Data Mining and Performance Prediction of Flexible Road Pavement Using Fuzzy Logic Theory: A Case of Nigeria

Adekunle Taiwo Olowosulu[1], Jibrin Mohammed Kaura[1], Abdulfatai Adinoyi Murana[1], and Paul Terkumbur Adeke[2](✉)

[1] Department of Civil Engineering, Faculty of Engineering, Ahmadu Bello University Zaria, Zaria, Kaduna State, Nigeria

[2] Department of Civil Engineering, College of Engineering, Federal University of Agriculture Makurdi, Makurdi, Benue State, Nigeria

adeke.pt@outlook.com

Abstract. The over dependence on road transport system to cater for the fast growing human population in some developing countries like Nigeria has necessitated the need for the development of an efficient and sustainable road pavement management system. This study used data mining techniques namely; Random Forest, Decision Tree and Naive Bayes algorithms to examine the inferred dataset on flexible road pavement performance attributes and surface condition classification in Nigeria. The data mining techniques were used to investigate hidden relationship between pavement performance variables and to authenticate the accuracy of subjective measurements that were used for pavement surface condition classification. The Random Forest and Decision Tree algorithms reported perfect classifications of road pavement sections into; Excellent, Good, Fair, Poor and Very poor. On the other hand, the Naïve Bayes algorithm yielded inaccurate classifications with some margin of errors which were attributed to missing and noisy entries in the dataset. This necessitated the use of Fuzzy logic theory for the performance prediction due to its capability to handle the imprecise dataset. It was used to develop Fuzzy Inference System (FIS) for performance prediction of flexible road pavement using attributes such as; the classified Initial Pavement Condition (IPC), Age of pavement, Resilient Modulus (M_R) of subgrade soil, Average Truck load per day, Average Annual Air Temperature and Rainfall to predict the Future Pavement Condition (FPC). The model was calibrated using the observed logical behaviour of road pavement to fit the engineering experience and judgement. A goodness-of-fit test between the observed and predicted FPC values showed high level of consistency – correlation coefficient at 90%. The research proposed 5120 mutually exclusive Fuzzy Logic Rules for performance prediction of road pavement based on permutation theory. Though, the required well-spread dataset for calibration of the model to cover all possible pavement conditions in Nigeria and subsequent validation were not available, a framework for performance prediction of flexible road pavement was developed, and a comprehensive guidelines on how to calibrate the FIS model using well-spread dataset was presented.

Keywords: Data mining · Random forest · Decision tree · Naive Bayes · Flexible pavement · Fuzzy logic

The original version of this chapter was revised: Table 5 has been updated. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-62586-3_14

H. Shehata and S. El-Badawy (Eds.): Sustainable Issues in Infrastructure Engineering, SUCI, pp. 163–192, 2021.
https://doi.org/10.1007/978-3-030-62586-3_11

1 Introduction

The philosophy of pavement usage and preservation states that; it is most cost effective to extend the time when pavements are in good condition with relatively low-cost treatments rather than letting the pavement condition deteriorate until more extensive work is required (Transport Research Board of the National Academies 2007). It is therefore very essential to develop an efficient and sustainable road pavement management system for countries like Nigeria whose major form of transportation is road transport. The practice of pavement management through performance monitoring and prediction for effective planning and timely maintenance and rehabilitation started decades ago. In 1950s, the AASHTO previously known as AASHO first carried out road test program to evaluate pavement conditions in America for efficient maintenance strategy (AASHTO 1993). Thereafter, researches on pavement management system worldwide continued till date (Fwa *et al.* 1994; Fwa and Shanmugam 1998; Cheu *et al.* 2004; Chassiakos 2006; Liu and Sun 2007; Golroo and Tighe 2009; Bianchini and Bandini 2010; Mubaraki 2010; Murana *et al.* 2012a, b; Thube 2011; Mahmood *et al.* 2013; Setyawan *et al.* 2015; Mahmood *et al.* 2015; Savio *et al.* 2016; Ziliute *et al.* 2016; Woo and Yeo 2016; Uglova and Tiraturyan 2016; Xiao and Wu 2016; Premkumar and Vavrik 2016; Pantuso *et al.* 2019; Marcelino *et al.* 2019; Santos *et al.* 2020).

A flexible pavement is an elastic structure made of different unbound layers of soil materials that exhibit nonlinear behaviour under the influence of traffic load, weather condition, material properties, age, etc. A typical cross-section of flexible road pavement structure built according to the Nigeria specifications for road construction is as presented in Fig. 1;

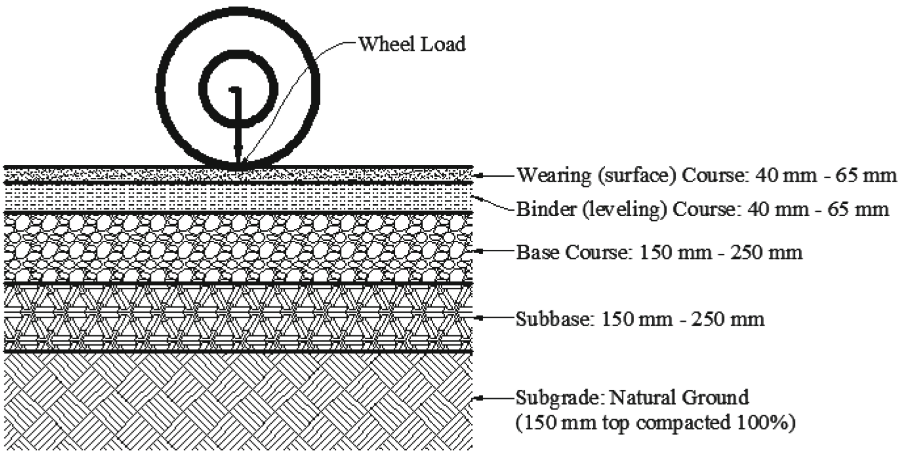


Fig. 1. Schematic of flexible pavement

The choice of flexible pavement for highway construction by most developing countries like Nigeria is due to its relatively cheap cost of construction, maintenance and rehabilitation (Taylor and Philip 2015); though its major disadvantages include relatively short life cycle and rapid deterioration over time (Yin 2007; Ayed 2016; Ziliute *et al.* 2016). The rate of deterioration of flexible pavement depends on several factors ranging from traffic load, climatic condition, material properties, maintenance policy, design of pavement structure and quality of workmanship during construction (Kirbas and Karasahin 2016; Murana 2016; Pantuso *et al.* 2019; Marcelino *et al.* 2019). Development of pavement management system is aimed at achieving good roads that could make cities more liveable and the highways safe for commuting. Threats caused by failed roads include extensive traffic noise and air pollution, massive wear and tear of vehicles, loss of human lives due to fatal auto crashes, high cost of travel due to congestions and fuel consumption, etc. Since huge funds are required for highway construction and maintenance, periodic and timely planning for repairs are essential for efficient management system (Miradi 2009; Abiola *et al.* 2010; Chikezie *et al.* 2011; Owolabi and Abiola 2011; Murana *et al.* 2012a, b; Adefemi and Ibrahim 2015; Mahmood 2015; Ayed 2016; Huang 2017; Gogoi and Dutta 2019).

The assessment of pavement performance usually employs different engineering techniques that require field data generated on-site or obtained from database of pavement management agencies, or from experienced engineers working on pavement distresses (Huang 2004; Miradi 2009; Saltan *et al.* 2011). The performance of road pavement could be classified on a scale of good to worse. The classification is a function of the severity level, quantity and frequency of surface defects caused by deterioration of the pavement materials and its drainage condition (Claros *et al.* 1986; ASTM D6433 2007). The periodic deterioration and consequent failure of road pavement has become a serious concern to pavement management agencies, hence the need for accurate condition classification for optimum performance prediction (Yin 2007; Mahmood 2015; Ayed 2016; Ziliute *et al.* 2016; Adeke *et al.* 2018a, b).

Following the advancement in research activities recently, the use of intelligent algorithms known as Artificial Intelligence (AI) techniques for developing pavement management systems has become a reliable research methodology (Miradi 2009; Russell and Norvig 2010; Mahmood 2015; Marcelino *et al.* 2019; Santos *et al.* 2020). This intelligent approach employs methods of data mining (knowledge discovery) such as Decision Tree, Random Forest, Naïve Bayes classifier, Rough Set Theory, etc. and machine learning (predictions based on known properties) such as; Fuzzy Logic, Genetic Algorithms (GA), Artificial Neural Networks (ANN), etc. (Smadi 2000; Munakata 2008; Witten *et al.* 2020; Fox 2018).

The development of pavement performance prediction models requires input data on pavement surface distress attributes such as; cracks, potholes, rutting, roughness, drainage condition, etc. Other parameters include material strength properties, environmental conditions (climate – rainfall and temperature), traffic load intensity and distribution pattern, and pavement structure are also essential input variables. The process involves generation of huge dataset on pavement behaviour over time (Yu 2005; Setyawan *et al.* 2015; Premkumar and Vavrik 2016; Hamed and Kakarash 2016). Intelligent algorithms are often used for data mining (clustering, classification, visualisation, prediction, etc.) to eliminate challenges of missing or incomplete dataset as relates to data collection and storage for efficient analysis (Claros *et al.* 1986; Miradi 2009; Road Sector Development Team 2014; Luca *et al.* 2016; Fox 2018; Dong *et al.* 2018).

1.1 The Concept of Data Mining

Data mining is a mathematical technique of machine learning which uses black-box models or exploratory techniques to examine dataset that define systems behaviour (Witten and Frank 2005; Fox 2018). It is an approximation technique used for classification, prediction and analysis of imprecise, uncertain or incomplete dataset and knowledge based elements with associated attributes to discover patterns and relationship between variables (Pawlak 2002). Other functions of data mining include to identify partial or total dependencies in a given dataset, eliminates redundant data, gives approach to null values, missing data, dynamic data, etc. The concept is based on the assumption that every object or data point of the universe of discourse is associated with some information (Miradi 2009). It is also believed that objects characterised by the same information are similar in view of their available information. Examples of intelligent algorithms used as classifiers include; Rough Set theory (Miradi 2009), Artificial Neural Network (ANN), Random SubSpace, Support Vector Machines, Pace Regression, Random Forest, Decision Tree (Gopalakrishnan *et al.* 2013; Sharma and Jain 2013), Naïve Bayes theorem (Inkoom *et al.* 2019), etc.

1.1.1 Random Forest

This classifier uses several decision trees in its algorithm to classify systems behaviour during training for accurate classification and predictions. Each tree represents a set of decision alternatives which contest for optimum votes within the dataset (Cigsar and Unal 2019; Li *et al.* 2019). Its works by constructing a group of randomly created decision trees and forecasting the class that is the mode of all classes (classification) or the mean (regression) of the individual trees. The Decision Trees also protect each other from erroneous alternatives since some are weak and unavoidably lead to inaccurate predictions. A collection of weak decision trees can yield better predictions since the prediction accuracy depends on the number of decision trees and predators used in the model. Random Forest has been used by previous researchers for interpretation, visualization, and handling complex non-linearity behaviour between predictors and response. Its limitations include over-fitting, not robust to outliers, and poor in predictive capacity compared to methods like the least square regression. The major advantage of the Random Forest algorithm is that, it does not permit overfitting (Kudjo *et al.* 2020). Li *et al.* 2019 used Random Forest algorithm for the identification of asphalt pavement distresses and condition classification to aid pavement management practice. Findings of the study confirmed its suitability due to relatively high degree of accuracy obtained as compared to the maximum likelihood classification and the support vector maching models. Gong *et al.* (2018) revealed that Random Forest algorithm significantly outperformed the linear regression model with coefficients of determination greater than 0.95 in both training and test sets when used for predicting the IRI of pavement. Pan *et al.* (2018), also used the Random Forest algorithm to detect potholes and caracks in asphlt pavement, and it performance was commendable.

