenance strategy optimization.

tree, SVM, kNN, and discriminant analysis, which all needs labeled dataset to train the models for prediction and classification. The supervised learnings include clustering, PCA, and factor analysis, which detect the correlation structures of an unlabeled dataset for classification. Reinforcement Learning (RL) is to determine the action of agents in an environment by maximizing its cumulative reward, similar to strategic optimization. In transportation engineering, RL has been used in traffic signal control to minimize vehicle delay, and traffic flow optimization to reduce congestion [231]. In pavement engineering, there is only one study using RL for long-term pavement maintenance strategy optimization using the data in Jiangsu, China. Fig. 12 shows the structure of the developed deep RL. The maintenance treatments, treatment performance models, and maintenance effectiveness, and 42 variables involving the pavement structures and materials, traffic loads, maintenance records, pavement conditions, etc., were treated as actions, environment, rewards, and states in the RL algorithm.

XXI. Bayesian Analysis

To improve model quality, we can use a large volume high-quality dataset for training, or to improve the accuracy of parameter estimates. Bayes' theorem updates the posterior probability based on the prior probabilities. The prior information combined with current data is used to obtain a posterior estimate of parameters in Bayesian analysis. Firstly, the prior probability distribution for a parameter is specified based on previous experience or knowledge. Then, the Bayes' law is used to provide a posterior probability distribution for the parameter [84]. In pavement engineering, the Bayesian analysis can be used to obtain and update the values of parameters based on the historical data.

The Bayesian Markov hazard model has been developed to estimate the condition changes of the civil infrastructures [232]. The mixed hazard model with Bayesian estimation could be used to define the change of pavement performance after repeated maintenance to support the decision-making system in asset management [233]. The parameter of the transition probability matrix of the Markov chain can be estimated through Bayesian analysis to predict the pavement performance [234].

A. Markov Chain Monte Carlo

Based on Bayes' law, the posterior probability density is proportional to the likelihood function multiplying the prior probability. After the posterior probability is obtained, the parameter can be estimated and the hypothesis can be tested. Generally, the integrals in the likelihood function can be used to obtain the mean value of parameter estimation. If the posterior probability is multivariate distributed, the multi-integral would become complicate which retards the application of Bayesian analysis. Hence, Monte Carlo simulation can approximate the integrals when the sample size is large enough. However, with the increase of dimensions and complexity of distribution, Monte Carlo simulation is not applicable. In this case, Markov Chain Monte Carlo (MCMC) is appropriate for simulation.

MCMC has been used to obtain the posterior distributions of the parameters in the Bayesian linear mixed-effects model to predict pavement rutting in accelerated pavement testing [235], in the Poisson hidden MCMC bay Markov model to predict the condition state [236], in Bayesian survival model to assess the pavement deterioration [237].

B. Metropolis-Hasting Sampling

In a Markov chain, the next state only depends on the current state and is independent of the previous states. The ergodic theorem of the Markov chain indicates that if the Markov chain is ergodic and the iteration is great enough, the distribution of samples is approximate to the true distribution no matter what the starting value is. An ergodic Markov Chain always has a stationary distribution. Therefore, the critical point of the MCMC is to construct an ergodic Markov chain with a stationary distribution. Metropolis-Hasting sampling and Gibbs sampling are the two widely-used sampling algorithms of MCMC.

Metropolis-Hastings (MH) is an iterative algorithm used to generate a sequence of serially correlated samples from the probability distributions that converge to a given target distribution. At each iteration, the acceptance ratio is used to decide whether the candidate is used or discarded in the next iteration. MH has been used to obtain the posterior distribution of the parameters of the sigmoidal equation to predict the longitudinal cracking [238], to update the parameters in LR to analyze the failure probability of pavement preventive maintenances [239], to estimate the parameters of non-homogenous Markov hazard model to evaluate the cracking condition states [240], to predict the pavement life through estimating the model based on the MEPDG model [241]. As shown in Fig. 13, the parameter estimate was significantly reduced.

C. Gibbs Sampling

Gibbs sampling is used to generate the posterior samples by sweeping through each variable to sample from its condition distribution with the remaining variables fixed. This iteration continues until it converges. Unlike the MH algorithm, all proposed samples are accepted. Gibbs sampling has been used to estimate the parameter distributions to forecast pavement deterioration [242], to estimate and predict the pavement layer thickness based on the GPR data [243], to investigate the propagation of transverse cracks on pavements [244], and to develop the probability model between IRI and the expected

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

TABLE II

SUMMARY OF THE CHARACTERISTICS OF DATA ANALYSIS METHODS

Data analysis methods		Characteristics
DOE	Full factorial	Examine all factors and possible interactions
	Partial factorial	Use a subset of a full factorial design to examine some of the main effects and 2-way interactions
	Taguchi	Least number of runs to examine some of the main effects and interactions
Significance tests	ANOVA	Examine the statistical differences on one continuous target by a predictor
	MANOVA	Examine the statistical differences on a weighted linear combination of targets by a predictor
	t-test	Test the differences between groups
	Paired t-test	Test the differences between the pairs from two groups
	Turkey's test	Find the individual that is significantly different by comparing the mean between each pairwise combination of groups
Linear and nonlinear regression	Multiple linear regression	Calculate the target based on a weighted sum of multiple predictors and interactions
	Nonlinear regression	Calculate based on a known nonlinear relationship between the targets and predictors
	Clusterwise	Partition observations into clusters and minimize the error computed over all the clusters
	MARS	Model non-linear relationships between the variables
	LASSO	Select and regularize variables
	Fuzzy logic	Consider variable the uncertainties using the membership functions
Logistic regression	Binary LR	Binary classification based on multiple predictors
	Ordinal LR	Multiple (>2) classification based on multiple predictors
	Ordered probit model	Model the relationship between an ordinal target and a set of predictors
Count data model	Poisson regression	Model counting data such number of cracking with Poisson distribution
	Zero-inflated model	Model counting data with too many zeros
Survival analysis	Product-limit model and life table method	Compare multiple survival times or curves
	Cox model	Evaluate the effect of predictors on survival curves with no assumption on the distribution of survival times
	Parametric model	Evaluate the effect of predictors on survival curves with the known distribution of survival times
Stochastic process	Markov chain	Model the probability of the transition between different grades of pavement condition
	Gaussian process regression	A Bayesian approach using prior distribution for prediction in a finite sub-collection of variables with a multivariate Gaussian distribution
Time series	AR	Predict future values considering previous values
	MA	Predict future values considering previous errors
	ARMA	Predict future values considering both previous values and errors
Neural networks	ANN	Nonlinear prediction with a large volume variables and samples with high accuracy
	DNN	ANN with multiple hidden layers and higher accuracy
	CNN	ANN with convolutional, pooling, and activation layers for handling imagery unstructured data
	RNN	ANN using previous output as input to handle sequential data for prediction
Decision tree	ID3, C4.5, C5.0, CART…	Partition by maximizing similarity within groups and differences between groups
Ensemble learning	Random forest	Use the mean of multiple decision trees
	GBT	Use a sequence of trees to fit the residuals from the previous steps
SVM		Binary and a nonlinear classifier based on hyperplane and support vectors by mapping the data to a higher-dimensional space
kNN		An instance-based learning classifier by comparing new sample against its neighbors
Discriminant analysis	Distance	Classify based on the distance to groups
	Bayesian	Classify considering the prior probability
	Linear	Classify by maximizing errors between different groups and minimizing errors within groups
Cluster analysis	Hierarchical	Clustering by merging the two nearest clusters according to distance metrics
	k-means	Clustering by minimizing distance within clusters and maximizing distance between groupings

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

TABLE II
(Continued.) SUMMARY OF THE CHARACTERISTICS OF DATA ANALYSIS METHODS

PCA		Determine principal components containing the most variance to using an orthogonal transformation
Factor analysis	EFA	Identify common factors for all variables
	CFA	Investigate variable loadings on specified factors
SEM		Investigate the relationship between unobservable variables and the relationship between unobservable variables and their observable measurements
Reinforcement learning		Strategy optimization by maximizing the cumulative reward of different actions in an environment
Bayesian	MCMC	Obtain the posterior distributions of the parameters based on sampling
	MH sampling	MCMC sampling by an iterative algorithm
	Gibbs sampling	MCMC sampling by sweeping each variable from its condition distribution with the remaining variables

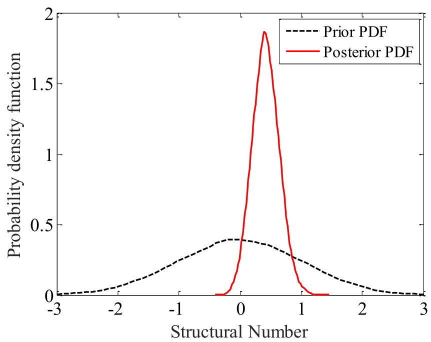


Fig. 13. Decease of uncertainty of parameter through MCMC for a LR regression [239].

pavement life through life-cycle cost analysis [245]. In addition to the Gibbs sampling, another efficient sampling called Hamiltonian Monte Carlo has also been introduced to further reduce the uncertainty of parameter estimates [246]–[248].

