Heliyon 10 (2024) e25625

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

A Bayesian decision support system for optimizing pavement management programs

Babitha Philip [a,b], Hamad AlJassmi [a,b,*]

[a] Department of Civil and Environmental Engineering, UAE University, Al Ain, P.O.Box 15551, United Arab Emirates
[b] Emirates Center for Mobility Research (ECMR), UAE University, Al Ain, P.O.Box 15551, United Arab Emirates

ARTICLE INFO

Keywords:
Bayesian belief networks
Machine-learning
Decision-support
Uncertainty

ABSTRACT

Over time, the pavement deteriorates due to traffic and the environment, resulting in poor riding quality and structural inadequacies. Evaluating pavement condition over time is thus a critical component of any pavement management system (PMS) to extend the service life of pavements. However, the uncertainty associated with the pavement deterioration process due to the heterogeneous nature of the pavement degradation factors makes the process difficult. The current work addresses this challenge of pavement management by developing an expert system framework based on Bayesian Belief Networks (BBN). This framework integrates data on existing road deterioration factors with knowledge gained from pavement experts to produce optimal decisions. The advantages of the BBN techniques lie in their ability to capture uncertainty, and probabilistically infer the values of variables in the domain, especially in the case of incomplete information where we only have data about some and not all variables. This has motivated the adoption of BBN in this study to optimize pavement maintenance decisions, on the basis of inferred road deterioration interpretations drawn from partial knowledge about road distress variables. This study presents the adoption of Bayesian methods to assist pavement maintenance engineers in determining the most successful and efficient maintenance and repair (M&R) tactics and the best time to apply them by means of a decision-support system. Data collected from 32 road sections in the United Arab Emirates in relation to road distress parameters (rutting, deflection, cracking, and international roughness index), as well as road characteristics, traffic, and environment data, has been used to demonstrate the applicability of the proposed decision-support tool.

1. Introduction

The demand for high-quality, well-maintained, and safe roads is often compromised by budgetary restrictions. Pavement management systems (PMS) are used by highway agencies to aid in the formulation of policies that could maximize resources and implement maintenance strategies that are appropriate for each road segment in a manner that benefits the overall pavement network. One of the primary requirements of a PMS is an effective pavement deterioration prediction model [1]. However, the inclusion of a feature that enables the optimization process for maintenance activities in addition to forecasting the future pavement conditions would ensure high potential pavement conditions [2]. Moreover, the combination of features that optimize maintenance activities

* Corresponding author. Emirates Center for Mobility Research (ECMR), UAE University, Al Ain, P.O.Box 15551, United Arab Emirates.
E-mail address: h.aljassmi@uaeu.ac.ae (H. AlJassmi).

https://doi.org/10.1016/j.heliyon.2024.e25625
Received 23 March 2023; Received in revised form 5 January 2024; Accepted 31 January 2024
Available online 6 February 2024
2405-8440/

based on future pavement conditions would significantly reduce the lifecycle costs of the pavement network. Nevertheless, researchers have focused either on developing an optimization tool for the pavement management or pavement deterioration prediction models [3].

Various maintenance and rehabilitation (M&R) activities are adopted by highway agencies to control the negative impacts of pavement deterioration. The majority of the traditional M&R procedures are time-consuming or depend upon the context of their implementation. Machine learning (ML) techniques with the ability to handle big data have taken up the top position in solving problems related to the transportation sector. Recent research has shown the effectiveness of adopting certain ML algorithms to solve the problem of road maintenance optimization and prioritization namely, Clustering Page Rank Algorithm (CPRA) [4], Pattern-Recognition algorithm [5], Genetic Algorithms (GA) [6], Particle Swarm Optimization (PSO) algorithm [7] and so on. Regarding the prediction models, the major two groups of prediction models are deterministic and probabilistic models [1]. Several approaches have been adopted to develop prediction models such as Regression analysis [8], Decision-trees [9], Artificial neural networks [10], Markov chain models [11], Bayesian approaches [12], and many others. Compared with deterministic models, probabilistic models have a better ability to handle the complexity associated with the pavement deterioration process [13].