1.1.2 Decision Trees

This is a flow-chart like structure with internal nodes which denote tests on variable sets, each branch represents an outcome of the test and the leaf node represents class distribution. The trees are generated from training data in a top-down, general-to-specific direction. The mechanism of decision tree is such that, whenever all data points have the same output or belong to the same class, there is no further decisions to be made with respect to partition of the data points, and the solution is complete. Contrary to when data points at a node have different outputs or belong to two or more classes, then a test is made at the node that will result into a split. The process is recursively repeated for each of the new intermediate nodes until a completely discriminating tree is obtained. The process is potentially an over-fitting solution, hence eliminates components that are too specific to noise and outliers that may be present in the training data (Miradi 2009; Gopalakrishnan *et al.* 2013).

The recent advanced model used for Decision Tree is kown as the J48 algorithm. Its structure is made of root or internal node which represents array of decision attributes, branches of the tree represent observations of the system performance based on possible outcomes and the leaves or terminal nodes are the target values (regression) or final class (classification) of the dependent variable. Estimations of J48 involve the manipulation of missing values, pruning, derivation of rules, etc. It uses predictive machine-learning model to calculate the resultant value of a new sample based on various attribute values of the available data. A decision tree can handle both binary and non binary classification (Ahishakiye *et al.* 2017; Cigsar and Unal 2019; Obuandike *et al.* 2015). Saravanan and Gayathri (2018) identified limitations of J48 algorithm to include; missing values of zero entries which make the tree wider and complicated, irrelevant entries to a class which lead to wrong classification and the presence of noisy data element which cause over fitting thereby increasing the margin of error.

On the other hand, according to Miradi (2009), advantages of the Decision Tree algorithm include; to handle both numerical and categorical variables, carryout classification/regression in simple form, stores old dataset and efficiently manipulated new dataset, carryout automatic stepwise variables selection and complexity reduction, extremely robust in handling outliers and easy to understand and interpret results. Cubero-Fernandez *et al.* (2017) employed the use of Decision Tree algorithms for pavement crack detection and classification into transverse longitudinal and alligator cracks. The study recommend the decision tree algorithm for further analysis due to its high level of accuracy. Also, Lin *et al.* (2013), Gopalakrishnan *et al.* (2013) and Inkoom *et al.* (2018) used decision tree algorithms for developing pavement maintenance and management system, results obtained by the algorithms yielded accurate estimations.

1.1.3 Naive Bayes Theorem

This algorithm is based on the Bayesian probability theorem which operates on conditional rules (Marianingsih and Utaminingrum, 2018; Inkoom *et al.* 2019). It is an intuitive approach which uses conditional probabilities of the occurrence of individual attributes in a given class label for making prediction of the targeted attribute. It could be used for attribute classifications, the Naïve Bayes algorithm is mostly used for predictive modelling using discrete events of a database without missing data to predict class labels (Alam and Pachauri 2017). It is an example of the Bayesian classification which computes or predictes the probability of an attribute belonging to a given class based on the assumption that the effect of that attribute is independent of the other attributes in the same class, hence is known as class conditional independence model (Tribhuvan *et al.* 2015). According to Alam and Pachauri (2017) the Bayes theorem calculates the probability of a targeted attribute by counting the frequency and combinations of values in the dataset, then estimating the parameter using method of maximum likelihood. The algorithm suites most complex real world problems. Also, its ability to estimate the parameter using small amount of training data is considered the major advantages of the algorithm. But its robustness is easily affected by noisy element in the dataset. According to Jang *et al.* (2015), the Naïve Bayes model is expressed as shown in Eq. 1;

$$P(y|x) = \frac{P(x|y).P(y)}{P(x)} \tag{1}$$

where $P(y|x)$ is the posterior probability of class y (target) given the predictor x (attribute), $P(x)$ is the prior probability of predictor, $P(y)$ is the prior probability of class and $P(x|y)$ is the likelihood which is the probability of predictor given class. Previous studies that used the Naïve Bayes algorithm in pavement management system recorded significant successes (Marianingsih and Utaminingrum 2018; Marianingsih *et al.* 2019; Inkoom *et al.* 2019); its inability to handle missing data and noisy data points yielded results with relatively significant margin of error (Shekharan 1998; Jang *et al.* 2015; Alam and Pachauri 2017; Gong *et al.* 2018).

1.2 The Concept of Fuzzy Logic Theory

Fuzzy Logic is an AI technique used for developing knowledge-based models which use human intuitive reasoning characterised by subjectivity and imprecision. Unlike the classic logic defined by binary values of 0 and 1, it is a multivalued logic model capable of measuring intermediate values between conventional evaluations of *on or off, true or false, yes or no,* etc. which could be formulated mathematically and processed using computers (Munakata 2008). Fuzzy Logic describes a system at varying degrees using adjectives like '*low', 'medium' and 'high'*. The theory itself is not imprecise in nature, but a mathematical theory which deals with subjectivity and uncertainties. It is capable of handling data inaccuracies, uncertainties and non-linearity. It uses intuitive and subjective expert knowledge or experience for Fuzzification process where there are no sufficient numeric data. The Fuzzy rules try to execute Fuzzy reasoning used for evaluating true degree of goal propositions (Hamed and Kakarash 2016). In this theory, the human vague thought and perception on interpretation of systems behaviour is expressed using appropriate mathematical models. Although probability theorems such as stochastic or Monte Carlo theorems have vast applications in analysing varying systems, they are subsets of the Fuzzy theorem (Bai *et al.* 2006; Munakata 2008). Fuzzy logic provides mathematical strength for emulating certain perceptual and linguistic attributes associated with human reasoning which are a function of experience over time in a systematic manner. Some quantities do not have sharp boundaries, hence best described using labels as; *'worse'*, *'many', 'tall', 'young', 'small'*, etc. which represent the fact known as Fuzzy concepts. They are usually true to some degree that cannot be easily quantified or are relative in nature as well as false on the other hand. The major limitation of Fuzzy logic is caused by uncertainties associated with total lack of information about the system behaviour. In contrast, the traditional binary logic usually has discrete set of options used for describing crisp events without intermediate classes.

1.2.1 Fuzzy Logic Modelling

A Fuzzy inference-based system is an intelligent method used for classification and prediction problems. The system interprets values in the input vector based on user-defined rules and assigns values to the output vector. The advantages of this approach is its knowledge representation in form of the *IF-THEN* logic rules, the mechanism of reasoning in human understandable terms, the capacity of taking linguistic information from human experts and combining it with numerical information, and the ability of approximating complicated nonlinear functions with simpler models. The Fuzzy rules are usually generated either from an expert knowledge or numerical dataset which makes the theory suitable for analysing several problems (Munakata 2008; Mahmood *et al.* 2013). A Fuzzy set is an extension of a crisp set since crisp set allows only full membership or no membership at all, a Fuzzy set allows for partial membership in the system (Bai *et al.* 2006). The Fuzzy Inference System (FIS) depends on its Membership Function (MF) which defines its correct value between 0 and 1. The degree to which any Fuzzy statement is true falls within the range of 0 and 1. Fuzzy logic provides an inference mechanism under cognitive uncertainty and mathematical computations using words. According to Chen and Pham (2001), for a conventional set, the MF of an element x in subset A, $\mu_A(x)$ is expressed as Eq. 2;

$$\mu_A(x) = \begin{cases} 1 & \textit{if and only if } x \in A \\ 0 & \textit{otherwise} \end{cases} \tag{2}$$

The degree of membership of each entity is a numerical measure from 0 to 1, where 0 denotes nonexistence of the element in the subset or set, 1 denotes full membership of the set and any value between 0 and 1 defines the partial degree of membership of the element in the set (Chen and Pham 2001; Bai *et al.* 2006). The basic mathematical relationship between subsets A and B include; union, intersect and complements, which are mathematically expressions as Eqs. 3–5;

$$A \cup B = \mu_A(x) \cup \mu_B(x) = \max(\mu_A(x), \mu_B(x)) \tag{3}$$

$$A \cap B = \mu_A(x) \cap \mu_B(x) = \min(\mu_A(x), \mu_B(x)) \tag{4}$$

$$A^c = 1 - \mu_A(x) \tag{5}$$

The Fuzzy MFs are commonly expressed graphically as; triangular, trapezoidal, Gaussian, bell-shaped, sigmoid, etc. curves. According to Mahmood *et al.* (2015), Santos *et al.* (2020) and Gogoi and Dutta (2019), the triangular MFs is suitable for pavement performance analysis. Chen and Pham (2001) and Bai *et al.* (2006) defined mathematical details of the triangular MF as shown in Fig. 2;

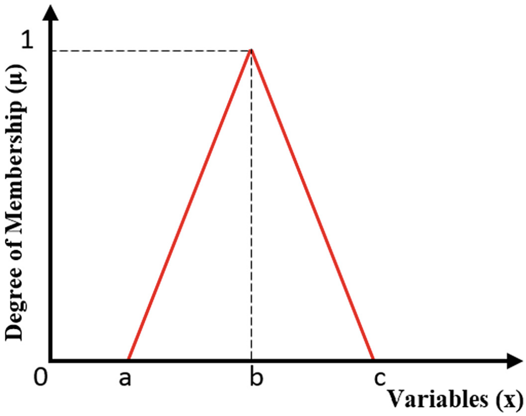


Fig. 2. Triangular Fuzzy MF

Figure 2 presents a triangular MF of Fuzzy set A, where only elements at point b assume full membership of 1. Models that define the degree of membership of elements of Fuzzy set A at various stages are as shown in Eqs. 6a and 6b;

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \\ 0 & x \geq c \end{cases} \tag{6a}$$

$$\mu_A(x; a, b, c) = max\left(min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \tag{6b}$$

The Mamdani framework is commonly used for executing FIS based on its set of MFs for both input and output variables in form of crisp values (Kaur and Tekkedil 2000; MATLAB 2015).

1.3 Analysis of Systems Behaviour Using Fuzzy Logic

Following guidelines stated by Bai *et al.* (2006), the three basic steps required for solving real life problems using the Fuzzy logic approach include Fuzzification, formation of Fuzzy logic rules and Defuzzification. Fuzzification of quantities is the process of translating the available numerical input variables into linguistic variable such as low, medium and high based on defined thresholds of membership or classifications of Fuzzy sets used for the development of Fuzzy logic rules. The Fuzzy logic rules are the decision rules based on experience and logic applied on the Fuzzified linguistic variables. Defuzzification of quantities is the end product which reverses the Fuzzification process to produce actual and meaningful output in linguistic form. The process involves reconverting the linguistic terms into crisp values according to their degree of membership using Fuzzy logic model such as the Mamdani model. According to Chen and Pham (2001) and Bai *et al.* (2006) some numerical methods of Defuzzification analysis include; the centre of gravity method, mean of maximum method, height method, etc. The centre of gravity method is commonly used, it defines the defuzzified value $\tilde{x}$ of a Fuzzy set as the Fuzzy centroid measured along the horizontal axis expressed as the centre of area under the MF curve using Eqs. 7a and 7b for discrete and continuous systems analysis respectively;

$$\tilde{x} = \frac{\sum_{i=1}^{n} x_i.\mu_i(x_i)}{\sum_{i=1}^{n} \mu_i(x_i)} \tag{7a}$$

$$\tilde{x} = \frac{\int_1^n x_i.\mu_i(x_i)dx}{\int_1^n \mu_i(x_i)dx} \tag{7b}$$

where, $\mu_i(x_i)$ is the MFs of area of the i th sampled element in the Fuzzy set and x_i is its centroid distance from the origin of the horizontal axis or crisp value.