XXII. CONCLUSION

Table II summarizes the data analysis methods used in pavement engineering and their characteristics. Generally, statistical models including linear, nonlinear, generalized linear, and logistic regression models, survival analysis, and stochastic process models are proper for significant factors quantification and pavement performance predictions with explicit model equation and clear coefficients meaning. The supervised machine learnings including ANN, decision trees, SVM, kNN, and discriminant analysis, etc. are powerful in prediction and classification and can deal with large data volume. The unsupervised machine learnings including PCA, factor analysis, and cluster analysis can be used to find the correlations between multiple variables to reduce the dimensionality and extract common factors, and therefore is usually used to pre-process data before modeling with supervised learnings.

Each model has its benefits or limitations. For example, linear models require normal distribution of the target and are less tolerant for collinearity. Zero-inflated models can handle large zero values in the target while Poisson models cannot. Survival analyses can deal with censored data. Time series models use previous conditions for prediction while Markov models only consider the current state for prediction. Decisions are tolerant of outliers. ANNs are more like black boxes while regression, decision trees, and SVMs can also show the explicit forms or parameters. CNNs are powerful to deal with image processes with automatic feature extraction. Therefore, it is necessary to select proper models based on the objectives and available data.

In future practices and studies, pavement engineers are facing three main challenges including pavement long-term preservation, pavement nondestructive testing, and pavement condition sensing. Accordingly, we will have accumulated pavement-related data in the PMSs; the big data from multiscale material characterizations, high-resolution images, and laser cloud points at macro and micro scales; and the dynamic monitoring data from pavement instrumentations. It will be the responsibility of pavement engineers to interpret those data to help pavement evaluation and maintenance. A trend is to select proper or combinations of two or more types of models for different stages of data clustering, feature extraction, and model training.

REFERENCES

[1] M. Huang, Q. Dong, F. Ni, and L. Wang, "LCA and LCCA based multi-objective optimization of pavement maintenance," *J. Cleaner Prod.*, vol. 283, Feb. 2021, Art. no. 124583.
[2] J. Rice and N. Martin, "Smart infrastructure technologies: Crowd-sourcing future development and benefits for Australian communities," *Technol. Forecasting Social Change*, vol. 153, pp. 119–256, Apr. 2020.
[3] A. Di Graziano, V. Marchetta, and S. Cafiso, "Structural health monitoring of asphalt pavements using smart sensor networks: A comprehensive review," *J. Traffic Transp. Eng. English Ed.*, vol. 7, no. 5, pp. 639–651, Oct. 2020.
[4] R. Mehmood, S. S. I. Katib, and I. Chlamtac, *Smart Infrastructure and Applications*. Berlin, Germany: Springer, 2020.
[5] D. Čygas, A. Laurinavičius, M. Paliukaitė, A. Motiejūnas, L. Žiliūtė, and A. Vaitkus, "Monitoring the mechanical and structural behavior of the pavement structure using electronic sensors," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 30, no. 4, pp. 317–328, Apr. 2015.
[6] H.-P. Wang, P. Xiang, and L.-Z. Jiang, "Optical fiber sensing technology for full-scale condition monitoring of pavement layers," *Road Mater. Pavement Design*, vol. 21, no. 5, pp. 1258–1273, Jul. 2020.
[7] Y. Bi, J. Pei, F. Guo, R. Li, J. Zhang, and N. Shi, "Implementation of polymer optical fibre sensor system for monitoring water membrane thickness on pavement surface," *Int. J. Pavement Eng.*, vol. 22, no. 7, pp. 872–881, Jun. 2021.
[8] X. Chapeleau *et al.*, "Assessment of cracks detection in pavement by a distributed fiber optic sensing technology Continuous health monitoring of pavement systems using smart sensing technology," *J. Civil Struct. Health Monit.*, vol. 7, no. 4, pp. 459–470, Jul. 2017.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[9] A. H. Alavi, H. Hasni, N. Lajnef, and K. Chatti, "Continuous health monitoring of pavement systems using smart sensing technology," *Construct. Building Mater.*, vol. 114, pp. 719–736, Jul. 2016.

[10] H. Bhuyan, A. Scheuermann, D. Bodin, and R. Becker, "Soil moisture and density monitoring methodology using TDR measurements," *Int. J. Pavement Eng.*, vol. 21, no. 10, pp. 1263–1274, Aug. 2020.

[11] C.-H. Ho, M. Snyder, and D. Zhang, "Application of vehicle-based sensing technology in monitoring vibration response of pavement conditions," *J. Transp. Eng. B, Pavements*, vol. 146, no. 3, Sep. 2020, Art. no. 04020053.

[12] W. Carey and P. Irick, "The pavement serviceability-performance concept," *Highway Res. Board Bull.*, vol. 250, pp. 1–20, 1960.

[13] K. Hall and C. Muñoz, "Estimation of present serviceability index from international roughness index," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1655, no. 1, pp. 93–99, Jan. 1999.

[14] F. H. Scrivner and W. R. Hudson, "A modification of the AASHO road test serviceability index formula," *Highway Res. Rec.*, no. 46, pp. 71–88, 1964.

[15] C. Liu and R. Herman, "New approach to roadway performance indices," *J. Transp. Eng.*, vol. 122, no. 5, pp. 329–336, Sep. 1996.

[16] *Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys*, Standard ASTM D6433-99, American Society for Testing and Materials, 1999.

[17] N. N. Eldin and A. B. Senouci, "A pavement condition-rating model using backpropagation neural networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 10, no. 6, pp. 433–441, Nov. 1995.

[18] N. C. Jackson, R. Deighton, and D. L. Huft, "Development of pavement performance curves for individual distress indexes in South Dakota based on expert opinion," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1524, no. 1, pp. 130–136, Jan. 1996.

[19] C. H. Juang and S. N. Amirkhanian, "Unified pavement distress index for managing flexible pavements," *J. Transp. Eng.*, vol. 118, no. 5, pp. 686–699, Sep. 1992.

[20] H. K. Koduru, F. Xiao, S. N. Amirkhanian, and C. H. Juang, "Using fuzzy logic and expert system approaches in evaluating flexible pavement distress: Case study," *J. Transp. Eng.*, vol. 136, no. 2, pp. 149–157, Feb. 2010.

[21] C. L. Saraf, "Pavement condition rating system: Review of PCR methodology," Resour. Int., Westerville, OH, USA, Tech. Rep. FHWA/OH-99/004, 1998.

[22] L. Sun and W. Gu, "Pavement condition assessment using fuzzy logic theory and analytic hierarchy process," *J. Transp. Eng.*, vol. 137, no. 9, pp. 648–655, Sep. 2011.

[23] A. Bianchini, "Fuzzy representation of pavement condition for efficient pavement management," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 27, no. 8, pp. 608–619, Sep. 2012.

[24] A. Golroo and S. L. Tighe, "Fuzzy set approach to condition assessments of novel sustainable pavements in the Canadian climate," *Can. J. Civil Eng.*, vol. 36, no. 5, pp. 754–764, May 2009.

[25] M. Karaşahin and S. Terzi, "Performance model for asphalt concrete pavement based on the fuzzy logic approach," *Transport*, vol. 29, no. 1, pp. 18–27, Mar. 2014.

[26] N.-F. Pan, C.-H. Ko, M.-D. Yang, and K.-C. Hsu, "Pavement performance prediction through fuzzy regression," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 10010–10017, Aug. 2011.

[27] N. G. Gharaibeh, Y. Zou, and S. Saliminejad, "Assessing the agreement among pavement condition indexes," *J. Transp. Eng.*, vol. 136, no. 8, pp. 765–772, Aug. 2010.

[28] J. F. Daieiden, A. Simpson, and J. B. Rauhut, *Rehabilitation Performance Trends: Early Observations From Long-Term Pavement Performance (LTPP) Specific Pavement Studies (SPS)*. McLean, VA, USA: Turner-Fairbank Highway Research Center, 1998.

[29] R. W. Perera and S. D. Kohn, "International roughness index of asphalt concrete overlays: Analysis of data from long-term pavement performance program SPS-5 projects," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1655, no. 1, pp. 100–109, Jan. 1999.

[30] E. Freitas, C. Freitas, and A. C. Braga, "The analysis of variability of pavement indicators: MPD, SMTD and IRI. A case study of Portugal roads," *Int. J. Pavement Eng.*, vol. 15, no. 4, pp. 361–371, Apr. 2014.

[31] S. Sachs, J. M. Vandenbossche, and M. B. Snyder, "Calibration of national rigid pavement performance models for the pavement mechanistic–empirical design guide," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2524, no. 1, pp. 59–67, Jan. 2015.

[32] A. Noory, F. M. Nejad, and A. Khodaii, "Effective parameters on interface failure in a geocomposite reinforced multilayered asphalt system," *Road Mater. Pavement Design*, vol. 19, no. 6, pp. 1458–1475, Aug. 2018.