As is evident in the ongoing research, considerable efforts have been taken to utilize Artificial Intelligence (AI) methods in decision-making for pavement maintenance. ML models are highly promising for present and future pavement management practice because they can better capture the complicated nonlinear correlations between numerous factors. However, the decision-making process faces certain challenges, such as expensive data collection procedures. For instance, the International Roughness Index (IRI) is frequently used to analyze the quality of the pavement, make repair choices, evaluate ride comfort, and calculate vehicle running expenses. Certain road classes are, however, omitted from IRI measurements at the network level since measuring IRI is often expensive. Yet another challenge faced in pavement management is the difficulty in simulating the different scenarios experienced by the pavement structure under realistic climatic and traffic conditions which makes the traditional laboratory testing unreasonable [14]. This study adopts Bayesian Belief Networks (BBN) to address the need for a comprehensive decision-making tool considering its ability to capture uncertainty and probabilistic estimation of unknown values to overcome the challenges in pavement management caused due to heterogeneous nature of the factors leading to pavement deterioration and requirement of huge inspection data by many of the statistical approaches [15–17]. This study proposes a novel framework to develop a decision support system based on BBN techniques to combine expert knowledge and available data to analyze different pavement scenarios to produce optimum decisions in relation to scheduling pavement maintenance and rehabilitation activities. Previous studies related to this work is also elaborated [18]. Bayesian decision support system presented in this study shows how the BBN method can be used to develop more realistic pavement management, which can be further enhanced with additional data in the future.

The remainder of this paper is organized as follows. Section 2 discusses different Bayesian approaches adopted by researchers in pavement studies. The proposed Bayesian platform for pavement management is described in Section 3. Section 4 demonstrates the working of the Bayesian DSS based on the data collected. The conclusions of this study are summarized in Section 5.

2. Bayesian methods in pavement management strategies

Bayesian techniques have been widely used in pavement studies for various applications such as estimating the correlation of pavement friction with the risk of traffic crashes at intersections [19], recovering missing traffic data from toll data and video surveillance data [20], predicting thermal conductivity in cement composites in pavements [21] and many others. The uncertainties associated with factors that have a detrimental impact on the long-term performance of road networks and the existence of significant heterogeneity among various locations and climate conditions have motivated researchers to deeply investigate the strength of BBN methods for better pavement management solutions [22].

Blumenfeld et al. [23] adopted Bayesian filters to estimate the current pavement condition and future deterioration. A section-based forecast model was developed by combining empirical models for pavement data collected from Germany, Austria, and Switzerland. The model developed had the capability to include uncertainty associated with both current and future measurements and is applicable for a wide variety of physical road condition characteristics.

Heba et al. [24] proposed the use of the Bayesian linear regression method that relies on the expertise of experts as a prior distribution to create a performance model in the absence of historical data for Libyan roads. In order to aid and support the input data for feeding the Bayesian model, experts from Libyan Road and Transportation Agency were interviewed. Further, a few inspections were carried out on-site to calculate the posterior distribution, which along with the expert knowledge determined the future pavement conditions.

Apart from the ability to deal with uncertainty and missing information, BBNs are more interpretable and transparent when compared to other AI methods such as Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS). ANN models are referred to as ‘black boxes’ due to less transparency in the prediction modeling. Similarly, inferences obtained from FIS are difficult to understand while dealing with complex relations. In contrast, BBN methods provide an efficient representation of the causal links among the variables, making them ideal for decision-making.

Although previous work has utilized Bayesian techniques to evaluate pavement conditions, this study contributes to the state-of-the-art by providing a framework for understanding the pavement performance by combining the principles of BBN and various ML algorithms. This enables those responsible for maintaining pavements to obtain better future estimates of the pavement characteristics, thereby allowing decision-makers to adopt the best decisions from among various scenarios. To achieve this objective, this study applies big data in cooperation with the government entity, the Ministry of Energy and Infrastructure (MoEI) in the United Arab Emirates (UAE).

3. Bayesian Decision Support System (BDSS)

Pavement management has grown increasingly crucial as pavements age and deteriorate, and funding levels have fallen owing to decreasing financing or increased competition for resources. Pavement management is essentially the practice of managing the pavement infrastructure in a cost-effective manner. Pavement management is a technique that consists of several phases that will assist the pavement maintainers in analyzing work plan alternatives. Many organizations are bound by limited funding and human resources. Despite these limits, highway authorities spend on pavement maintenance because pavement management equips an agency with the tools it needs to address management challenges, mainly the optimal allocation and usage of available resources [25].