1.4 Pavement Management System Using Fuzzy Logic

The subjective measurement of pavement performance attributes for other prediction models is seen as major limitation since its creates significant margin of error in estimation, this disadvantage is better handled using the Fuzzy logic theory due to its ability to Fuzzify variables (Liu and Sun 2007; Karagahin and Terzi 2014; Mahmood *et al.* 2015). Some existing pavement performance prediction models are based on pavement surface condition classification such as; Condition Rating Survey (CRS) and Pavement Condition Index (PCI) methodologies among others (Dabous *et al.* 2019a, b). The use of such models require the initial subjective assessment of pavement distress conditions as input variable (Setyawan *et al.* 2015; Premkumar and Vavrik 2016; Hamed and Kakarash 2016). In Nigeria, dataset on road pavement surface condition evaluation are not readily available. This is basically due to the cost of operation, lack of machineries and operational skills, government policies, etc. Though the challenge is not peculiar to Nigeria (Dabous *et al.* 2019a, b). Investigating the impact of various damaging factors on road pavement require large data on all the influencing factors which is usually expensive in terms of cost and time (Abiola 2012; Surendrakuma *et al.* 2013; Mane *et al.* 2016). Salpisoth (2014) identified major challenges in road pavement management system to include; the ability to carry out routine monitoring and evaluation of pavement condition regularly with limited budget and expertise; and how to predict pavement performance using appropriate models based on incomplete data for optimum maintenance and repair decisions. The use of Fuzzy logic method in pavement management system in terms of condition classification, performance predictions and decision prioritisation has gained wide commendation in recent times. This assertion can be proven using previous studies as listed in Table 1;

This study therefore aims at developing a framework for data mining and performance prediction of flexible road pavement using Fuzzy Logic theory. Objectives of the

Table 1. The use of fuzzy logic theory in pavement management system

Author(s) and date	Methodology	Findings
Shoukry *et al.* (1997)	Pavement condition classification	Recommended
Fwa and Shanmugam (1998)	Pavement condition classification	Recommended
Bandara and Gunaratne (2001)	Pavement condition classification	Recommended
Arliansyah *et al.* (2003)	Pavement condition classification	Recommended
Golroo and Tighe (2009)	Pavement condition classification	Recommended
Koduru *et al.* (2010)	Pavement condition classification	Recommended
Mahmood *et al.* (2013)	Pavement condition classification	Recommended
Mahmood (2015)	Pavement condition classification	Recommended
Mahmood *et al.* (2015)	Pavement condition classification	Recommended
Kaur and Tekkedil (2000)	Pavement performance prediction	Recommended
Liu and Sun (2007)	Pavement performance prediction	Recommended
Miradi (2009)	Pavement performance prediction	Recommended
Bianchini and Bandini (2010)	Pavement performance prediction	Recommended
Wang and Li (2011)	Pavement performance prediction	Recommended
Karagahin and Terzi (2014)	Pavement performance prediction	Recommended
Aggarwal and Kumar (2015)	Pavement performance prediction	Recommended
Jeong *et al.* (2017)	Pavement performance prediction	Recommended
Cheu *et al.* (2004)	Maintenance decision prioritisation	Recommended
Chassiakos (2006)	Maintenance decision prioritisation	Recommended
Chandran et al. (2007)	Maintenance decision prioritisation	Recommended
Chen and Flintsch (2008)	Maintenance decision prioritisation	Recommended
Moazami *et al.* (2011)	Maintenance decision prioritisation	Recommended
Gogoi and Dutta (2019)	Maintenance decision prioritisation	Recommended
Nawir and Prihartanto (2019)	Maintenance decision prioritisation	Recommended
Santos *et al.* (2020)	Maintenance decision prioritisation	Recommended

study include; to carryout data mining for optimum initial road pavement surface condition classification based on attributes that affect road pavement performance in Nigeria using intelligent algorithms of Waikato Environment for Knowledge Analysis (WEKA) software, to develop and simulate a Fuzzy Logic Inference System for performance prediction of flexible road pavement in Nigeria implemented using MATLAB software and to develop a decision framework for performance prediction of flexible road pavement in Nigeria.

2 Methodology

2.1 Description of Study Area

Nigeria is among the most populous countries on the Africa continent with an estimated human population of 206 million persons. Its land mass is estimated at 923,768.00 square kilometres of surface area. Road transport is her major form of land transportation system. The distribution of Federal Highways in Nigeria by length sums to 34,340.95 km (Federal Government of Nigeria 2012). Details of some selected links of Federal Highway in Nigeria is as shown in Table 2;

Table 2. Selected federal highway links in Nigeria (Claros *et al.* 1986)

S/N	State	Link ID	Length (km)	No. of lanes	Age (years)	Surface finishing
1.	Anambra	89	43.7	2	6	Asphalt concrete
2.	Bauchi	285	51.3	2	8	Asphalt concrete
3.	Benue	112	53.6	2	6	Asphalt concrete
4.	Borno	716	50.7	2	9	Surfaced dressed
5.	Kaduna	130	46.8	2	5	Asphalt concrete
6.	Kwara	5	22.8	2	9	Asphalt concrete
7.	Ogun	17	18.3	2	12	Asphalt concrete
8.	Ogun	332	30.9	2	10	Asphalt concrete
9.	Plateau	136	10.0	2	9	Asphalt concrete
10.	Plateau	138	23.2	2	5	Asphalt concrete
11.	Imo	370	45.6	2	2	Asphalt concrete
12.	Rivers	144	18.0	2	8	Asphalt concrete
13.	Sokoto	255	57.3	2	8	Asphalt concrete
14.	Oyo	22	23.4	2	10	Asphalt concrete

2.2 Data Collection and Description

The study considered secondary dataset obtained from the database of pavement condition evaluation unit of the Federal Ministry of Power, Works and Housing Nigeria as reported in Claros *et al.* (1986) for data mining and model development. The dataset captured condition attributes of 102 links of flexible pavement of Federal Highways across Nigeria. The relevant attributes used for the study included;

1. **Average daily truck load:** this was reported as the average number of daily truck load known as heavy goods or commercial vehicles in each direction that weighed at least 1500 kg as estimated from the Average Daily Traffic (ADT) volume travelling on the Federal Highway facility during the investigation period.

2. **Strength property of subgrade soil:** the Resilient Modulus (M_R) of subgrade soil determined based on procedures stated in AASHO T-274-82 as reported by Claros *et al.* (1986) was used as the subsoil strength representative variable.
3. **Weather condition:** weather parameters of interest included the Average Annual Air Temperature and Rainfall within the classified Zones.
4. **Age of pavement:** this was the period in years measured from when the Highway facility was first opened to traffic or rehabilitated till the date of condition evaluation survey.
5. **Initial pavement condition (IPC):** this classification used the Condition Survey Rating Scores (CSRS) which measures from 0–100 for Very poor to Excellent based on the average rut depth (cm), percentage area of alligator or fatigue cracking and the drainage condition per road segment. The classifications of CSRS scale is as presented in Table 3;

Table 3. Classification of pavement surface condition

Classification of pavement condition	Limits of CSRS
Excellent	>81
Good	66–81
Fair	46–65
Poor	25–45
Very poor	<25

2.3 Data Mining for IPC Classification

Data mining algorithms were used to determine the accuracy of initial surface condition classification of pavement using WEKA software since the measure of some attributes was subjective and characterised by missing and incomplete dataset. The Random Forest, Decision Tree and Naïve Bayes classifiers were used for optimum training (Witten *et al.* 2015). Aside the dataset used for IPC classification which was vague in nature, other performance attributes such as; traffic load, temperature and rainfall, etc. had no sufficient dataset that necessitated data mining process, hence were manually classified based on guidelines specified in manuals and experience (Claros *et al.* 1986; Road Sector Development Team 2014). Table 4 presents the summary of data attributes and types used (Claros *et al.* 1986);

Table 4. Data attributes and type

S/N	Attributes	Data type
1.	Average rut depth (cm)	Numerical
2.	Fatigue cracking (%)	Numerical
3.	Drainage condition	Nominal
4.	CSRS (%)	Numerical
5.	Pavement condition	Nominal

2.4 Development of FIS

The Fuzzy logic tool in MATLAB programming software was used to develop Mamdani FIS. The Fuzzification process involved identifying essential attributes into Fuzzy sets as input and output for performance prediction of road pavement. The creation of Fuzzy MF involved identifying suitable function (curve) based on classified attributes in the datasets, then calibrating it using representative parameters of the attribute (Mahmood 2015; Nawir and Prihartanto 2019; Santo *et al.* 2020). Table 5 presents the proposed classifications of attributes and ranges to form Fuzzy sets used for the development of Fuzzy MF;

Table 5. Development of fuzzy sets for attributes

S/N	Parameters	Classifications	Limits	Ranges of FMF
1.	Initial pavement condition (IPC)	Very poor	<25	[0 15 25]
		Poor	25–45	[20 35 50]
		Fair	46–65	[40 50 65]
		Good	66–81	[60 70 87]
		Excellent	>81	[80 90 100]
2.	Age of pavement (years)	New	<2	[0 1 3]
		Recent	2–5	[2 5 7]
		Old	6–12	[6 9 13]
		Very old	>12	[12 16 20]
3.	Truck load (veh/day)	Low	<200	[0 125 280]
		Medium	200–1000	[195 600 1000]
		High	1001–2000	[850 1400 2000]
		Very high	>2000	[1700 2000 2500]
4.	M_R of subgrade soil (kg/m^2)	Low	$<0.7 \times 10^7$	$[0\ 0.4\ 0.75] \times 10^7$
		Medium	$[0.7–1.3] \times 10^7$	$[0.6\ 1.0\ 1.425] \times 10^7$
		High	$[1.3–2.1] \times 10^7$	$[1.2\ 1.8\ 2.4] \times 10^7$
		Very high	$>2.1 \times 10^7$	$[2.0\ 2.48\ 3.0] \times 10^7$
5.	Temperature ($^\circ$C)	Low	<20	[0 10 20]
		Medium	20–35	[15 25 35]
		High	36–45	[30 35 42]
		Very high	>45	[40 45 50]
6.	Rainfall (mm)	Low	<600	[0 250 650]
		Medium	600–1200	[500 820 1200]
		High	1201–1800	[1100 1500 1900]
		Very high	>1800	[1700 2000 2500]
7.	Future pavement condition (FPC)	Very poor	<25	[0 15 25]
		Poor	25–45	[20 35 50]
		Fair	46–65	[40 50 65]
		Good	66–81	[60 70 87]
		Excellent	>81	[80 90 100]