[33] L. Yongjie, Y. Shaopu, and W. Jianxi, "Research on pavement longitudinal crack propagation under non-uniform vehicle loading," *Eng. Failure Anal.*, vol. 42, pp. 22–31, Jul. 2014.

[34] J.-C. Du, "Evaluation of asphalt pavement layer bonding stress," *J. Civil Eng. Manage.*, vol. 21, no. 5, pp. 571–577, May 2015.

[35] A. Rahman, H. Huang, C. Ai, H. Ding, C. Xin, and Y. Lu, "Fatigue performance of interface bonding between asphalt pavement layers using four-point shear test set-up," *Int. J. Fatigue*, vol. 121, pp. 181–190, Apr. 2019.

[36] Z. Zhao, J. Jiang, F. Ni, Q. Dong, J. Ding, and X. Ma, "Factors affecting the rutting resistance of asphalt pavement based on the field cores using multi-sequenced repeated loading test," *Construct. Building Mater.*, vol. 253, Aug. 2020, Art. no. 118902.

[37] N. Sabahfar, A. Ahmed, S. R. Aziz, and M. Hossain, "Cracking resistance evaluation of mixtures with high percentages of reclaimed asphalt pavement," *J. Mater. Civil Eng.*, vol. 29, no. 4, Apr. 2017, Art. no. 06016022.

[38] D. Mishra and E. Tutumluer, "Aggregate physical properties affecting modulus and deformation characteristics of unsurfaced pavements," *J. Mater. Civil Eng.*, vol. 24, no. 9, pp. 1144–1152, Sep. 2012.

[39] M. Fakhri and E. Amoosoltani, "The effect of reclaimed asphalt pavement and crumb rubber on mechanical properties of roller compacted concrete pavement," *Construct. Building Mater.*, vol. 137, pp. 470–484, Apr. 2017.

[40] K. Gaspard, Z. Zhang, and M. A. Elseifi, "Integration of rolling wheel deflectometer deflection measurements into pavement management systems: Use of multivariate statistical methods and fuzzy logic," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2366, no. 1, pp. 25–33, Jan. 2013.

[41] L. Fuentes, M. Gunaratne, and D. Hess, "Evaluation of the effect of pavement roughness on skid resistance," *J. Transp. Eng.*, vol. 136, no. 7, pp. 640–653, Jul. 2010.

[42] S. W. Haider and K. Chatti, "Effect of design and site factors on fatigue cracking of new flexible pavements in the LTPP SPS-1 experiment," *Int. J. Pavement Eng.*, vol. 10, no. 2, pp. 133–147, Apr. 2009.

[43] Q. Dong and B. Huang, "Evaluation of effectiveness and cost-effectiveness of asphalt pavement rehabilitations utilizing LTPP data," *J. Transp. Eng.*, vol. 138, no. 6, pp. 681–689, Jun. 2012.

[44] K. Park, N. E. Thomas, and K. W. Lee, "Applicability of the international roughness index as a predictor of asphalt pavement condition," *J. Transp. Eng.*, vol. 133, no. 12, pp. 706–709, Dec. 2007.

[45] G. Yang, K. C. P. Wang, and J. Q. Li, "Multiresolution analysis of three-dimensional (3D) surface texture for asphalt pavement friction estimation," *Int. J. Pavement Eng.*, pp. 1–10, 2020, doi: 10.1080/10298436.2020.1726350.

[46] H. K. Salama, K. Chatti, and R. W. Lyles, "Effect of heavy multiple axle trucks on flexible pavement damage using in-service pavement performance data," *J. Transp. Eng.*, vol. 132, no. 10, pp. 763–770, Oct. 2006.

[47] M. M. de Farias and R. O. de Souza, "Correlations and analyses of longitudinal roughness indices," *Road Mater. Pavement Des.*, vol. 10, no. 2, pp. 399–415, Jan. 2009.

[48] W. Huang, S. Liang, and Y. Wei, "Surface deflection-based reliability analysis of asphalt pavement design," *Sci. China Technolog. Sci.*, vol. 63, no. 9, pp. 1824–1836, Sep. 2020.

[49] S. O. Obazee-Igbinedion and O. Owolabi, "Pavement sustainability index for highway infrastructures: A case study of Maryland," *Frontiers Struct. Civil Eng.*, vol. 12, no. 2, pp. 192–200, Jun. 2018.

[50] S.-S. Wu, "Developing a quantitative rating system for continuously reinforced concrete pavement," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1699, no. 1, pp. 11–15, Jan. 2000.

[51] S. Kodippily, T. F. P. Henning, and J. M. Ingham, "Detecting flushing of thin-sprayed seal pavements using pavement management data," *J. Transp. Eng.*, vol. 138, no. 5, pp. 665–673, May 2012.

[52] O. Sylvestre, J.-P. Bilodeau, and G. Doré, "Effect of frost heave on long-term roughness deterioration of flexible pavement structures," *Int. J. Pavement Eng.*, vol. 20, no. 6, pp. 704–713, Jun. 2019.

[53] N. Abdelaziz, R. T. A. El-Hakim, S. M. El-Badawy, and H. A. Afify, "International roughness index prediction model for flexible pavements," *Int. J. Pavement Eng.*, vol. 21, no. 1, pp. 88–99, Jan. 2020.

[54] X. Chen, H. Zhu, Q. Dong, and B. Huang, "Optimal thresholds for pavement preventive maintenance treatments using LTPP data," *J. Transp. Eng., A, Syst.*, vol. 143, no. 6, Jun. 2017, Art. no. 04017018.

[55] K. A. Small and C. Winston, "Optimal highway durability," *Amer. Econ. Rev.*, vol. 78, no. 3, pp. 560–569, 1988.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[56] M. Y. Shahin, K. A. Cation, and M. R. Broten, *Micro Paver Concept and Development Airport Pavement Management System*. Mobile, AL, USA: Computer Programs, 1987.

[57] J. A. Prozzi and S. M. Madanat, "Development of pavement performance models by combining experimental and field data," *J. Infrastruct. Syst.*, vol. 10, no. 1, pp. 9–22, Mar. 2004.

[58] H. Kerali, J. B. Odoki, and E. E. Stannard, *Overview of HDM-4* (The Highway Development and Management Series). Washington, DC, USA, 2000, pp. 1–43.

[59] Q. Dong, B. Huang, and S. H. Richards, "Calibration and application of treatment performance models in a pavement management system in Tennessee," *J. Transp. Eng.*, vol. 141, no. 2, Feb. 2015, Art. no. 04014076.

[60] A. Garciadiaz and M. Riggins, "Serviceability and distress methodology for predicting pavement performance," *Transp. Res. Rec.*, vol. 997, pp. 56–61, Jan. 1984.

[61] R. E. Smith *et al.*, *Integration of Network- and Project-Level Performance Models for TXDOT PMIS*. Boston, MA, USA: Mathematical Models, 2001.

[62] K. Wu, "Development of PCI-based pavement performance model for management of road infrastructure system," Ph.D. dissertation, Dept. Civil Eng., Arizona State Univ., Phoenix, AS, USA, 2015.

[63] F. S. Albuquerque and W. P. Núñez, "Development of roughness prediction models for low-volume road networks in northeast Brazil," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2205, no. 1, pp. 198–205, Jan. 2011.

[64] S. T. Hernán De, H. S. Priscila, and S. T. Mauricio, "Calibration of performance models for surface treatment to Chilean conditions: The HDM-4 case," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1819, no. 1, pp. 285–293, Jan. 2003.

[65] J. Li *et al.*, "The highway development and management system in Washington state: Calibration and application for the department of transportation road network," *Transp. Res. Rec.*, vol. 1933, no. 1, pp. 53–61, 2005.

[66] G. Rohde, I. Wolmarans, and E. Sadzik, "The calibration and validation of HDM performance models in the Gauteng PMS," in *Proc. SATC*, 2002, pp. 1–11.

[67] G. T. Rohde *et al.*, "The calibration and use of HDM-IV performance models in a pavement management system," in *Proc. 4th Int. Conf. Manag. Pavements*, 1998, pp. 1312–1331.

[68] P. Sebaaly *et al.*, "Nevada's approach to pavement management," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1524, pp. 109–117, Jan. 1996.

[69] S. Hussan, M. A. Kamal, I. Hafeez, N. Ahmad, S. Khanzada, and S. Ahmed, "Modelling asphalt pavement analyzer rut depth using different statistical techniques," *Road Mater. Pavement Des.*, vol. 21, no. 1, pp. 117–142, Jan. 2020.

[70] W. Wang, Y. Qin, X. Li, D. Wang, and H. Chen, "Comparisons of faulting-based pavement performance prediction models," *Adv. Mater. Sci. Eng.*, vol. 2017, pp. 1–9, Sep. 2017.

[71] W. Zhang and P. L. Durango-Cohen, "Explaining heterogeneity in pavement deterioration: Clusterwise linear regression model," *J. Infrastruct. Syst.*, vol. 20, no. 2, Jun. 2014, Art. no. 04014005.

[72] Z. Luo and H. Yin, "Probabilistic analysis of pavement distress ratings with the clusterwise regression method," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2084, no. 1, pp. 38–46, Jan. 2008.