One important issue that highway and transportation organizations face is that the funding they receive is frequently insufficient to effectively repair and rehabilitate every deteriorated roadway section. The problem is worsened by the fact that roads can often be used when in poor condition, making it simple to postpone maintenance operations until conditions become unacceptable. Roadway deterioration is typically caused by the inevitable wear and tear that occurs over time, rather than by bad design and construction techniques. Many variables contribute to the progressive deterioration of a pavement, including changes in climate, drainage, soil conditions, and truck traffic. Lack of funding frequently prevents timely maintenance and renovation of transportation facilities, resulting in a larger problem with more significant consequences.

Because funding and manpower are frequently insufficient to meet needs, many transportation agencies are forced to balance their work programs between preventative maintenance activities and problems requiring rapid corrective action. As preventive maintenance has been overlooked, roads will worsen over time, and the level of complaints by road users will serve as the basis for restoration. The road-users are unwilling to accept low-quality pavements that create vibration and severe damage to their automobiles. Poor pavement quality may result in crashes, increasing end-user costs dramatically. Preventive maintenance should be performed in an ordered and methodical manner, as this is the least expensive approach in the long term. Nevertheless, if funds are limited, authorities frequently react to either the most critical problems or those that draw the most public complaints.

Thus, a systematic pavement management procedure is a critical part of an effective pavement network design and management. The major steps involved in a PMS are illustrated in Fig. 1.

The process of pavement management begins with the data collection and ends with updating the database with the new data. The core of any PMS lies from step 2 to step 4. This study proposes the adoption of BBN techniques to carry out steps 2 to step 4 in an efficient way capturing the uncertainty associated with the pavement deterioration process. This study is carried out based on the principles of BBN. A BBN structure is a directed acyclic graph, in which the domain variables are represented as nodes and the dependency among the variables is represented by arcs. The variables involved in a BBN structure can be either continuous, discrete, or both. Bayesian statistics aid in learning from our data and incorporating new information into future research. We do not rely on the usual (i.e., frequentist) framework's notion of infinitely repeating an event (or experiment). Instead, we add existing information and human judgment into the process to help with parameter estimates. Thus, the adoption of Bayesian statistics reduces the complexity of the model, allows the incorporation of background knowledge, and can produce inferences based on limited available information and data.

Bayesian Belief Networks (BBN) are efficient to represent uncertain knowledge regarding interrelationships among variables in a complex system [26]. The BBN structure functions as an inference mechanism to facilitate the probabilistic estimation of unknown values [27]. In scenarios with low-data, BBN enables inverse modeling, avoiding overfitting problems and providing insights into unobserved variables [28]. Due to the uncertainties inherent in road deterioration processes, researchers have increasingly adopted BBN for pavement data analysis. A significant advantage of BBN over other methods is its ability to combine the combination of prior knowledge and experience with data to yield improved insights, a critical aspect in the assessment of road performance [29]. Beyond managing uncertainties, BBN demonstrates proficiency in handling complex and nonlinear relationships among diverse factors [26]. The principles of BBN for the development of a DSS are described below.

3.1. Principles of Bayesian statistics

Consider two events A and B in a sample space. The probability of event A is represented by P(A) and the probability of event B is P(B). If events A and B share some outcomes but have certain outcomes that the other does not, they are said to be intersecting events. The intersection of the events is represented by $A \cap B$. According to the rules of probability, if event A has occurred, the updated probability of B can be calculated based on conditional probability. Hence, the conditional probability of event B given event A is given in Equation (1). Using the multiplication rule of probability, the joint probability $P(A \cap B)$ is substituted as in Equation (2), known as Bayes' theorem, discovered by Reverend Thomas Bayes [30].

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{1}$$

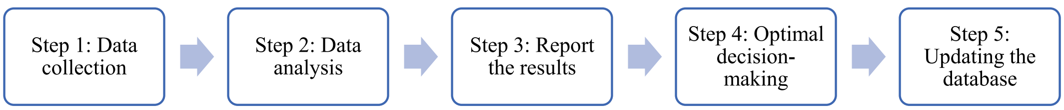

Fig. 1. Steps involved in PMS database development.

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)} \tag{2}$$

Thus, the application of Bayes' rule enables us to calculate the probability of an unknown value by observing an interrelated value. Generally, this type of analysis reduces the data collection efforts. In the current scenario where the number of motor vehicles is expected to increase in the coming decades, rehabilitation of the highway system is becoming a challenging task. Further, the shortage of highway budgets is motivating highway agencies to focus on tools and techniques to preserve the existing highway assets [31]. This study proposes a method for maintaining, preserving, and rehabilitating highway assets in an economical way by diminishing the quantity of data collection.