Using Fuzzy sets presented in Table 5, the Fuzzification and creation of Fuzzy logic rules for implementation using MATLAB software based on the data source was as shown in Table 6;

Table 6 revealed that highway links whose properties satisfied a given trend or had similar pattern of logical behaviour were identified and grouped into blocks with the observed crisp values that defined the Future Pavement Condition (FPC). At the instances of missing data and pattern mismatch, simple extrapolation techniques (guided by expe-

Table 6. Fuzzification and fuzzy logic rules

Rule No.	Link Code	IPC	Age (Years)	Truck Load (Veh/Day)	MR of Subgrade (kg/m2)	Temp. (°C)	Rainfall (mm)	FPC	Crisp Value
1.	370	Excellent	New	Low	High	Medium	V.High	Good	76
2.	615	Excellent	Recent	Low	High	Medium	High	Excellent	85
3.	61	Excellent	Recent	Medium	Medium	Medium	High	Excellent	93
4.	565	Excellent	Recent	Medium	Medium	Medium	V.High	Excellent	87
5.	130	Excellent	Recent	Medium	High	Medium	High	Excellent	92
6.	298	Excellent	Recent	Medium	High	Medium	V.High	Good	72
7.	736	Excellent	Recent	Medium	V.High	Medium	High	Good	78
8.	405	Excellent	Old	Low	Low	Medium	High	Good	76
9.	371	Excellent	Old	Low	Medium	Medium	V.High	Good	76
10.	716	Excellent	Old	Low	Medium	Medium	Medium	Fair	60
11.	27	Excellent	Old	Low	Medium	Medium	Medium	Excellent	92
12.	472	Excellent	Old	Low	High	Medium	Medium	Excellent	86
13.	287	Excellent	Old	Low	High	Medium	Medium	Good	77
14.	551	Excellent	Old	Low	High	Medium	High	Excellent	90
15.	630	Excellent	Old	Low	High	Medium	V.High	Good	78
16.	289	Excellent	Old	Medium	Medium	Medium	Medium	Excellent	83
17.	219	Excellent	Old	Medium	Medium	Medium	High	Excellent	94
18.	136	Excellent	Old	Medium	Medium	Medium	High	Good	75
19.	252	Excellent	Old	Medium	High	Medium	Medium	Excellent	86
20.	290	Excellent	Old	Medium	High	Medium	Medium	Good	81
21.	17	Excellent	Old	Medium	High	Medium	High	Good	69
22.	275	Excellent	Old	Medium	High	Medium	High	Excellent	90
23.	176	Excellent	Old	High	Medium	Medium	High	Excellent	92
24.	50	Excellent	Old	Medium	High	Medium	V.High	Excellent	92
25.	186	Excellent	Old	Medium	High	Medium	V.High	Good	76
26.	211	Excellent	Old	High	High	Medium	High	Excellent	84
27.	144	Excellent	Old	High	High	Medium	V.High	Good	76
28.	187	Excellent	Old	High	V.High	Medium	V.High	Good	76
29.	173	Excellent	Old	V.High	High	Medium	High	Good	78
30.	210	Excellent	Old	V.High	High	Medium	V.High	Good	79
31.	121	Excellent	Old	High	High	Medium	V.High	Excellent	85
32.	89	Excellent	Old	High	High	Medium	High	Excellent	88
33.	91	Excellent	Old	High	Medium	Medium	High	Excellent	86
34.	123	Excellent	V.Old	Low	High	Medium	V.High	Fair	59

rience) were used to fill some targeted classes that were necessary to aid logical decisions (Chen and Flintsch 2008). The process filtered identical rules and eliminated repeated entries and entries with missing data leading to the developing of 34 unique Fuzzy logic rules (out of 102 links considered) for implementation using MATLAB software. A plot of triangular Fuzzy logic membership functions MF for inputs and output variables used is as shown in Fig. 3;

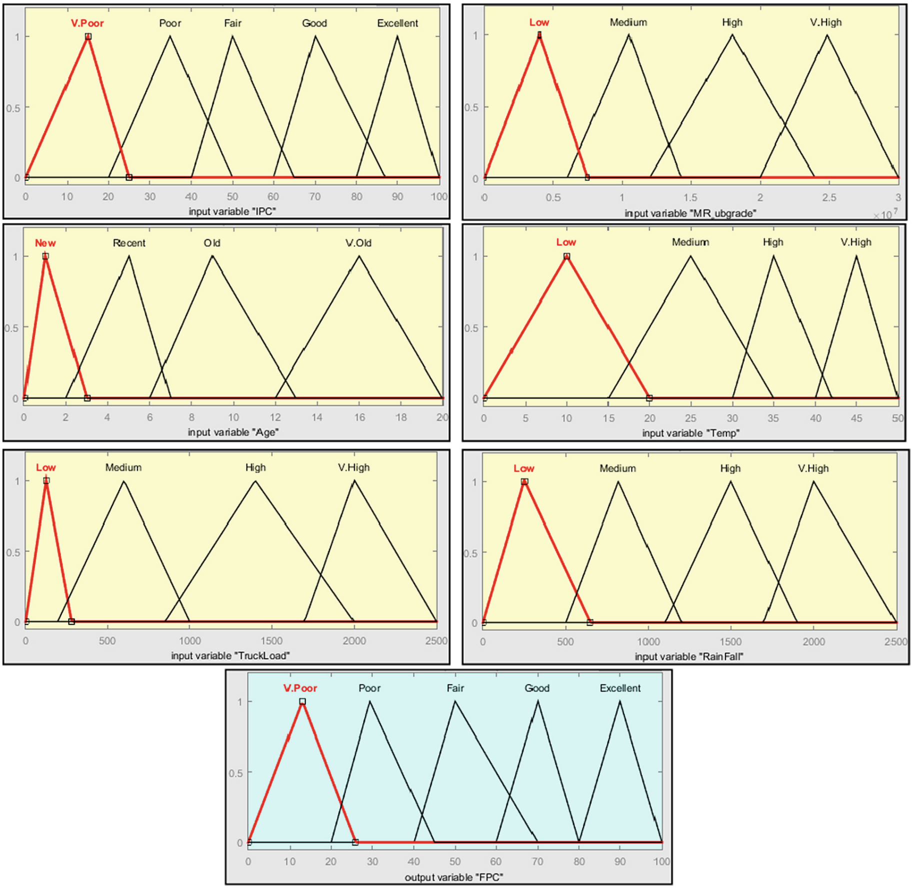


Fig. 3. MF plots for inputs and output variables

The *IF . . . THEN . . .* logic rule was further used to build conditional logic that defined pavement behaviour over time, then the model was simulated and defuzzified to yield FPC as the output. An illustration of this methodology is as presented in Fig. 4;

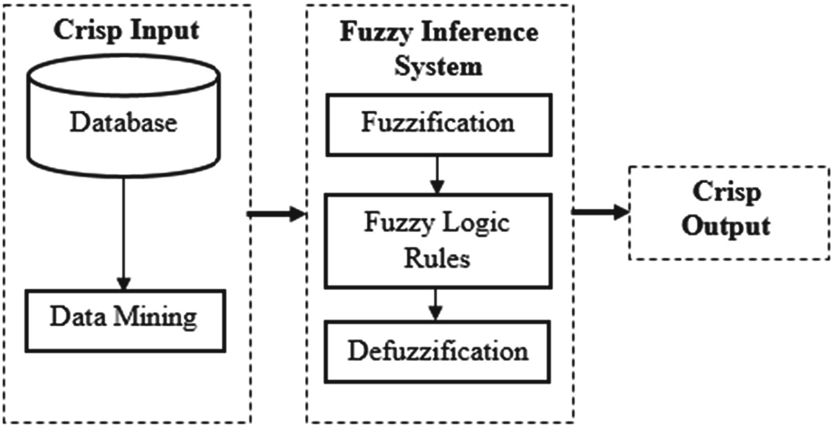


Fig. 4. Schematic of methodology

3 Results and Discussion

3.1 Data Mining for IPC Classifications

Detailed results of initial pavement surface condition classification using intelligent algorithms - Random Forest, Decision Tree and Naïve Bayes classifier for the sampled links were as presented in Fig. 5;

Figure 5 present confusion matrices of classifications where entries with perfect diagonal matrices indicated accurate classification, while those forming upper and lower triangular matrices showed incorrect classifications. Instances considered were correctly or incorrectly classified into Excellent, Good, Fair, Poor and Very Poor conditions by the Random Forest, Decision Tree and Naïve Bayes algorithms. The perfect diagonal matrices for Fig. 5 (a) to (n) classifications using Random Forest model showed its weakness and insensitivity in handling missing and noisy data elements, while the Decision Tree model failed to attain perfect diagonal matrices for classifications of Fig. 5 (a), (h), (m) and (n) to show its relatively low sensitivity to discrepancies in the dataset. On the other hand, the Naïve Bayes model had perfect diagonal matrix at Fig. 5 (j) classification only. This implied that all instances were fittingly classified into the defined classifications of the pavement surface condition based on the attributes used. The Random Forest and Decision Tree classifiers present relatively perfect diagonal matrices while the Naïve Bayes classifier gave substantially dispersed matrices. The dispersed trend of confusion matrices is an indication of incorrect classifications, and is attributed to high sensitivity of the Naïve Bayes classifier to the presence of missing data and noisy data elements in the dataset, which affected the accuracy of its classifications (Inkoom *et al.* 2019; Kudjo

Random Forest | **Decision Tree** | **Naive Bayes**

Random Forest:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 12  0  0 |  a = Good
  0 93  0 |  b = Excellent
  0  0  1 |  c = Fair
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 12  0  0 |  a = Good
  0 93  0 |  b = Excellent
  1  0  0 |  c = Fair
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 10  2  0 |  a = Good
  6 87  0 |  b = Excellent
  0  0  1 |  c = Fair
```

(a) Link 89 in Anambra State

Random Forest:
```
=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 51  0  0  0  0 |  a = Fair
  0 35  0  0  0 |  b = Poor
  0  0 35  0  0 |  c = Excellent
  0  0  0 42  0 |  d = Good
  0  0  0  0 10 |  e = Very Poor
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 51  0  0  0  0 |  a = Fair
  0 35  0  0  0 |  b = Poor
  0  0 35  0  0 |  c = Excellent
  0  0  0 42  0 |  d = Good
  0  0  0  0 10 |  e = Very Poor
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 43  2  0  6  0 |  a = Fair
  2 27  0  0  6 |  b = Poor
  0  0 34  1  0 |  c = Excellent
  0  0  1 41  0 |  d = Good
  0  1  0  0  9 |  e = Very Poor
```

(b) Link 716 in Borno State

Random Forest:
```
=== Confusion Matrix ===

  a  b   <-- classified as
 69  0 |  a = Excellent
  0  5 |  b = Good
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b   <-- classified as
 69  0 |  a = Excellent
  0  5 |  b = Good
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b   <-- classified as
 68  1 |  a = Excellent
  0  5 |  b = Good
```

(c) Link 5 in Kwara State

Random Forest:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 21  0  0 |  a = Excellent
  0 16  0 |  b = Good
  0  0 10 |  c = Fair
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 21  0  0 |  a = Excellent
  0 16  0 |  b = Good
  0  0 10 |  c = Fair
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 21  0  0 |  a = Excellent
  0 16  0 |  b = Good
  0  2  8 |  c = Fair
```