[73] Z. Luo and E. Y. J. Chou, "Pavement condition prediction using clusterwise regression," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1974, no. 1, pp. 70–77, Jan. 2006.

[74] M. Khadka, A. Paz, and A. Singh, "Generalised clusterwise regression for simultaneous estimation of optimal pavement clusters and performance models," *Int. J. Pavement Eng.*, vol. 21, no. 9, pp. 1122–1134, Jul. 2020.

[75] M. Khadka, A. Paz, C. Arteaga, and D. K. Hale, "Simultaneous generation of optimum pavement clusters and associated performance models," *Math. Problems Eng.*, vol. 2018, pp. 1–17, Dec. 2018.

[76] M. Khadka and A. Paz, "Comprehensive clusterwise linear regression for pavement management systems," *J. Transp. Eng. B, Pavements*, vol. 143, no. 4, Dec. 2017, Art. no. 04017014.

[77] N. O. Attoh-Okine, S. Mensah, and M. Nawaiseh, "A new technique for using multivariate adaptive regression splines (MARS) in pavement roughness prediction," *Proc. Inst. Civil Eng. Transp.*, vol. 156, no. 1, pp. 51–55, Feb. 2003.

[78] U. Kırbaş and M. Karaşahin, "Performance models for hot mix asphalt pavements in urban roads," *Construct. Building Mater.*, vol. 116, pp. 281–288, Jul. 2016.

[79] U. Kırbaş, "IRI sensitivity to the influence of surface distress on flexible pavements," *Coatings*, vol. 8, no. 8, p. 271, Aug. 2018.

[80] H. Gong, Y. Sun, and B. Huang, "Estimating asphalt concrete modulus of existing flexible pavements for mechanistic-empirical rehabilitation analyses," *J. Mater. Civil Eng.*, vol. 31, no. 11, Nov. 2019, Art. no. 04019252.

[81] K. Alland, J. M. Vandenbossche, and J. Brigham, "Statistical model to detect voids for curled or warped concrete pavements," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2639, no. 1, pp. 28–38, Jan. 2017.

[82] P. Marcelino, M. D. L. Antunes, and E. Fortunato, "Comprehensive performance indicators for road pavement condition assessment," *Struct. Infrastruct. Eng.*, vol. 14, no. 11, pp. 1433–1445, Nov. 2018.

[83] K. C. P. Wang and Q. Li, "Pavement smoothness prediction based on fuzzy and gray theories," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 26, no. 1, pp. 69–76, Dec. 2009.

[84] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis*. Boca Raton, FL, USA: CRC Press, 2011.

[85] N.-D. Hoang, "Automatic detection of asphalt pavement raveling using image texture based feature extraction and stochastic gradient descent logistic regression," *Autom. Construct.*, vol. 105, Sep. 2019, Art. no. 102843.

[86] Q. Dong, C. Dong, and B. Huang, "Statistical analyses of field serviceability of throw-and-roll pothole patches," *J. Transp. Eng.*, vol. 141, no. 9, Sep. 2015, Art. no. 04015017.

[87] W. Zhang *et al.*, "Development of predictive models for initiation and propagation of field transverse cracking," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2524, no. 1, pp. 92–99, Jan. 2015.

[88] N. Alaswadko, R. Hassan, D. Meyer, and B. Mohammed, "Probabilistic prediction models for crack initiation and progression of spray sealed pavements," *Int. J. Pavement Eng.*, vol. 20, no. 1, pp. 1–11, Jan. 2019.

[89] S. Shen, W. Zhang, L. Shen, and H. Huang, "A statistical based framework for predicting field cracking performance of asphalt pavements: Application to top-down cracking prediction," *Construct. Building Mater.*, vol. 116, pp. 226–234, Jul. 2016.

[90] R. Hassan, O. Lin, and A. Thananjeyan, "Probabilistic modelling of flexible pavement distresses for network management," *Int. J. Pavement Eng.*, vol. 18, no. 3, pp. 216–227, Mar. 2017.

[91] X. Luo *et al.*, "Factor analysis of maintenance decisions for warranty pavement projects using mixed-effects logistic regression," *Int. J. Pavement Eng.*, pp. 1–12, Jan. 2020, doi: 10.1080/10298436.2020.1766039.

[92] Y. Wang, "Ordinal logistic regression model for predicting AC overlay cracking," *J. Perform. Constructed Facilities*, vol. 27, no. 3, pp. 346–353, Jun. 2013.

[93] D. Chen, T. Cavalline, and N. Mastin, "Development of piecewise linear performance models for flexible pavements using PMS data," *J. Perform. Constructed Facilities*, vol. 29, no. 6, Dec. 2015, Art. no. 04014148.

[94] D. Rys, J. Judycki, M. Pszczola, M. Jaczewski, and L. Mejlun, "Comparison of low-temperature cracks intensity on pavements with high modulus asphalt concrete and conventional asphalt concrete bases," *Construct. Building Mater.*, vol. 147, pp. 478–487, Aug. 2017.

[95] R. Hassan, O. Lin, and A. Thananjeyan, "A comparison between three approaches for modelling deterioration of five pavement surfaces," *Int. J. Pavement Eng.*, vol. 18, no. 1, pp. 26–35, Jan. 2017.

[96] N. Sanabria, V. Valentin, and S. Bogus, "Comparing neural networks and ordered probit models for forecasting pavement condition in New Mexico," presented at the Transp. Res. Board 96th Annu. Meeting, Washington, DC, USA, Jan. 2017.

[97] M. S. Yamany and D. M. Abraham, "Hybrid approach to incorporate preventive maintenance effectiveness into probabilistic pavement performance models," *J. Transp. Eng. B, Pavements*, vol. 147, no. 1, Mar. 2021, Art. no. 04020077.

[98] S. McNeil and F. Humplick, "Evaluation of errors in automated pavement-distress data acquisition," *J. Transp. Eng.*, vol. 117, no. 2, pp. 224–241, Mar. 1991.

[99] A. Pozarycki, P. Górnas, and J. Fengier, "Pavement fatigue degradation phenomenon assessment based on multi-load FWD data and stochastic process evaluation," *Periodica Polytechnica, Civil Eng.*, vol. 60, no. 4, pp. 471–477, 2016.