The nature of the unknown parameter estimated in a statistical model is the basic difference between Bayesian statistics and

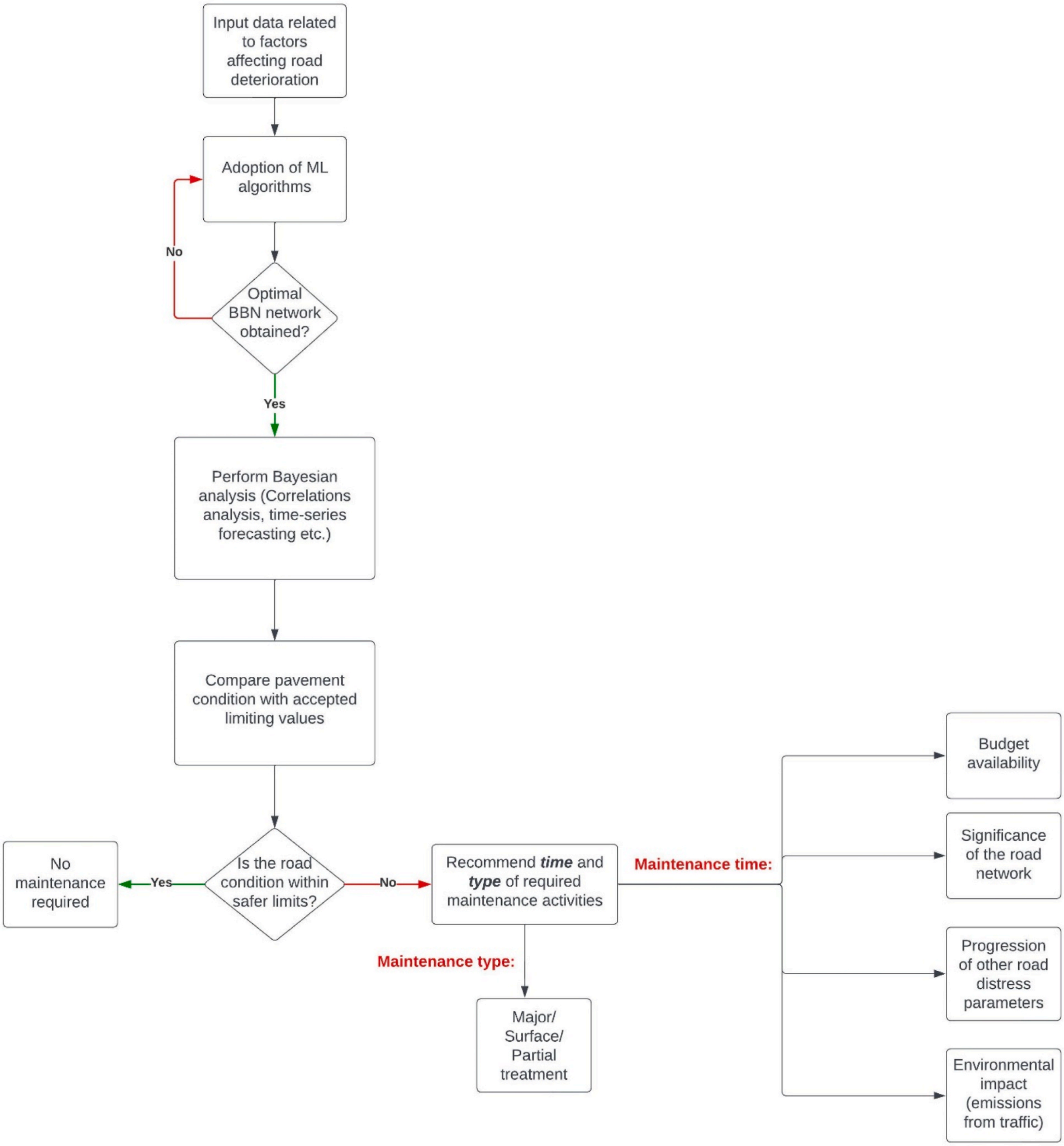

Fig. 2. Framework of BDSS.

conventional statistics. In conventional statistics, the unknown parameter is considered to be fixed and unknown. However, from a Bayesian perspective, the unknown parameter can incorporate uncertainty which can be described by a probability distribution. Thus, the outcome of a Bayesian analysis is not a single outcome, but instead a distribution with the probability of the unknown value. That is, it is assumed that each parameter has a distribution that captures uncertainty about its value. This uncertainty is reflected by the prior distribution (or prior), which is defined before witnessing the data. The observed evidence is then stated in terms of the data's likelihood function. The likelihood is then utilized to modify the prior, obtaining the posterior distribution, which is a mixture of the prior distribution and the likelihood function. Thus, the prior, likelihood, and posterior are the basic components employed in Bayes' theorem and hence the core principle of Bayesian statistics [32].