(d) Link 370 in Imo State

Random Forest:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  6  0 |  a = Good
  0 99 |  b = Excellent
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  6  0 |  a = Good
  0 99 |  b = Excellent
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  6  0 |  a = Good
  1 98 |  b = Excellent
```

(e) Link 332 in Ogun State

Random Forest:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  7  0 |  a = Good
  0 70 |  b = Excellent
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  7  0 |  a = Good
  0 70 |  b = Excellent
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b   <-- classified as
  7  0 |  a = Good
  2 68 |  b = Excellent
```

(f) Link 285 in Bauchi State

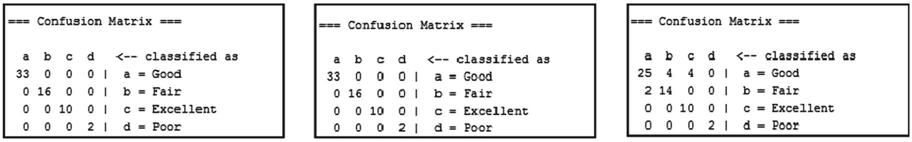
Random Forest:
```
=== Confusion Matrix ===

   a   b   <-- classified as
 186   0 |  a = Excellent
   0  14 |  b = Good
```
Decision Tree:
```
=== Confusion Matrix ===

   a   b   <-- classified as
 186   0 |  a = Excellent
   0  14 |  b = Good
```
Naive Bayes:
```
=== Confusion Matrix ===

   a   b   <-- classified as
 184   2 |  a = Excellent
   1  13 |  b = Good
```

(g) Link 255 in Sokoto State

Random Forest:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 29  0  0 |  a = Good
  0 26  0 |  b = Excellent
  0  0  1 |  c = Fair
```
Decision Tree:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 29  0  0 |  a = Good
  0 26  0 |  b = Excellent
  1  0  0 |  c = Fair
```
Naive Bayes:
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 27  2  0 |  a = Good
  1 25  0 |  b = Excellent
  0  0  1 |  c = Fair
```

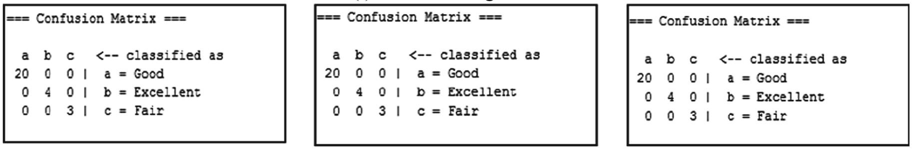
(h) Link 22 in Oyo State

Fig. 5. Surface condition classification of pavement using confusion matrices



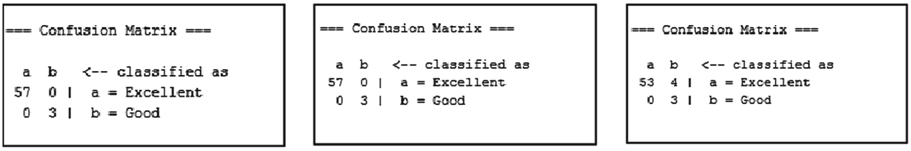
```
=== Confusion Matrix ===

  a  b  c  d   <-- classified as
 33  0  0  0 |  a = Good
  0 16  0  0 |  b = Fair
  0  0 10  0 |  c = Excellent
  0  0  0  2 |  d = Poor

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
 33  0  0  0 |  a = Good
  0 16  0  0 |  b = Fair
  0  0 10  0 |  c = Excellent
  0  0  0  2 |  d = Poor

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
 25  4  4  0 |  a = Good
  2 14  0  0 |  b = Fair
  0  0 10  0 |  c = Excellent
  0  0  0  2 |  d = Poor
```

(i) Link 17 in Ogun State



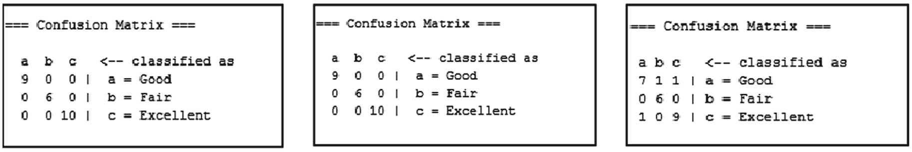
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 20  0  0 |  a = Good
  0  4  0 |  b = Excellent
  0  0  3 |  c = Fair

=== Confusion Matrix ===

  a  b  c   <-- classified as
 20  0  0 |  a = Good
  0  4  0 |  b = Excellent
  0  0  3 |  c = Fair

=== Confusion Matrix ===

  a  b  c   <-- classified as
 20  0  0 |  a = Good
  0  4  0 |  b = Excellent
  0  0  3 |  c = Fair
```

(j) Link 144 in Rivers State



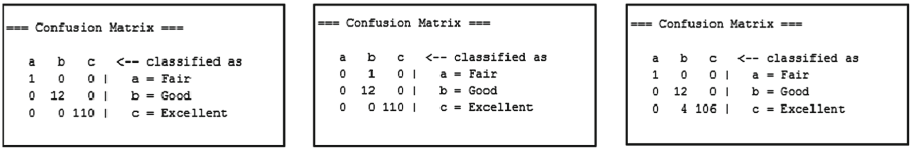
```
=== Confusion Matrix ===

  a  b   <-- classified as
 57  0 |  a = Excellent
  0  3 |  b = Good

=== Confusion Matrix ===

  a  b   <-- classified as
 57  0 |  a = Excellent
  0  3 |  b = Good

=== Confusion Matrix ===

  a  b   <-- classified as
 53  4 |  a = Excellent
  0  3 |  b = Good
```

(k) Link 138 in Plateau State



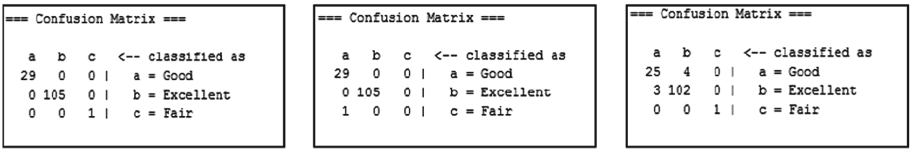
```
=== Confusion Matrix ===

  a  b  c   <-- classified as
  9  0  0 |  a = Good
  0  6  0 |  b = Fair
  0  0 10 |  c = Excellent

=== Confusion Matrix ===

  a  b  c   <-- classified as
  9  0  0 |  a = Good
  0  6  0 |  b = Fair
  0  0 10 |  c = Excellent

=== Confusion Matrix ===

 a b c   <-- classified as
 7 1 1 | a = Good
 0 6 0 | b = Fair
 1 0 9 | c = Excellent
```

(l) Link 136 in Plateau State



```
=== Confusion Matrix ===

  a   b   c   <-- classified as
  1   0   0 |  a = Fair
  0  12   0 |  b = Good
  0   0 110 |  c = Excellent

=== Confusion Matrix ===

  a   b   c   <-- classified as
  0   1   0 |  a = Fair
  0  12   0 |  b = Good
  0   0 110 |  c = Excellent

=== Confusion Matrix ===

  a   b   c   <-- classified as
  1   0   0 |  a = Fair
  0  12   0 |  b = Good
  0   4 106 |  c = Excellent
```

(m) Link 130 in Kaduna State



```
=== Confusion Matrix ===

  a   b   c   <-- classified as
 29   0   0 |  a = Good
  0 105   0 |  b = Excellent
  0   0   1 |  c = Fair

=== Confusion Matrix ===

  a   b   c   <-- classified as
 29   0   0 |  a = Good
  0 105   0 |  b = Excellent
  1   0   0 |  c = Fair

=== Confusion Matrix ===

  a   b   c   <-- classified as
 25   4   0 |  a = Good
  3 102   0 |  b = Excellent
  0   0   1 |  c = Fair
```

(n) Link 112 in Benue State

Fig. 5. (*continued*)

et al. 2020). In order words, failure to achieve perfect classification was attributed to the number of incorrectly classified instances due to significant disparities observed in the dataset which were possibly due to missing or noisy data entries (Witten *et al.* 2016; Marianingsih and Utaminingrum 2018). Models performance through percentage accuracy of classifiers was estimated using the correct and incorrect classifications as shown in Fig. 6;

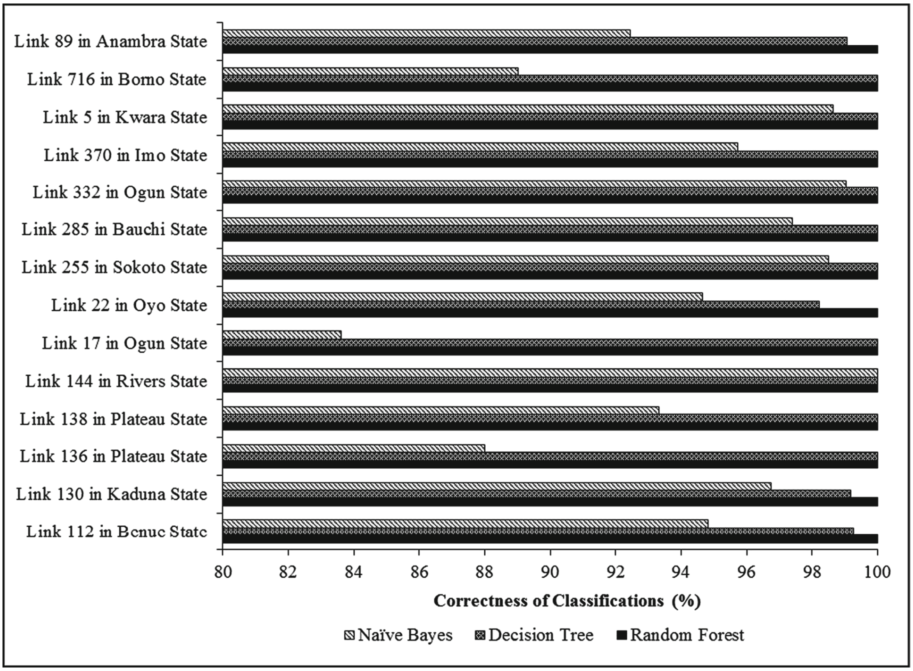


Fig. 6. Accuracy of classifications

Figure 6 revealed that the Random Forest and Decision Tree models made classifications with relatively high accuracy, while the Naïve Bayes classifier made relatively significant number of incorrect classifications. This was attributed to their insensitivity to the challenge of missing data compared to the Naïve Bayes algorithms which recorded percentage errors due to incorrect classification caused by noise data points except for Link 144 in River State (Shekharan 1998; Jang *et al.* 2015; Alam and Pachauri 2017; Gong *et al.* 2018; Inkoom *et al.* 2019). Also, the coefficient of correlations analysis of classifications presented by Kappa statistics of WEKA software is as shown in Fig. 7;

Figure 7 revealed that classifications by the Random Forest model assumed accurate performance at 100% while those by the Decision tree model had some margin of error with that of Naïve Bayes theorem having relatively low accuracy, except for link 144 in River state. The classification errors are further quantified using the Root Relative Squared Error (%) as shown in Fig. 8;

Results presented in Fig. 8 revealed that assertions established in Fig. 7 were correct due to the inversely proportional relationship between the coefficient of correction and the estimated root relative squared errors (%) of classification outputs of the models for links.