[100] H.-W. Ker, Y.-H. Lee, and C.-H. Lin, "Prediction models for transverse cracking of jointed concrete pavements: Development with long-term pavement performance database," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2068, no. 1, pp. 20–31, Jan. 2008.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[101] H.-W. Ker, Y.-H. Lee, and P.-H. Wu, "Development of fatigue cracking prediction models using long-term pavement performance database," *J. Transp. Eng.*, vol. 134, no. 11, pp. 477–482, Nov. 2008.
[102] S. Madanat, S. Bulusu, and A. Mahmoud, "Estimation of infrastructure distress initiation and progression models," *J. Infrastruct. Syst.*, vol. 1, no. 3, pp. 146–150, Sep. 1995.
[103] Q. Dong, X. Jiang, B. Huang, and S. H. Richards, "Analyzing influence factors of transverse cracking on LTPP resurfaced asphalt pavements through NB and ZINB models," *J. Transp. Eng.*, vol. 139, no. 9, pp. 889–895, Sep. 2013.
[104] A. Pantuso, G. W. Flintsch, S. W. Katicha, and G. Loprencipe, "Development of network-level pavement deterioration curves using the linear empirical Bayes approach," *Int. J. Pavement Eng.*, vol. 22, no. 6, pp. 780–793, May 2021.
[105] R. Winfrey, *Economic Analysis for Highways*. Scranton, PA, USA: International Textbook Co, 1969.
[106] W. Paterson and A. D. Chesher, "On predicting pavement surface distress with empirical models of failure times," *Transp. Res. Rec.*, no. 1095, pp. 45–56, 1986.
[107] J. A. Prozzi and S. M. Madanat, "Using duration models to analyze experimental pavement failure data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1699, no. 1, pp. 87–94, Jan. 2000.
[108] H. C. Shin and S. Madanat, "Development of a stochastic model of pavement distress initiation," *Doboku Gakkai Ronbunshu*, vol. 2003, no. 744, pp. 61–67, 2003.
[109] S. Inkoom and J. O. Sobanjo, "Competing risks models for the deterioration of highway pavement subject to hurricane events," *Struct. Infrastruct. Eng.*, vol. 15, no. 6, pp. 837–850, Jun. 2019.
[110] N. G. Gharaibeh and M. I. Darter, "Probabilistic analysis of highway pavement life for Illinois," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1823, no. 1, pp. 111–120, Jan. 2003.
[111] S. A. Romanoschi and J. B. Metcalf, "Evaluation of probability distribution function for the life of pavement structures," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1730, no. 1, pp. 91–98, Jan. 2000.
[112] X. Luo, F. Wang, N. Wang, X. Qiu, and J. Tao, "Evaluation of warranty and nonwarranty pavement performance using survival analysis," *J. Transp. Eng. B, Pavements*, vol. 146, no. 1, Mar. 2020, Art. no. 04019035.
[113] Q. Dong and B. Huang, "Evaluation of influence factors on crack initiation of LTPP resurfaced-asphalt pavements using parametric survival analysis," *J. Perform. Constructed Facilities*, vol. 28, no. 2, pp. 412–421, Apr. 2014.
[114] B.-W. Tsai, J. T. Harvey, and C. L. Monismith, "Application of Weibull theory in prediction of asphalt concrete fatigue performance," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1832, no. 1, pp. 121–130, Jan. 2003.
[115] B. Ebrahimi, H. Wallbaum, K. Svensson, and D. Gryteselv, "Estimation of Norwegian asphalt surfacing lifetimes using survival analysis coupled with road spatial data," *J. Transp. Eng. B, Pavements*, vol. 145, no. 3, Sep. 2019, Art. no. 04019017.
[116] J. Yu, E. Y. Chou, and J.-T. Yau, "Estimation of the effects of influential factors on pavement service life with Cox proportional hazards method," *J. Infrastruct. Syst.*, vol. 14, no. 4, pp. 275–282, Dec. 2008.
[117] A. Vargas-Nordcbeck and F. Jalali, "Life-extending benefit of crack sealing for pavement preservation," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 1, pp. 272–281, Jan. 2020.
[118] K. Svenson, "Estimated lifetimes of road pavements in Sweden using time-to-event analysis," *J. Transp. Eng.*, vol. 140, no. 11, Nov. 2014, Art. no. 04014056.
[119] F. Jalali, A. Vargas-Nordcbeck, and M. Nakhaei, "Role of preventive treatments in low-volume road maintenance program: Full-scale case study," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 12, pp. 855–862, Dec. 2019.
[120] K. Svenson, Y. Li, Z. Macuchova, and L. Rönnegård, "Evaluating needs of road maintenance in Sweden with the mixed proportional hazards model," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2589, no. 1, pp. 51–58, Jan. 2016.
[121] Q. Dong and B. Huang, "Failure probability of resurfaced preventive maintenance treatments: Investigation into long-term pavement performance program," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2481, no. 1, pp. 65–74, Jan. 2015.
[122] Y. Wang, K. C. Mahboub, and D. E. Hancher, "Survival analysis of fatigue cracking for flexible pavements based on long-term pavement performance data," *J. Transp. Eng.*, vol. 131, no. 8, pp. 608–616, Aug. 2005.
[123] L. Li, S. I. Guler, and E. T. Donnell, "Pavement friction degradation based on Pennsylvania field test data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2639, no. 1, pp. 11–19, Jan. 2017.
[124] C. Chen *et al.*, "Assessment of composite pavement performance by survival analysis," *J. Transp. Eng.*, vol. 141, no. 9, Sep. 2015, Art. no. 04015018.
[125] T. S. Kiessé, T. Lorino, and H. Khraibani, "Discrete nonparametric kernel and parametric methods for the modeling of pavement deterioration," *Commun. Statist. Theory Methods*, vol. 43, no. 6, pp. 1164–1178, Mar. 2014.
[126] L. Gao, J. P. Aguiar-Moya, and Z. Zhang, "Bayesian analysis of heterogeneity in modeling of pavement fatigue cracking," *J. Comput. Civil Eng.*, vol. 26, no. 1, pp. 37–43, Jan. 2012.
[127] V. Donev and M. Hoffmann, "Condition prediction and estimation of service life in the presence of data censoring and dependent competing risks," *Int. J. Pavement Eng.*, vol. 20, no. 3, pp. 313–331, Mar. 2019.
[128] Z. Li, "A probabilistic and adaptive approach to modeling performance of pavement infrastructure," Ph.D. dissertation, Dept. Center Transp. Res., Univ. Texas Austin, Austin, TX, USA, 2005.
[129] J. V. Camahan, W. J. Davis, M. Y. Shahin, P. L. Keane, and M. I. Wu, "Optimal maintenance decisions for pavement management," *J. Transp. Eng.*, vol. 113, no. 5, pp. 554–572, Sep. 1987.
[130] S. Madanat, R. Mishalani, and W. H. W. Ibrahim, "Estimation of infrastructure transition probabilities from condition rating data," *J. Infrastruct. Syst.*, vol. 1, no. 2, pp. 120–125, Jun. 1995.
[131] K. A. Abaza, "Stochastic approach for design of flexible pavement: A case study for low volume roads," *Road Mater. Pavement Des.*, vol. 12, no. 3, pp. 663–685, Jan. 2011.
[132] K. A. Abaza, "Back-calculation of transition probabilities for Markovian-based pavement performance prediction models," *Int. J. Pavement Eng.*, vol. 17, no. 3, pp. 253–264, Mar. 2016.
[133] K. A. Abaza, S. A. Ashur, and I. A. Al-Khatib, "Integrated pavement management system with a Markovian prediction model," *J. Transp. Eng.*, vol. 130, no. 1, pp. 24–33, Jan. 2004.
[134] K. A. Abaza, "Iterative linear approach for nonlinear nonhomogenous stochastic pavement management models," *J. Transp. Eng.*, vol. 132, no. 3, pp. 244–256, Mar. 2006.
[135] J. Yang, M. Gunaratne, J. J. Lu, and B. Dietrich, "Use of recurrent Markov chains for modeling the crack performance of flexible pavements," *J. Transp. Eng.*, vol. 131, no. 11, pp. 861–872, Nov. 2005.
[136] A. V. Moreira, J. Tinoco, J. R. M. Oliveira, and A. Santos, "An application of Markov chains to predict the evolution of performance indicators based on pavement historical data," *Int. J. Pavement Eng.*, vol. 19, no. 10, pp. 937–948, Oct. 2018.
[137] S. Nasseri, M. Gunaratne, J. Yang, and A. Nazef, "Application of improved crack prediction methodology in Florida's highway network," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2093, no. 1, pp. 67–75, Jan. 2009.
[138] P. Saha, K. Ksaibati, and R. Atadero, "Developing pavement distress deterioration models for pavement management system using Markovian probabilistic process," *Adv. Civil Eng.*, vol. 2017, pp. 1–9, Jan. 2017.
[139] S. Alimoradi, A. Golroo, and S. M. Asgharzadeh, "Development of pavement roughness master curves using Markov chain," *Int. J. Pavement Eng.*, pp. 1–11, 2020, doi: 10.1080/10298436.2020.1752917.
[140] H. Pérez-Acebo, N. Mindra, A. Railean, and E. Rojí, "Rigid pavement performance models by means of Markov chains with half-year step time," *Int. J. Pavement Eng.*, vol. 20, no. 7, pp. 830–843, Jul. 2019.
[141] O. Thomas and J. Sobanjo, "Comparison of Markov chain and semi-Markov models for crack deterioration on flexible pavements," *J. Infrastruct. Syst.*, vol. 19, no. 2, pp. 186–195, Jun. 2013.
[142] A. Golroo and S. L. Tighe, "Development of pervious concrete pavement performance models using expert opinions," *J. Transp. Eng.*, vol. 138, no. 5, pp. 634–648, May 2012.
[143] A. Shah, S. Tighe, and A. Stewart, "Development of a unique deterioration index, prioritization methodology, and foreign object damage evaluation models for Canadian airfield pavement management," *Can. J. Civil Eng.*, vol. 31, no. 4, pp. 608–618, Aug. 2004.
[144] D. Chimba, E. Kidando, and M. Onyango, "Evaluating the service life of thermoplastic pavement markings: Stochastic approach," *J. Transp. Eng. B, Pavements*, vol. 144, no. 3, Sep. 2018, Art. no. 04018029.