3.2. Framework of BDSS

Initially, the data gathered from the highway agencies (historical data) related to the road segments, and any other source of information such as expert knowledge, are used to calculate the prior probability distribution of each variable. Once the prior probabilities have been calculated, appropriate ML algorithms are adopted to develop an accurate Bayesian structure. The ML algorithms are majorly of two types, supervised and unsupervised. Supervised is adopted when the analysis has to be performed with respect to a target variable and unsupervised is adopted when there is no target. Hence, suitable algorithms are selected accordingly. The basic workflow of the proposed BDSS is given in Fig. 2.

Different algorithms produce different Bayesian structures, the best one among them is decided mainly based on the minimum description length (MDL) score. The MDL score is calculated by combining the complexity of the structure and the data it produces. The MDL score's guiding principles state that every data set may be trained using a model based on how well it can compress the data, and that the score indicates how generalizable the model is. The model with the lowest MDL score is considered to be the best one; this score is determined using Equation (3), where L(Data|Model) is the length of the data description provided to the model and L(Model) is the length of the model description.

$$MDL = L\ (Data|Model) + L(Mod\ el) \tag{3}$$

Finding the optimal network is followed by Bayesian analysis, which is the core part of the DSS, where different types of analysis can be performed. Exploring the interrelationships among the nodes is beneficial during missing data which also reduces the uncertainty associated with variables [18]. Another major strength of the Bayesian statistics is to capture the temporal aspect of the variables, which enables to forecast the future values of the nodes, as it follows the Markovian property and invariant transition probability [33].

Finally, the pavement maintainers are capable to utilize the inferences obtained through the Bayesian analysis for decision-making. The prior probabilities known from the previous knowledge are updated on the basis of evidence to revise the beliefs. The posterior probability of an event is the combination of the prior belief and the evidence (likelihood) obtained from the occurrence of another event [30]. The posterior probability distribution thus gives the decision-makers better insights regarding the pavement condition. Several factors affect the final decisions such as budget availability, the significance of the road network, the progression of other road

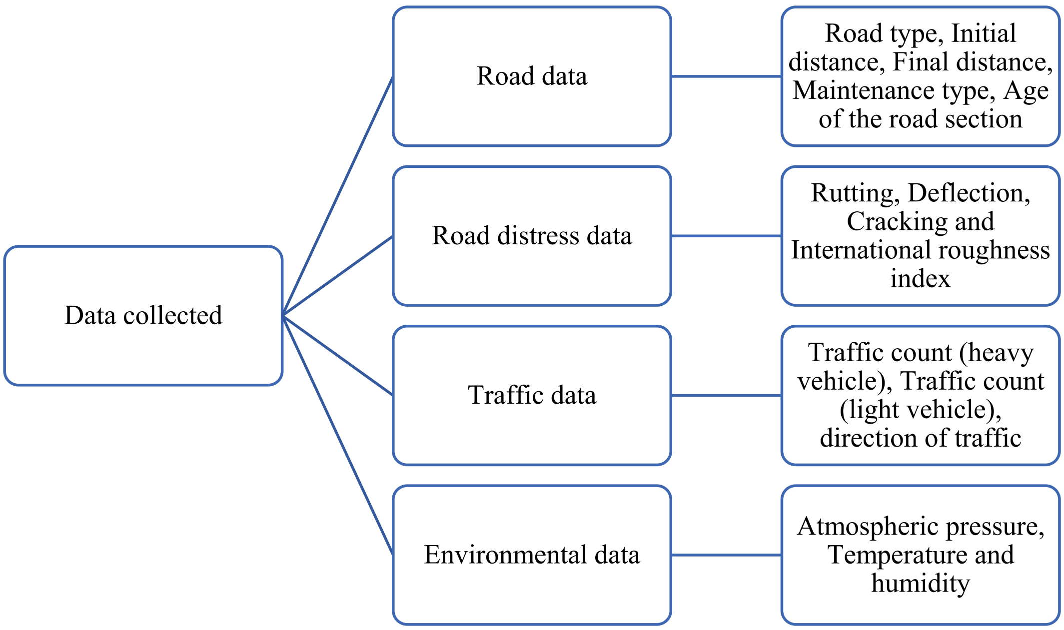

Fig. 3. Data used in this study to develop a BDSS.

distress parameters, and the impact on the environment.