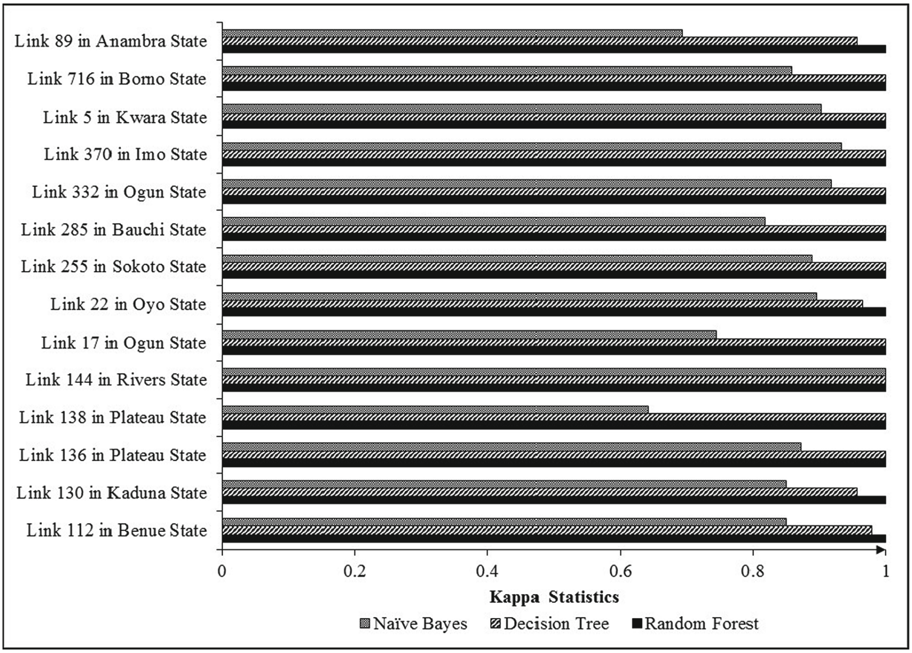


Fig. 7. Coefficient of correlation

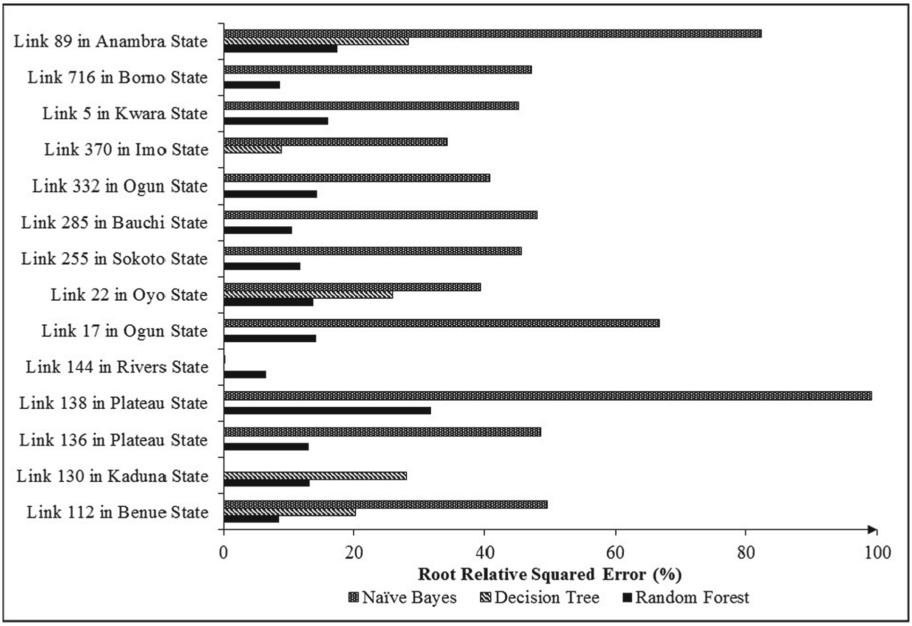


Fig. 8. Error of classification

3.2 Calibration and Validation of the FIS Model

Parameters used for the development of the FIS were adjusted to conform to realities of flexible road pavement behaviour and extreme cases of pavement conditions in Nigeria based on current conditions (Jeong *et al.* 2017; Cheu *et al.* 2004), it fitted the model inputs and outputs within limits of variables in the dataset that described the system behaviour (Chen and Flintsch 2008; Mahmood *et al.* 2013). According to Mahmood *et al.* (2013), the process requires a well spread dataset to be able to capture all possible events in the system behaviour. The process was guided by the original dataset used for model development, logical experience and engineering judgement – in other words, fitting experience to the observed logics or patterns (Chen and Flintsch 2008; Karagahin and Terzi 2014; Aggarwal and Kumar 2015; Mahmood 2015).

3.2.1 Goodness of Fit Test

Figure 9 presents a curve fitting plot of the observed and predicted results of the built FIS;

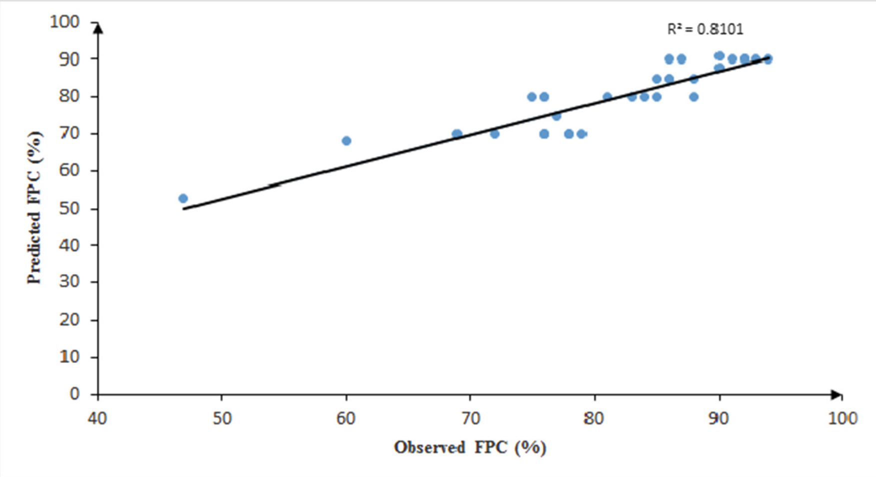


Fig. 9. Curve fitting plot

Figure 9 revealed the proportion of variance accounted for between the observed and predicted FPC values through the coefficient of determination, $R^2 = 0.8101$ with coefficient of alienation $(1 - R^2)$ at 0.1899. This defined the extent of similarity between the observed and predicted values of FPC at 81% level of accuracy. In other words, the strength of the best-fit line known as the correlation coefficient was estimated at 90% consistency (Heiman 2011).

The ‘independent’ dataset required for the validation of the built model was not readily available. But the relatively high degree of accuracy shown from the goodness of fit test explained the high level of certainty in using the model. Therefore, being a linguistic and experienced based model, the process adopted permutation principles to derive 5120 Fuzzy Logic Rules which satisfied possible pavement conditions in Nigeria for performance prediction (See extract in Appendix).

4 Conclusion

This study used the Random Forest, Decision Tree and Naive Bayes algorithms of data mining to examine the inferred dataset on pavement performance attributes and surface condition classification in Nigeria based on reports of pavement surface condition evaluation survey carried out in 1986. The data mining techniques were used to investigate hidden relationship between pavement performance attributes and to authenticate the accuracy of subjective measurements that were used for pavement surface condition classification. The Random Forest and Decision Tree algorithms reported relatively perfect classifications of road pavement sections into; Excellent, Good, Fair, Poor and Very poor. On the other hand, the Naïve Bayes algorithm yielded inaccurate classifications with some margin of errors which were attributed to missing and noisy entries in the dataset.

Considering outputs of the Naïve Bayes classifier, this necessitated the use of Fuzzy logic theory for performance prediction of the road pavement due to its capability to handle the imprecise dataset. It was used to develop Fuzzy Inference System (FIS) for performance prediction of flexible road pavement using attributes such as; the classified Initial Pavement Condition (IPC), Age of pavement, Resilient Modulus (M_R) of subgrade soil, Average Truck load per day, Average Annual Air Temperature and Rainfall to predict the Future Pavement Condition (FPC). The model was calibrated using the observed logical behaviour of road pavement to fit the engineering experience and judgement. A goodness-of-fit test between the observed and predicted FPC values showed high level of consistency at 90%. There were no sufficient and well-spread dataset that described pavement behavioural pattern in present Nigeria to calibrate further the Fuzzy logic rules. The process proposed 5120 mutually exclusive Fuzzy logic rules for performance prediction of road pavement based on permutation theory. Though, the required well-spread dataset for calibration of the model to cover all possible pavement conditions in Nigeria and subsequent validation were not available, a framework for performance prediction of flexible road pavement was developed, and the study presented comprehensive guidelines on how to calibrate the FIS model using well-spread dataset.

Appendix

Rule No.	Initial pavement condition (IPC)	Age (years)	Truck load (veh/day)	MR of subgrade (kg/m^2)	Temp. ($^\circ$C)	Rainfall (mm)	Final pavement condition (FPC)
1.	Excellent	New	Low	Low	Low	Low	Excellent
2.	Excellent	New	Low	Low	Low	Medium	Excellent
3.	Excellent	New	Low	Low	Low	High	Excellent
4.	Excellent	New	Low	Low	Low	V. High	Excellent
5.	Excellent	New	Low	Low	Medium	Low	Excellent
6.	Excellent	New	Low	Low	Medium	Medium	Excellent
7.	Excellent	New	Low	Low	Medium	High	Excellent
8.	Excellent	New	Low	Low	Medium	V. High	Excellent
9.	Excellent	New	Low	Low	High	Low	Excellent
…	…	…	…	…	…	…	…
1656.	Good	Old	Medium	V. High	Medium	V. High	Poor
1657.	Good	Old	Medium	V. High	High	Low	Poor
1658.	Good	Old	Medium	V. High	High	Medium	Poor
1659.	Good	Old	Medium	V. High	High	High	Poor
1660.	Good	Old	Medium	V. High	High	V. High	Poor
1661.	Good	Old	Medium	V. High	V.High	Low	Poor
1662.	Good	Old	Medium	V. High	V.High	Medium	Poor
1663	Good	Old	Medium	V. High	V.High	High	Poor
…	…	…	…	…	…	…	…
2370.	Fair	Recent	Medium	Low	Low	Medium	Poor
2371.	Fair	Recent	Medium	Low	Low	High	Poor
2372.	Fair	Recent	Medium	Low	Low	V. High	Poor
2373.	Fair	Recent	Medium	Low	Medium	Low	Poor
2374.	Fair	Recent	Medium	Low	Medium	Medium	Poor
2375.	Fair	Recent	Medium	Low	Medium	High	Poor
2376.	Fair	Recent	Medium	Low	Medium	V. High	Poor

(*continued*)

(*continued*)