[145] M. Putu *et al.*, "A stochastic-based performance prediction model for road network pavement maintenance," *Road Transp. Res., A J. Austral. New Zealand Res. Pract.*, vol. 21, no. 3, p. 34, 2012.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[146] H. Gao and X. Zhang, "A Markov-based road maintenance optimization model considering user costs," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 28, no. 6, pp. 451–464, Jul. 2013.
[147] L. Yao, Q. Dong, J. Jiang, and F. Ni, "Deep reinforcement learning for long-term pavement maintenance planning," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 35, no. 11, pp. 1230–1245, Nov. 2020.
[148] J. Q. Retherford and M. McDonald, "Unified approach for uncertainty analysis using the AASHTO mechanistic-empirical pavement design guide," *J. Transp. Eng.*, vol. 138, no. 5, pp. 657–664, May 2012.
[149] N. Karballaeezadeh, H. G. Tehrani, D. M. Shadmehri, and S. Shamshirband, "Estimation of flexible pavement structural capacity using machine learning techniques," *Frontiers Struct. Civil Eng.*, vol. 14, no. 5, pp. 1083–1096, Oct. 2020.
[150] A. S. Hosseini, P. Hajikarimi, M. Gandomi, F. M. Nejad, and A. H. Gandomi, "Optimized machine learning approaches for the prediction of viscoelastic behavior of modified asphalt binders," *Construct. Building Mater.*, vol. 299, Sep. 2021, Art. no. 124264.
[151] L. Titus-Glover, "Unsupervised extraction of patterns and trends within highway systems condition attributes data," *Adv. Eng. Informat.*, vol. 42, Oct. 2019, Art. no. 100990.
[152] Z. Luo, "Pavement performance modelling with an auto-regression approach," *Int. J. Pavement Eng.*, vol. 14, no. 1, pp. 85–94, Jan. 2013.
[153] M. Fang, C. Han, Y. Xiao, Z. Han, S. Wu, and M. Cheng, "Prediction modelling of rutting depth index for asphalt pavement using de-noising method," *Int. J. Pavement Eng.*, vol. 21, no. 7, pp. 895–907, Jun. 2020.
[154] C.-Y. Chu and P. L. Durango-Cohen, "Estimation of infrastructure performance models using state-space specifications of time series models," *Transp. Res. C, Emerg. Technol.*, vol. 15, no. 1, pp. 17–32, Feb. 2007.
[155] H. M. Javad, N. Akbar, and A. Seyedjalil, "Pavement deterioration modeling for forest roads based on logistic regression and artificial neural networks," *Croatian J. Forest Eng.*, vol. 39, no. 2, pp. 271–287, Jan., 2018.
[156] Z. Wang *et al.*, "Prediction of highway asphalt pavement performance based on Markov chain and artificial neural network approach," *J. Supercomput., An Int. J. High-Perform. Comput. Design, Anal., Use*, vol. 77, no. 2, pp. 1–23, 2020.
[157] H. Ziari, J. Sobhani, J. Ayoubinejad, and T. Hartmann, "Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods," *Int. J. Pavement Eng.*, vol. 17, no. 9, pp. 776–788, Oct. 2016.
[158] H. Gong, Y. Sun, W. Hu, and B. Huang, "Neural networks for fatigue cracking prediction using outputs from pavement mechanistic-empirical design," *Int. J. Pavement Eng.*, vol. 22, no. 2, pp. 162–172, Jan. 2021.
[159] J. Yang, J. J. Lu, M. Gunaratne, and B. Dietrich, "Modeling crack deterioration of flexible pavements: Comparison of recurrent Markov chains and artificial neural networks," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1974, no. 1, pp. 18–25, Jan. 2006.
[160] S. Chandra, C. R. Sekhar, A. K. Bharti, and B. Kangadurai, "Relationship between pavement roughness and distress parameters for Indian highways," *J. Transp. Eng.*, vol. 139, no. 5, pp. 467–475, May 2013.
[161] N. N. Eldin and A. B. Senouci, "Use of neural networks for condition rating of jointed concrete pavements," *Adv. Eng. Softw.*, vol. 23, no. 3, pp. 133–141, Jan. 1995.
[162] F. Gu, X. Luo, Y. Zhang, Y. Chen, R. Luo, and R. L. Lytton, "Prediction of geogrid-reinforced flexible pavement performance using artificial neural network approach," *Road Mater. Pavement Des.*, vol. 19, no. 5, pp. 1147–1163, Jul. 2018.
[163] M. Ling, X. Luo, S. Hu, F. Gu, and R. L. Lytton, "Numerical modeling and artificial neural network for predicting J-Integral of top-down cracking in asphalt pavement," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2631, no. 1, pp. 83–95, Jan. 2017.
[164] H. Gong, Y. Sun, Z. Mei, and B. Huang, "Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks," *Construct. Building Mater.*, vol. 190, pp. 710–718, Nov. 2018.
[165] L. Yao, Q. Dong, J. Jiang, and F. Ni, "Establishment of prediction models of asphalt pavement performance based on a novel data calibration method and neural network," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 1, pp. 66–82, Jan. 2019.
[166] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, May 2017.
[167] H. Nhat-Duc, Q.-L. Nguyen, and V.-D. Tran, "Automatic recognition of asphalt pavement cracks using Metaheuristic optimized edge detection algorithms and convolution neural network," *Autom. Construct.*, vol. 94, pp. 203–213, Oct. 2018.
[168] F. M. Nejad and H. Zakeri, "An optimum feature extraction method based on Wavelet–Radon transform and dynamic neural network for pavement distress classification," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9442–9460, Aug. 2011.
[169] Y. Hou *et al.*, "The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis," *Engineering*, vol. 7, no. 6, pp. 845–856, Jun. 2021.
[170] Z. Tong, J. Gao, A. Sha, L. Hu, and S. Li, "Convolutional neural network for asphalt pavement surface texture analysis," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1056–1072, Dec. 2018.
[171] W. Ye, W. Jiang, Z. Tong, D. Yuan, and J. Xiao, "Convolutional neural network for pothole detection in asphalt pavement," *Road Mater. Pavement Des.*, vol. 22, no. 1, pp. 42–58, Jan. 2021.
[172] Z. Tong, J. Gao, and H. Zhang, "Innovative method for recognizing subgrade defects based on a convolutional neural network," *Construct. Building Mater.*, vol. 169, pp. 69–82, Apr. 2018.
[173] J. Zhang, X. Yang, W. Li, S. Zhang, and Y. Jia, "Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method," *Autom. Construct.*, vol. 113, May 2020, Art. no. 103119.
[174] Y. Hou *et al.*, "MobileCrack: Object classification in asphalt pavements using an adaptive lightweight deep learning," *J. Transp. Eng. B, Pavements*, vol. 147, no. 1, Mar. 2021, Art. no. 04020092.
[175] F. Yang, L. Zhang, S. Yu, D. V. Prokhorov, X. Mei, and H. Ling, "Feature pyramid and hierarchical boosting network for pavement crack detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1525–1535, Apr. 2020.
[176] Y. Du *et al.*, "Pavement distress detection and classification based on YOLO network," *Int. J. Pavement Eng.*, vol. 22, no. 13, pp. 1659–1672, Nov. 2021.
[177] V. P. Tran, T. S. Tran, H. J. Lee, K. D. Kim, J. Baek, and T. T. Nguyen, "One stage detector (RetinaNet)-based crack detection for asphalt pavements considering pavement distresses and surface objects," *J. Civil Struct. Health Monitor.*, vol. 11, no. 1, pp. 205–222, Feb. 2021.
[178] Z. Tong, D. Yuan, J. Gao, Y. Wei, and H. Dou, "Pavement-distress detection using ground-penetrating radar and network in networks," *Construct. Building Mater.*, vol. 233, Feb. 2020, Art. no. 117352.
[179] B. Yu, X. Meng, and Q. Yu, "Automated pixel-wise pavement crack detection by classification-segmentation networks," *J. Transp. Eng. B, Pavements*, vol. 147, no. 2, Jun. 2021, Art. no. 04021005.
[180] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using fully convolutional network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1090–1109, Dec. 2018.
[181] S. L. H. Lau, E. K. P. Chong, X. Yang, and X. Wang, "Automated pavement crack segmentation using U-Net-based convolutional neural network," *IEEE Access*, vol. 8, pp. 114892–114899, 2020.
[182] J. Huyan, W. Li, S. Tighe, Z. Xu, and J. Zhai, "CrackU-Net: A novel deep convolutional neural network for pixelwise pavement crack detection," *Struct. Control Health Monitor.*, vol. 27, no. 8, Aug. 2020, Art. no. e2551.
[183] Y. Zhang, B. Chen, J. Wang, J. Li, and X. Sun, "APLCNet: Automatic pixel-level crack detection network based on instance segmentation," *IEEE Access*, vol. 8, pp. 199159–199170, 2020.
[184] H. Huang *et al.*, "Deep learning-based instance segmentation of cracks from shield tunnel lining images," *Struct. Infrastruct. Eng.*, pp. 1–14, 2020, doi: 10.1080/15732479.2020.1838559.
[185] N. Tabatabaee, M. Ziyadi, and Y. Shafahi, "Two-stage support vector classifier and recurrent neural network predictor for pavement performance modeling," *J. Infrastruct. Syst.*, vol. 19, no. 3, pp. 266–274, Sep. 2013.
[186] S. Choi and M. Do, "Development of the road pavement deterioration model based on the deep learning method," *Electronics*, vol. 9, no. 1, p. 3, Dec. 2019.
[187] A. Zhang *et al.*, "Automated pixel-level pavement crack detection on 3D asphalt surfaces with a recurrent neural network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 34, no. 3, pp. 213–229, Mar. 2019.
[188] M. Arifuzzaman, U. Gazder, M. S. Alam, O. Sirin, and A. A. Mamun, "Modelling of Asphalt's adhesive behaviour using classification and regression tree (CART) analysis," *Comput. Intell. Neurosci.*, vol. 2019, pp. 1–7, Aug. 2019.
[189] Q. Dong, X. Chen, B. Huang, and X. Gu, "Analysis of the influence of materials and construction practices on slurry seal performance using LTPP data," *J. Transp. Eng. B, Pavements*, vol. 144, no. 4, Dec. 2018, Art. no. 04018046.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[190] X. Wang, D. Montgomery, and E. Owusu-Antwi, "Application of regression trees to LTPP data analysis," in *Proc. Int. Conf. Highway Pavement Data, Anal. Mechanistic Design Appl.* Columbus, OH, USA: Ohio Res. Inst. Transp. Environ. Ohio Univ., 2003, pp. 323–333.