4. Working of the proposed BDSS

4.1. Data description

The data used in this study is collected from the Road Department of the Ministry of Energy and Infrastructure (MoEI), United Arab Emirates. The data collected from the ministry include road data, road distress data, and traffic data from year 2013–2019 along 32 prominent road segments in the country. The environmental data were collected from online resources. A summary of the data collected is given in Fig. 3.

Several pre-processing procedures were applied to the data to correct differences in measurement intervals for different road distress parameters (for example, cracking was measured at an interval of every 10 m and deflection was measured at an interval of approximately 100 m) and to fill the gaps in the data. Several assumptions related to maintenance type and age of the road section were also adopted as explained in our previous work [18]. The final dataset for analysis had 3272 data points, each data point represents a 10-m length road segment.

4.2. Data analysis

The final dataset with 3272 data points is exported to the different stages of BDSS as mentioned in Fig. 2. This study has adopted BayesiaLab 10.2 software to perform the analysis [34]. Once the data is prepared for analysis, the prior probabilities of each variable are calculated. The prior probability distribution of the 15 variables involved in this study is given in Fig. 4.

Different machine-learning algorithms are adopted to train the data. In this study, the objective is to obtain a Bayesian structure to understand the current pavement condition on an overall basis and not with respect to a particular variable. Hence, unsupervised learning algorithm is adopted. The different algorithms applied to the data are Taboo learning, Maximum weight spanning tree, Taboo order, Equivalence class learning (EQ), SopLEQ, and Taboo EQ.

Corresponding to each of the algorithms adopted in the previous step, a Bayesian network is obtained. The optimal network can be