Rule No.	Initial pavement condition (IPC)	Age (years)	Truck load (veh/day)	MR of subgrade (kg/m^2)	Temp. ($^\circ$C)	Rainfall (mm)	Final pavement condition (FPC)
2377.	Fair	Recent	Medium	Low	High	Low	Poor
…	…	…	…	…	…	…	…
4089.	Poor	V.Old	V.High	V. High	High	Low	V.Poor
4090.	Poor	V.Old	V.High	V. High	High	Medium	V.Poor
4091.	Poor	V.Old	V.High	V. High	High	High	V.Poor
4092.	Poor	V.Old	V.High	V. High	High	V. High	V.Poor
4093.	Poor	V.Old	V.High	V. High	V.High	Low	V.Poor
4094.	Poor	V.Old	V.High	V. High	V.High	Medium	V.Poor
4095.	Poor	V.Old	V.High	V. High	V.High	High	V.Poor
4096.	Poor	V.Old	V.High	V. High	V.High	V. High	V.Poor
…	…	…	…	…	…	…	…
5114.	V.Poor	V.Old	V.High	V. High	High	Medium	V.Poor
5115.	V.Poor	V.Old	V.High	V. High	High	High	V.Poor
5116.	V.Poor	V.Old	V.High	V. High	High	V. High	V.Poor
5117.	V.Poor	V.Old	V.High	V. High	V.High	Low	V.Poor
5118.	V.Poor	V.Old	V.High	V. High	V.High	Medium	V.Poor
5119.	V.Poor	V.Old	V.High	V. High	V.High	High	V.Poor
5120.	V.Poor	V.Old	V.High	V. High	V.High	V. High	V.Poor

References

AASHTO: AASHTO Guide for the Design of Pavement Structures. American Association of State Highway and Transportation Officials, Washington, D.C. (1993)

Abiola, O.S., Owolabi, A.O., Odunfa, S.O., Olusola, A.: Investigation into causes of premature failure of highway pavements in Nigeria and remedies. In: Proceedings of the Nigeria Institution of Civil Engineers (NICE) Conference (2010)

Abiola, O.S., Owolabi, A.O., Sadiq, O.M., Aiyedun, P.O.: Application of dynamic artificial neural network for modelling ruts depth for lagos-ibadan expressway, Nigeria. Asian Research Publishing Network (ARPN)-J. Eng. Appl. Sci. **7**(8), 986–991 (2012)

Adefemi, B.A., Ibrahim, A.A.: Flexible pavement assessment of selected highways in Ifelodun local government, Ikirun-Osun, South – Western Nigeria. Int. J. Eng. Technol. **5**(8), 475–484 (2015)

Adeke, P.T., Atoo, A.A., Joel, E.: A policy framework for efficient and sustainable road transport system to boost synergy between urban and rural settlements in developing countries: a case of Nigeria. In: 1st International Civil Engineering Conference (ICEC 2018), Department of Civil Engineering, Federal University of Technology, Minna, Nigeria (2018a)

Adeke, P.T., Atoo, A.A., Orga, S.G.: Assessment of pavement condition index: a case of flexible road pavements on the university of agriculture Makurdi Campus. Niger. J. Technol. **38**(1), 15–21 (2018b)

African Development Bank Group: Rail infrastructure in Africa – Financing Policy Options. International d'Abidjan, Abidjan, Côte d'Ivoire (2015)

Aggarwal, P., Kumar, N.: Fuzzy model for road roughness index. In: International Conference on Biological, Civil and Environmental Engineering, Indonesia, 3–4 February 2015

Ahishakiye, E., Taremwa, D., Omulo, E.O., Niyonzima, I.: Crime prediction using decision tree (J48) classificaiton algorithm. Int. J. of Comp. and Info. Techn. **6**(3), 188–195 (2017)

Alam, F., Pachauri, S.: Comparative study of J48, naïve bayes and one-r classification techniques for credit card fraud detection using WEKA. Adv. in Comp. Sci. and Tech. **10**(6), 1731–1743 (2017)

American Society for Testing and Materials (ASTM): Standard Practice for Road and Parking Lots Pavement Condition Index Survey. D6433-07, Philadelphia (2007)

Arliansyah, J., Maruyama, T., Takahashi, O.: A development of fuzzy pavement condition assessment. Proc. Jpn. Soc. Civ. Eng. **746**, 275–285 (2003)

ASTM D6433-07 2007: Standard Practice for Road and Parking Lots Pavement Condition Index Survey. American Standard for Testing and Materials, Philadelphia

Ayed, A.: Development of Empirical and Mechanistic Empirical performance Models at Project and Network Levels. PhD Thesis, Department of Civil Engineering, University of Waterloo, Canada (2016)

Bandara, N., Gunaratne, M.: Current and future pavement maintenance prioritisation based on rapid visual condition evaluation. J. Transp. Eng. ASCE **127**, 116–123 (2001)

Bianchini, A., Bandini, P.: Prediction of pavement performance through Neuro-fuzzy reasoning. Comput. Aided Civ. Infrastruct. Eng. **25**, 39–54 (2010)

Chandran, S., Isaac, K.P., Veeraragavan, A.: Prioritization of low-volume pavement sections for maintenance by using fuzzy logic. Transp. Res. Rec. J. Transp. Res. Board **1989-1**(1), 53–60 (2007). https://doi.org/10.3141/1989-06

Chassiakos, A.P.: A fuzzy-based system for maintenance planning or road pavements. In: Proceedings of the 10th WSEAS International Conference on Computers, Vouliagmeni, Athens, Greece, pp. 535–540 (2006)

Chen, C., Flintsch, G.W.: Calibrating fuzzy-logic-based rehabilitation decision models using the LTPP database. In: 7th International Conference on Managing Pavement Assets (2008)

Chen, G., Pham, T.T.: Introduction to Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems. CRC Press, Boca Raton (2001)

Cheu, R.L., Wang, Y., Fwa, T.F.: Genetic algorithm-simulation methodology for pavement maintenance scheduling. Comput. Aided Civ. Infrastructural Eng. **19**, 446–455 (2004)

Chikezie, C.U., Olowosulu, A.T., Abejide, O.S., Kolo, B.A.: Review of application of genetic algorithms in optimization of flexible pavement maintenance and rehabilitation in Nigeria. World J. Eng. Pure Appl. Sci. **1**(3), 68–76 (2011)

Claros, G., Carmichael, R.F., Harvey, J.: Development of Pavement Evaluation Unit and Rehabilitation Procedure for Overlay Design method. Lagos: Texas Research and Development Foundation for the Nigeria Federal Ministry of Works and Housing (1986)

Cubero-Fernandez, A., Rodriguez, F.J., Villatoro, R., Olivares, J., Palomares, J.M.: Efficient pavement crack detection and classification. EURASIP J Image Video Process. **39**, 2–11 (2017)

Dabous, S.A., Al-Khayyat, G., Feroz, S.: Utility-based road maintenance prioritization method using pavement condition rating. Balt. J. Road Bridg. Eng. **15**(1), 126–146 (2019a)

Dabous, S.A., Zeiada, W., Al-Ruzouq, R., Hamad, K., Al-Khayyat, G.: Distress-based evidential reasoning method for pavement infrastructure condition assessment and rating. Int. J. Pavement Eng., 1–12 (2019b). https://doi.org/10.1080/10298436.2019.1622012

Dong, S., Zhong, J., Hao, P., Zhang, W., Chen, J., Lei, Y., Schneider, A.: Mining multiple association rules in LTPP database: an analysis of asphalt pavement thermal cracking distress. Constr. Build. Mater. **191**, 837–852 (2018)

Fox, C.: Data Science for Transport; Self-Study Guide with Computer Exercises. Springer, Switzerland (2018)

Fwa, T.F, Tan, C.Y., Chan, W.T.: Road-maintenance planning using genetic algorithms. II: analysis. J. Transp. Eng. **120**(5), 710–722 (1994)

Fwa, T.F., Shanmugam, R.: Fuzzy logic technique for pavement condition rating and maintenance-needs assessment. In: Fourth International Conference on Managing Pavements, Durban, South Africa, pp. 465–476 (1998)

Gogoi, R., Dutta, B.: Maintenance prioritisation of interlocking concrete block pavement using fuzzy logic. Int. J. Pavement Res. Technol. **13**(2020), 168–175 (2019). https://doi.org/10.1007/s42947-019-0098-9

Golroo, A., Tighe, S.L.: Fuzzy Set approach to condition assessments of novel sustainable pavements in the Canadian climate. Can. J. Civ. Eng. **36**, 754–764 (2009)

Gong, H., Sun, Y., Shu, X., Huang, B.: Use of random forest regression for predicting IRI of asphalt pavements. Const. and Building Mat. **189**, 890–897 (2018)

Gopalakrishnan, K., Agrawal, A., Ceylan, H., Kim, S., Choudhary, A.: Knowledge discovery and data mining in pavement inverse analysis. Transport **28**(1), 1–10 (2013). https://doi.org/10.3846/16484142.2013.777941

Government of The Federal Republic of Nigeria: General Specifications (Roads and Bridges), vol. II, Revised, Abuja, Nigeria (2016)

Hamed, R.I., Kakarash, Z.A.: Evaluate the asphalt pavement performance of rut depth based on intelligent method. Int. J. Eng. Comput. Sci. **5**(1), 15474–15481 (2016)

Heiman, G.W.: Basic Statistics for Behavioral Sciences, 6th edn. Cengage Learning, Belmont (2011)

Highway Research Board of the NAS-NRC Division of Engineering and Industrial Research Special Report 61G: The AASHO Road Test, Report 7 Summary Report, National Academy of Sciences – National Research Council, Washington, D.C. (1962)

Huang, Y.: Evaluating pavement response and performance with different simulative tests. PhD Thesis, Virginia Polytechnic Institute and state University, Virginia (2017)

Huang, Y.H.: Pavement Analysis and Design, 2nd edn. Pearson Prentice Hall, Inc., Upper Saddle River (2004)

Inkoom, S., Sobanjo, J., Barbu, A., Niu, X.: Prediction of the Crack condition of highway pavements using machine learning models. Struc. and Infrast. Eng. (2018). https://doi.org/10.1080/15732479.2019.1581230

Inkoom, S., Sobanjo, J., Barbu, A., Niu, X.: Pavement crack rating using machine learning frameworks: partitioning, bootstrap forest, boosted tree, Naïve Bayes, and K-Nearest neighbors. J. Transp. Eng., Part B: Pavement **145**(3), 1–12 (2019)

Jang, W., Lee, J.K., Lee, J., Han, S.H.: Naïve bayesian classifier for selecting good/bad projects during the early stage of international construction bidding decisions. Math. Prob. Eng. **2015**, 1–12 (2015)

Jeong, H., Kim, H., Kim, K., Kim, H.: Prediction of flexible pavement deterioration in relation to climate change using fuzzy logic. J. Infrastruct. Syst. Am. Soc. Civ. Eng. (ASCE) **23**(4), 04017008-1 (2017). https://doi.org/10.1061/(asce)is.1943-555x.0000363

Karagahin, M., Terzi, S.: Performance model for asphalt concrete pavement based on the fuzzy logic approach. Transport **29**(1), 18–27 (2014)

Kaur, D., Tekkedil, D.: Fuzzy expert system for ashpahl pavement performance prediction. In: IEEE Intelligent Transport Systems, Conference Proceedings, Dearborn, USA (2000)

Kirbas, U., Karasahin, M.: Performance models of hot mix asphalt pavement in urban roads. Constr. Build. Mater. **116**, 281–288 (2016)

Koduru, H.K., Xiao, F., Amirkhanian, S.N., Juang, C.H.: Using fuzzy logic and expert system approaches in evaluating flexible pavement distress: case study. J. Transp. Eng. ASCE **136**, 149–157 (2010)