[191] M. Kang, M. Kim, and J. H. Lee, "Analysis of rigid pavement distresses on interstate highway using decision tree algorithms," *KSCE J. Civil Eng.*, vol. 14, no. 2, pp. 123–130, Mar. 2010.

[192] H. Gong, Y. Sun, X. Shu, and B. Huang, "Use of random forests regression for predicting IRI of asphalt pavements," *Construct. Building Mater.*, vol. 189, pp. 890–897, Nov. 2018.

[193] H. Gong, Y. Sun, W. Hu, P. A. Polaczyk, and B. Huang, "Investigating impacts of asphalt mixture properties on pavement performance using LTPP data through random forests," *Construct. Building Mater.*, vol. 204, pp. 203–212, Apr. 2019.

[194] P. Nitsche, R. Stütz, M. Kammer, and P. Maurer, "Comparison of machine learning methods for evaluating pavement roughness based on vehicle response," *J. Comput. Civil Eng.*, vol. 28, no. 4, Jul. 2014, Art. no. 04014015.

[195] A. Ashrafian *et al.*, "Classification-based regression models for prediction of the mechanical properties of roller-compacted concrete pavement," *Appl. Sci.*, vol. 10, no. 11, p. 3707, May 2020.

[196] Y. Pan, X. Zhang, G. Cervone, and L. Yang, "Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3701–3712, Oct. 2018.

[197] H. Gong *et al.*, "An efficient and robust method for predicting asphalt concrete dynamic modulus," *Int. J. Pavement Eng.*, pp. 1–12, 2021, doi: 10.1080/10298436.2020.1865533.

[198] R. Guo, D. Fu, and G. Sollazzo, "An ensemble learning model for asphalt pavement performance prediction based on gradient boosting decision tree," *Int. J. Pavement Eng.*, pp. 1–14, 2021, doi: 10.1080/10298436.2021.1910825.

[199] H. Gong, Y. Sun, and B. Huang, "Gradient boosted models for enhancing fatigue cracking prediction in mechanistic-empirical pavement design guide," *J. Transp. Eng. B, Pavements*, vol. 145, no. 2, Jun. 2019, Art. no. 04019014.

[200] L. Barua, B. Zou, M. Noruzoliaee, and S. Derrible, "A gradient boosting approach to understanding airport runway and taxiway pavement deterioration," *Int. J. Pavement Eng.*, vol. 22, no. 13, pp. 1673–1687, Jan. 2020.

[201] K. Gopalakrishnan and S. Kim, "Support vector machines approach to HMA stiffness prediction," *J. Eng. Mech.*, vol. 137, no. 2, pp. 138–146, Feb. 2011.

[202] P. Georgiou, C. Plati, and A. Loizos, "Soft computing models to predict pavement roughness: A comparative study," *Adv. Civil Eng.*, vol. 2018, pp. 1–8, Jul. 2018.

[203] H. Ziari, M. Maghrebi, J. Ayoubinejad, and S. T. Waller, "Prediction of pavement performance: Application of support vector regression with different kernels," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2589, no. 1, pp. 135–145, Jan. 2016.

[204] N. Kargah-Ostadi and S. M. Stoffels, "Framework for development and comprehensive comparison of empirical pavement performance models," *J. Transp. Eng.*, vol. 141, no. 8, Aug. 2015, Art. no. 04015012.

[205] N. Karballaeezadeh, D. Mohammadzadeh S, S. Shamshirband, P. Hajikhodaverdikhan, A. Mosavi, and K.-W. Chau, "Prediction of remaining service life of pavement using an optimized support vector machine (case study of Semnan–Firuzkuh road)," *Eng. Appl. Comput. Fluid Mech.*, vol. 13, no. 1, pp. 188–198, Jan. 2019.

[206] N.-D. Hoang and Q.-L. Nguyen, "A novel method for asphalt pavement crack classification based on image processing and machine learning," *Eng. With Comput.*, vol. 35, no. 2, pp. 487–498, Apr. 2019.

[207] N.-D. Hoang and Q.-L. Nguyen, "Automatic recognition of asphalt pavement cracks based on image processing and machine learning approaches: A comparative study on classifier performance," *Math. Problems Eng.*, vol. 2018, pp. 1–16, Jul. 2018.

[208] R. B. Mallick, N. M. Kottayi, R. K. Veeraragavan, E. Dave, C. DeCarlo, and J. E. Sias, "Suitable tests and machine learning approach to predict moisture susceptibility of hot-mix asphalt," *J. Transp. Eng. B, Pavements*, vol. 145, no. 3, Sep. 2019, Art. no. 04019030.

[209] S. M. Piryonesi and T. E. El-Diraby, "Role of data analytics in infrastructure asset management: Overcoming data size and quality problems," *J. Transp. Eng. B, Pavements*, vol. 146, no. 2, Jun. 2020, Art. no. 04020022.

[210] S. Inkoom, J. Sobanjo, A. Barbu, and X. Niu, "Pavement crack rating using machine learning frameworks: Partitioning, bootstrap forest, boosted trees, Naïve Bayes, and k-nearest neighbors," *J. Transp. Eng. B, Pavements*, vol. 145, no. 3, Sep. 2019, Art. no. 04019031.

[211] M. S. Kaseko, Z. Lo, and S. G. Ritchie, "Comparison of traditional and neural classifiers for pavement-crack detection," *J. Transp. Eng.*, vol. 120, no. 4, pp. 552–569, Jul. 1994.

[212] S. Mokhtari, L. Wu, and H.-B. Yun, "Comparison of supervised classification techniques for vision-based pavement crack detection," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2595, no. 1, pp. 119–127, Jan. 2016.

[213] M. Castro and J. A. Sánchez, "Fatigue and healing of asphalt mixtures: Discriminate analysis of fatigue curves," *J. Transp. Eng.*, vol. 132, no. 2, pp. 168–174, Feb. 2006.

[214] F. Sayyady, J. R. Stone, K. L. Taylor, F. M. Jadoun, and Y. R. Kim, "Clustering analysis to characterize mechanistic–empirical pavement design guide traffic data in north Carolina," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2160, no. 1, pp. 118–127, Jan. 2010.

[215] W. Wang, S. Wang, D. Xiao, S. Qiu, and J. Zhang, "An unsupervised cluster method for pavement grouping based on multidimensional performance data," *J. Transp. Eng. B, Pavements*, vol. 144, no. 2, Jun. 2018, Art. no. 04018005.

[216] C. W. Yi, Y. T. Chuang, and C. S. Nian, "Toward crowdsourcing-based road pavement monitoring by mobile sensing technologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1905–1917, Aug. 2015.

[217] L. Li, W. Luo, K. Wang, G. Liu, and C. Zhang, "Automatic groove measurement and evaluation with high resolution laser profiling data," *Sensors*, vol. 18, no. 8, p. 2713, Aug. 2018.

[218] J. David *et al.*, "Detection of road pavement quality using statistical clustering methods," *J. Intell. Inf. Syst., Integrating Artif. Intell. Database Technol.*, vol. 54, no. 3, p. 483, 2020.

[219] Y. Jiao, Y. Zhang, M. Zhang, L. Fu, and L. Zhang, "Investigation of fracture modes in pervious asphalt under splitting and compression based on acoustic emission monitoring," *Eng. Fract. Mech.*, vol. 211, pp. 209–220, Apr. 2019.

[220] C. Jing, J. Zhang, and B. Song, "An innovative evaluation method for performance of in-service asphalt pavement with semi-rigid base," *Construct. Building Mater.*, vol. 235, Feb. 2020, Art. no. 117376.

[221] A. Bianchini, "Pavement maintenance planning at the network level with principal component analysis," *J. Infrastruct. Syst.*, vol. 20, no. 2, Jun. 2014, Art. no. 04013013.

[222] P. Ghasemi *et al.*, "Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: Elimination of correlated inputs and extrapolation in modeling," *Struct. Multidisciplinary Optim.*, vol. 59, no. 4, p. 1335, 2019.

[223] Y. Zhang, J. G. McDaniel, and M. L. Wang, "Estimation of pavement macrotexture by principal component analysis of acoustic measurements," *J. Transp. Eng.*, vol. 140, no. 2, Feb. 2014, Art. no. 04013004.

[224] P. Tian, A. Shukla, L. Nie, G. Zhan, and S. Liu, "Characteristics' relation model of asphalt pavement performance based on factor analysis," *Int. J. Pavement Res. Technol.*, vol. 11, no. 1, pp. 1–12, Jan. 2018.

[225] X. Chen, Q. Dong, H. Zhu, B. Huang, and E. G. Burdette, "Contributions of condition measurements on the latent pavement condition by confirmatory factor analysis," *Transportmetrica A, Transp. Sci.*, vol. 15, no. 1, pp. 2–17, Feb. 2019.

[226] M. Ben-Akiva and R. Ramaswamy, "An approach for predicting latent infrastructure facility deterioration," *Transp. Sci.*, vol. 27, no. 2, pp. 174–193, May 1993.