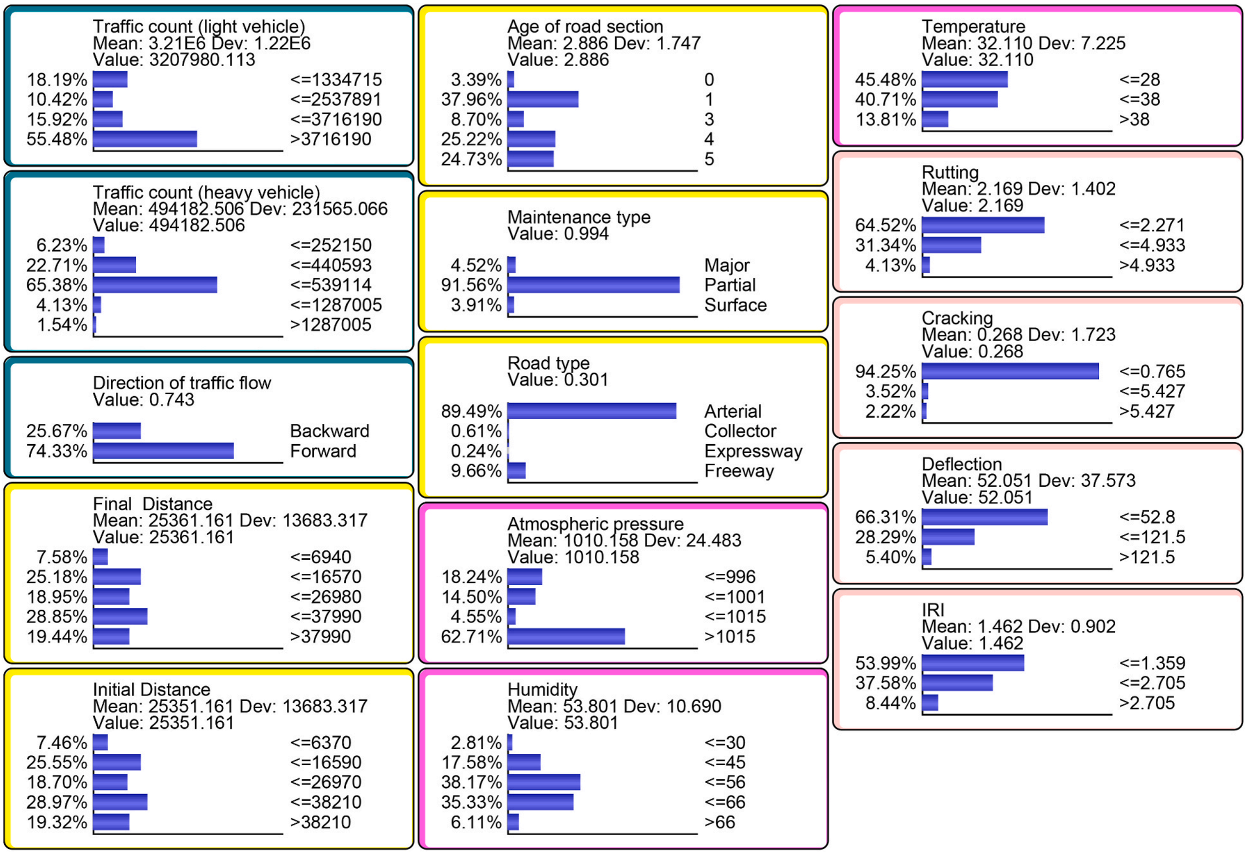

Fig. 4. Prior probability distribution of the variables.
*Variables coming under the same category are given similar color.

decided based on the MDL score, as calculated according to Equation (3). The MDL scores corresponding to the algorithms adopted in this study are given in Table 1.

Among the six structures, the Bayesian network structure corresponding to taboo order learning is observed to attain the least MDL score. Hence, the corresponding structure is considered to be the optimal one and is selected for further analysis.

Once the optimal network is obtained, several analyses can be performed. The observed evidence which is expressed as the likelihood function of the data, combines with the prior data based on Bayes' theorem generating posterior distribution of the variables. Future investigation of the data is carried out on the posterior distribution which reflects the updated knowledge of the variable balancing the prior knowledge and the observed data.

One type of analysis to be performed is the correlation analysis, which aids in understanding the interrelationships among the variables and the intensity of each relation. The amount of information related to one variable with respect to another interrelated variable is calculated by mutual information [35]. Fig. 5 represents the interrelationships obtained among the variables under study with mutual information on each arc.

The analysis can be further restricted based on the variables of interest, for instance, to explore the impact of heavy vehicle loads on road distress parameters as shown in Fig. 6.

Fig. 6 shows that the customized inferences required by pavement maintainers can be deduced, which favors better pavement management. The Kullback-Leibler (KL) divergence, relative weight of each relation, overall impact of each relation, and Pearson's correlation are additional metrics to represent the impact of each relations obtained as shown in Table 2.

The results indicate that the increase in the number of heavy vehicles may improve the chances of occurrence of deflection, while, cracking is least affected by the quantity of heavy vehicles.

Further extension of the pavement management capabilities of the proposed framework is the risk assessment of the road condition by comparing the current values of the road distress parameters with the prescribed safe limit. Since the allowed limit of road deterioration criteria varies from place to place, the optimal decision may be unique to each highway agency. To execute decisions, several function nodes can be added into the Bayesian structure in connection with the road distress parameter nodes. The limiting values adopted by the MoEI as per the discussion with the officials at the ministry in the year 2020 were: '5' for rutting, '2' for cracking, '40' for deflection, and '2' for IRI. In this study, decision nodes (function nodes) were included in the analysis by writing corresponding codes. Two sets of function nodes were added, one to return the mean value of the road distress parameter and the other to return the decision based on threshold values of the distress parameters. The mean value is returned by adding the inference function "MeanValue (v)", where (v) represents the distress parameter of interest. Explanation of the two sets of function nodes corresponding to the distress parameter rutting is given below:

- Function node set 1 (Mean value nodes): Will return the mean value of the road distress parameter
- Function node set 2 (Decision nodes): Will return the current road condition based on the acceptable value of the parameter, i.e., the function node will return 'crossed the safe limit' if the mean value of the road distress parameter is above the acceptable safe limit and 'within the safe limit' in the opposite case.

An illustration of the decision-making process is provided in Fig. 7, for the Bayesian structure obtained corresponding to the road distress parameters in our previous work [18].