Kudjo, P.K., Chen, J., Mensah, S., Amankwah, R., Kudjo, C.: The Effect of bellwether analysis on software vulnerability severity prediction models. Softw. Qual. J. (2020). https://doi.org/10.1007/s11219-019-09490-1

Li, Z., Chenge, C., Kwan, M., Tong, X., Tian, S.: Identifying Asphalt pavement distress using UAV LiDAR point cloud data and random forest classification. Int. J. Geo-Info. **8**(39), 2–26 (2019)

Lin, J.D., Huang, W.H., Hung, C.T., Chen, C.T., Lee, J.C.: Using decesion tree for data minig of pavement maintenace and management. Appl. Mech and Mat. **330**(2013), 1015–1019 (2013)

Liu, Y., Sun, M.: Fuzzy optimization BP neural network model for pavement performance assessment. In: IEEE International Conference on Grey Systems and Intelligent Services GSIS, Nanjing, China, pp. 1031–1034 (2007)

Luca, M.D., Abbondati, F., Pirozzi, M., Zilioniene, D.: Preliminary study on runway pavement friction decay using data mining. In: 6th Transport Research Area, 18–21 April 2016, vol. 14, pp. 3751–3760 (2016)

Mane, A.S., Gujarathi, S.N., Arkatkar, S.S., Sarkar, A.K., Singh, A.P.: Methodology for pavement condition and maintenance of rural roads. In: A National Conference on Fifteen Years of Transport Engineering Group, Civil Engineering Department, Indian Institution of Technology Roorkee, Roorkee – 247667 Uttarakhand (2016)

Mahmood, M.S.: Network-level maintenance decisions for flexible pavement using a soft computing-based framework. Ph.D. thesis, Nottingham Trent University, United Kingdom (2015)

Mahmood, M., Rahman, M., Nolle, L., Mathavan, S.: A Fuzzy logic approach for pavement section classification. Int. J. Pavement Res. Technol. Chin. Soc. Pavement Eng. **6**(5), 620–626 (2013). https://doi.org/10.6135/ijprt.org.tw/2013.6(5).620

Mahmood, M., Rahmood, M., Mathavan, S., Nolle, L.: pavement management: data centric rules and uncertainty management in section classification by a fuzzy inference system. Bitum. Mix. Pavement **6**, 533–541 (2015)

Marcelino, P. Antunes, M. de. L., Fortunato, E., Gomes, M.C.: Machine learning approach for pavement performance prediction. Int. J. Pavement Eng., 1–15 (2019). https://doi.org/10.1080/10298436.2019.1609673

Marianingsih, S., Utaminingrum, F.: Comparison of support vector machine classifier and naïve bayes classifier on road surface type classification (2018). https://doi.org/10.1109/SIET.2018.8693113

Marianingsih, S., Utaminingrum, F., Bachtiar, F.A.: Road surface types classification using combined of K-Nearest neighbor and naïve bayes based on GLCM. Int. J. Adv. Soft. Comp. Appl. **11**(2), 15–27 (2019)

MATLAB: User Guide Manual. Mathworks Inc. USA (2015)

Miradi, M.: Knowledge discovery and pavement performance: intelligent data mining, a Ph.D. thesis submitted to the Section of Road and Railway Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands (2009)

Moazami, D., Behbahani, H., Muniandy, R.: Pavement rehabilitation and maintenance prioritization of urban roads using fuzzy logic. Expert Syst. Appl. **38**, 12869–12879 (2011)

Mubaraki, M.A.: Predicting deterioration for the Saudi Arabia Urban road network. Ph.D. thesis, Department of Civil Engineering, University of Nottingham, United Kingdom (2010)

Munakata, T.: Fundamentals of New Artificial Intelligence, Neural Evolutionary, Fuzzy and More, 2nd edn. Computer and Information Science Department Cleveland State University, USA (2008)

Murana, A.A., Olowosulu, A.T., Otuoze, H.S.: Minimum threshold of monte carlo cycles for Nigerian empirical-mechanistic pavement analysis and design system. Niger. J. Technol. **31**(3), 321–328 (2012)

Murana, A.A.: Characterisation of subgrade materials from some Nigerian sources for use in the Nigeria empirical-mechanistic pavement analysis and design system, Ph.D. thesis, Department of Civil Engineering, Faculty of Engineering, Ahmadu Bello Universiy Zaria, Nigeria (2016)

Murana, A.A., Olowosulu, A.T.: Evaluation of rutting models using reliability for mechanistic-empirical design of flexible pavement. J. Eng. Appl. Sci. **7**(2), 123–127 (2012b)

Nawir, D., Prihartanto, E.: Decision-making analysis of road maintenance in north kalimantan region using technology of fuzzy logic (case study: Liang Bunyu Street Section, West Sebatik Sub-District, Nunukan District). IOP Conf. Ser. Earth Environ. Sci. **353,** 012054 (2019). https://doi.org/10.1088/1755-1315/353/1/012054

Obuandike, G.N., Audu, I., John, A.: Analytical Study of some selected classification algorithms in WEKA Using real crime dataset. Int. J. Adv. Res. Artif. Intell. **4**(12), 44–48 (2015)

Pan, Y., Zhang, X., Cervone, G., Yang, L.: Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. **1–12**, (2018). https://doi.org/10.1109/JSTARS.2018.2865528

Pantuso, A., Flintsch, G.W., Katicha, S.W., Loprencipe, G.: development of network-level pavement deterioration curves using the linear empirical Bayes approach. Int. J. Pavement Eng. (2019). https://doi.org/10.1080/10298436.2019.1646912

Pawlak, Z.: Rough sets and intelligent data analysis. Inf. Sci. **147**, 1–12 (2002)

Premkumar, L., Vavrik, W.R.: Enhancing pavement performance prediction models for the Illinois Tollway system. Int. J. Pavement Res. Technol. **9**, 14–19 (2016). https://doi.org/10.1016/j.ijprt.2015.12.002

Road Sector Development Team: Configuration and Calibration of HDM-4 to Nigeria Conditions, Government of the Federal Republic of Nigeria, Nigeria, p. 33 (2014)

Russell, S.J., Norvig, P.: Artificial Intelligence. A modern Approach, 3rd edn. Pearson Education Inc., London (2010)

Salpisoth, H.: Simple evaluation methods for road pavement management in developing countries. Ph.D. thesis - Graduate School of Engineering, Kyoto University, Japan (2014)

Saltan, M., Terzi, S., Kucuksille, E.U.: Backcalculation of pavement layer moduli and poisson's ratio using data mining. Expert Syst. Appl. **38**, 2600–2608 (2011)

Santos, J., Torres-Machi, C., Morillas, S., Cerezo, V.: A Fuzzy Logic expert system for selecting optimal and sustainable life cycle maintenance and rehabilitation strategies for road pavements. Int. J. Pavement Eng. 1–13 (2020). https://doi.org/10.1080/10298436.2020.1751161
Saravanan, N., Gayathri, V.: Performance and classification evauation of J48 algorithm and kendalls based J48 algorithm (KNJ48). Int. J. Comp. Intell. Inf. **7**(4), 188–198 (2018)
Savio, D., Nivitha, M.R., Bindhu, B.K., Krishnan, J.M.: Overloading analysis of Bituminous pavements in India using M-E pavement design guide. In: 11th Transport Planning and Implementation Methodologies for Developing Countries, Mubai, India (2016)
Setyawan, A., Nainggolan, J., Budiarto, A.: Predicting the remaining service life of road using pavement condition index. In: The 5th International Conference of Euro Asia Civil Engineering Forum (2015)
Sharma, T.C., Jain, M.: WEKA approach for comparative study of classification algorithm. Int. J. Adv. Res. Comput. Commun. Eng. **2**(4), 1925–1931 (2013)
Shekharan, A.R.: Effect of noisy data on pavement performance prediction by artificial neural networks. Trans. Res. Rec. **1643**, 7–13 (1998)
Shoukry, S.N., Martinelli, D.R., Reigle, J.A.: Universal Pavement distress evaluator based on fuzzy sets. Trans Res. Rec. **1592**, 180–186 (1997)
Smadi, O.G.: Knowledge based expert system pavement management optimisation. Paper based on Ph.D. Dissertation, Iowa State University, Ames, Iowa, United States (2000)
Surendrakuma, K., Prashant, N., Mayuresh, P.: Application of Markovian probabilistic process to develop a decision support system for Pavement maintenance management. Int. J. Sci. Technol. Res. **2**(8), 295–303 (2013)
Taylor, M.A.P., Philip, M.L.: Investigating the impact of maintenance regimes on the design life of road pavements in a changing climate and the implications for transport policy. Transp. Policy **41**, 117–135 (2015)
Thube, D.T.: Artificial Neural Network (ANN) based pavement deterioration models for low volume roads in India. Int. J. Pavement Res. Technol. **5**(2), 115–120 (2011)
Transport Research Board of the National Academies: Transport Research Circular E-C118 - Pavement Lessons Learned from the AASHO Road Test and Performance of the Interstate Highway System, Washington, DC 20001 (2007)
Tribhuvan, A.P., Tribhuvan, P.P., Gade, J.G.: Applying naïve bayesian classifier for predicting performance of a student using WEKA. Adv. Comp. Res. **7**(1), 239–242 (2015)
Uglova, E.V., Tiraturyan, A.N.: Interlayer bond evaluation in the flexible pavement structures using a non-destructive testing method. In: International Conference on Industrial Engineering (2016)
Wang, K.C.P., Li, L.: Pavement smoothness prediction based on fuzzy and gray theories. Comput. Aided Civ. Infrastruct. Eng. **26**(1), 69–76 (2011)
WEKA: Waikato Environment for Knowledge Analysis: User's Manual, The University of Waikato, New Zealand (2018)
Witten, I.H., Frank, E.: Data Mining – Practical Machine learning Tools and Techniques, 2nd edn. Morgan Kaufmann, Elsevier, London (2005)
Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: The WEKA Workbench – Data Mining Practical Machine Learning Tools and Techniques, 4th edn. Morgan Kaufmann, Elsevier, London (2016). Online Appendix, Retrieved on 6th February, 2020. From: www
Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: The WEKA Workbench – Data Mining Practical Machine Learning Tools and Techniques, 4th edn. Morgan Kaufmann, Elsevier, London (2016)
Woo, S., Yeo, H.: Optimization of pavement inspection schedule with traffic demand prediction. In: 11th International Conference of the International Institute for Infrastructure Resilience and Reconstruction (13R2) (2016). Procedia – Social and Behavioural Sciences 218, 95–103
Xiao, D.X., Wu, Z.: Using systematic indices to relate traffic load spectra to pavement performance. Int. J. Pavement Res. Technol. **9**, 302–312 (2016)

Yin, H.: Integrating instrumentation data in probabilistic performance prediction of flexible pavement, Ph.D. thesis in the Department of Civil and Environmental Engineering, Graduate School, The Pennsylvania State University (2007)

Yu, J.: Pavement service life estimation and condition prediction. Ph.D. thesis, Department of Civil Engineering, University of Toledo (2005)

Ziliute, L., Motiejunas, A., Kleiziene, R., Bribulis, G., Kravcovas, I.: Temperature and Moisture Variation in Pavement Structures of the Test Road. 6th Transport Research Arena (2016)