[227] X. Chen, Q. Dong, H. Zhu, and B. Huang, "Development of distress condition index of asphalt pavements using LTPP data through structural equation modeling," *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 58–69, Jul. 2016.

[228] R. Ramaswamy, "Estimation of latent pavement performance from damage measurements," Ph.D. dissertation, Dept. Civil Eng., Massachusetts Inst. Technol., Cambridge, MA, USA, 1989.

[229] C.-Y. Chu and P. L. Durango-Cohen, "Incorporating maintenance effectiveness in the estimation of dynamic infrastructure performance models," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 23, no. 3, pp. 174–188, Apr. 2008.

[230] Q. Dong, X. Chen, and H. Gong, "Performance evaluation of asphalt pavement resurfacing treatments using structural equation modeling," *J. Transp. Eng. B, Pavements*, vol. 146, no. 1, Mar. 2020, Art. no. 04019043.

[231] I. Arel, C. Liu, T. Urbanik, and A. G. Kohls, "Reinforcement learning-based multi-agent system for network traffic signal control," *IET Intell. Transp. Syst.*, vol. 4, no. 2, pp. 128–135, 2010.

[232] D. Han, K. Kaito, and K. Kobayashi, "Application of Bayesian estimation method with Markov hazard model to improve deterioration forecasts for infrastructure asset management," *KSCE J. Civil Eng.*, vol. 18, no. 7, pp. 2107–2119, Nov. 2014.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.

[233] D. Han, K. Kaito, K. Kobayashi, and K. Aoki, "Management scheme of road pavements considering heterogeneous multiple life cycles changed by repeated maintenance work," *KSCE J. Civil Eng.*, vol. 21, no. 5, pp. 1747–1756, Jul. 2017.

[234] N. Tabatabaee and M. Ziyadi, "Bayesian approach to updating Markov-based models for predicting pavement performance," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2366, no. 1, pp. 34–42, Jan. 2013.

[235] A. Onar, F. Thomas, B. Choubane, and T. Byron, "Bayesian degradation modeling in accelerated pavement testing with estimated transformation parameter for the response," *J. Transp. Eng.*, vol. 133, no. 12, pp. 677–687, Dec. 2007.

[236] N. Lethanh, K. Kaito, and K. Kobayashi, "Infrastructure deterioration prediction with a Poisson hidden Markov model on time series data," *J. Infrastruct. Syst.*, vol. 21, no. 3, Sep. 2015, Art. no. 04014051.

[237] S. Inkoom, J. O. Sobanjo, E. Chicken, D. Sinha, and X. Niu, "Assessment of deterioration of highway pavement using Bayesian survival model," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 6, pp. 310–325, Jun. 2020.

[238] E. S. Park, R. E. Smith, T. J. Freeman, and C. H. Spiegelman, "A Bayesian approach for improved pavement performance prediction," *J. Appl. Statist.*, vol. 35, no. 11, pp. 1219–1238, Nov. 2008.

[239] X. Chen, Q. Dong, X. Gu, and Q. Mao, "Bayesian analysis of pavement maintenance failure probability with Markov chain Monte Carlo simulation," *J. Transp. Eng. B, Pavements*, vol. 145, no. 2, Jun. 2019, Art. no. 04019001.

[240] K. Kobayashi, K. Kaito, and N. Lethanh, "A competing Markov model for cracking prediction on civil structures," *Transp. Res. B, Methodol.*, vol. 68, pp. 345–362, Oct. 2014.

[241] D. M. Dilip and G. L. S. Babu, "Methodology for pavement design reliability and back analysis using Markov chain Monte Carlo simulation," *J. Transp. Eng.*, vol. 139, no. 1, pp. 65–74, Jan. 2013.

[242] F. Hong and J. A. Prozzi, "Estimation of pavement performance deterioration using Bayesian approach," *J. Infrastruct. Syst.*, vol. 12, no. 2, pp. 77–86, Jun. 2006.

[243] L. O. Mills and N. Attoh-Okine, "Analysis of ground penetrating radar data using hierarchical Markov chain Monte Carlo simulation," *Can. J. Civil Eng.*, vol. 41, no. 1, pp. 9–16, Jan. 2014.

[244] L. N. O. Mills, N. O. Attoh-Okine, and S. McNeil, "Hierarchical Markov chain Monte Carlo simulation for modeling transverse cracks in highway pavements," *J. Transp. Eng.*, vol. 138, no. 6, pp. 700–705, Jun. 2012.

[245] H. Wang, Z. Wang, J. Zhao, and J. Qian, "Life-cycle cost analysis of pay adjustment for initial smoothness of asphalt pavement overlay," *J. Test. Eval.*, vol. 48, no. 2, pp. 1350–1364, Mar. 2020.

[246] J. B. Nagel and B. Sudret, "Hamiltonian Monte Carlo and borrowing strength in hierarchical inverse problems," *ASCE-ASME J. Risk Uncertainty Eng. Syst. A, Civil Eng.*, vol. 2, no. 3, Sep. 2016, Art. no. B4015008.

[247] L. H. Nguyen, I. Gaudot, and J. Goulet, "Uncertainty quantification for model parameters and hidden state variables in Bayesian dynamic linear models," *Struct. Control Health Monitor.*, vol. 26, Dec. 2018, Art. no. e2309.

[248] L. Chen *et al.*, "Investigation of influential factors of tire/pavement noise: A multilevel Bayesian analysis of full-scale track testing data," *Construct. Building Mater.*, vol. 270, Feb. 2021, Art. no. 121484.

Qiao Dong (Member, IEEE) received the B.S. degree in civil engineering and the M.S. degree in roadway and railway engineering from Southeast University, Nanjing, China, in 2003 and 2006, respectively, and the Ph.D. degree in civil and environmental engineering from The University of Tennessee, Knoxville, USA, in 2011.

From 2011 to 2016, he was a Research Associate with The University of Tennessee. Since 2016, he joined Southeast University as a Professor. He has published more than 70 journal articles, and served as the Primary Investigator for over ten research projects, including the National Natural Science Foundation of China and the Natural Science Foundation of Jiangsu Province. His research interests include pavement asset management based on data analysis and artificial intelligence, pavement distress non-destructive evaluation, pavement materials multiscale characterization and simulation, and pavement performance sensor technology. He is an Active Member of the Bituminous Materials Committee (BMC) of the American Society of Civil Engineers, the Pavement Maintenance Committee (AHD20) of Transportation Research Board (TRB), and the Pavement Performance Evaluation Committee of the World Transportation Congress.

Xueqin Chen received the B.S. and Ph.D. degrees in civil engineering from Tongji University, Shanghai, China, in 2012 and 2018, respectively.

From 2014 to 2016, she was a Visiting Scholar with The University of Tennessee, Knoxville, USA. She is currently an Assistant Professor with the Department of Civil Engineering, Nanjing University of Science and Technology, Nanjing, China. Her research interests include performance evaluation and prediction of the life cycle of infrastructure, infrastructure maintenance decisions and asset management, smart infrastructure (IS3), and pavement management systems. She was the First Prize Winner of the International Contest on LTPP Data Analysis in 2015. She serves as a Reviewer for *Journal of Transportation Engineering*, *Journal of Cleaner Production*, and *Tunnelling and Underground Space Technology*.

Shi Dong received the Ph.D. degree from Chang'an University in 2019, and jointly trained by the University of Waterloo, Canada, under the supervision of Prof. Peiwen Hao and Prof. Susan L. Tighe.

Since 2019, he has been serving as a full-time Assistant Professor for the College of Transportation Engineering, Chang'an University. Since 2020, he has also been serving as the Deputy Director for the Research Center of Digital Construction and Management for Transportation Infrastructure of Shaanxi Province. Since 2021, he has been serving as the Deputy Director for the Institute of Transportation Big-data and Artificial Intelligence, Chang'an University. He has been granted or participated in ten research projects funded by governments. He has published 30 academic papers. His research interests include intelligent maintenance of highway infrastructure based on multi-source big data analyzing, semi-rigid base pavement performance prediction and rehabilitation decision, mechanics-empirical pavement design guide (MEPDG) calibration and optimization, and building information modeling (BIM) for transportation infrastructure at servicing stage. He has reviewing articles for the *International Journal of Pavement Engineering* and *Construction and Building Materials*.

Fujian Ni received the B.S., M.S., and Ph.D. degrees from Southeast University, Nanjing, China.

From 1997 to 1999, he was supported by the China Ministry to held a post-doctoral position with Nagaoka University of Technology, Japan. He is currently a Professor with Southeast University. He has served as the Primary Investigator for over 100 research projects, including the National Natural Science Foundation of China and the Natural Science Foundation of Jiangsu Province and has published more than 100 journal articles. His research interests include construction materials, pavement management systems, material research and mechanics analysis of steel deck pavement, and cold in-place recycling. He is a member of the Transportation Research Board (TRB), the China Highway Society, and the National Professional Standardization Committee of China.

Authorized licensed use limited to: Boise State University. Downloaded on May 24,2023 at 16:52:09 UTC from IEEE Xplore. Restrictions apply.