Fig. 7(a) represents the Bayesian model which involves the parameters under study represented as nodes and the function nodes; function node set 1 representing the mean values of the road distress parameters and the function node set 2 representing the decisions or inferences of the model. The probability distributions of the parameters are shown in Fig. 7(b) and the output of the function nodes representing the mean value obtained from the probability distributions are shown in Fig. 7(c). It can be seen that the mean value of the IRI is 1.462 mm/m, deflection is 52.051 mm/100, rutting is 2.169 mm and cracking is 0.268 %. These values are compared to the limiting values adopted in the region. Fig. 7(d) gives the inferences related to the current status of the road condition, which is the output of the function nodes representing the decisions.

The mean values of the distress parameters are obtained based on the probability distribution of the node (road distress parameter), which is interpreted based on the historical data. It has to be noted that since the road distress parameters are expressed in terms of probability distributions other than a single value, the correlations drawn among the variables based on Bayesian statistics will aid in obtaining the probability distributions even if the available knowledge on the road distress parameters is limited. This would enable the decision-makers to have an estimate of the road distress parameters even without performing the actual measurement in the field. This can be considered to be the prominent advantage of a BDSS, which helps in reducing the data collection efforts and related

Table 1
MDL score corresponding to the machine-learning algorithms.

Machine-learning algorithm	MDL score
Taboo	48,556.48
Maximum weight spanning tree	53,225.539
Taboo order	48,431.897
EQ	48,567.216
SopLEQ	48,731.296
TabooEQ	48,567.216

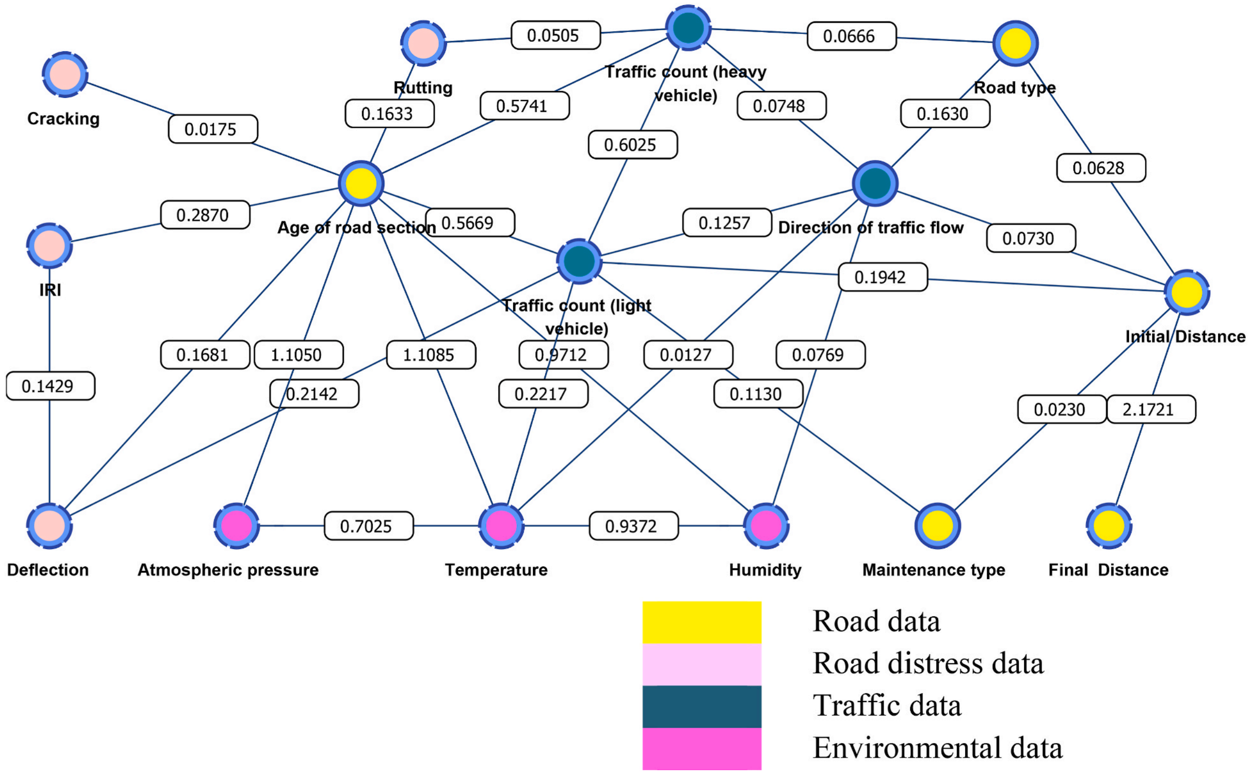

Fig. 5. Relationship analysis among the factors related to road deterioration.

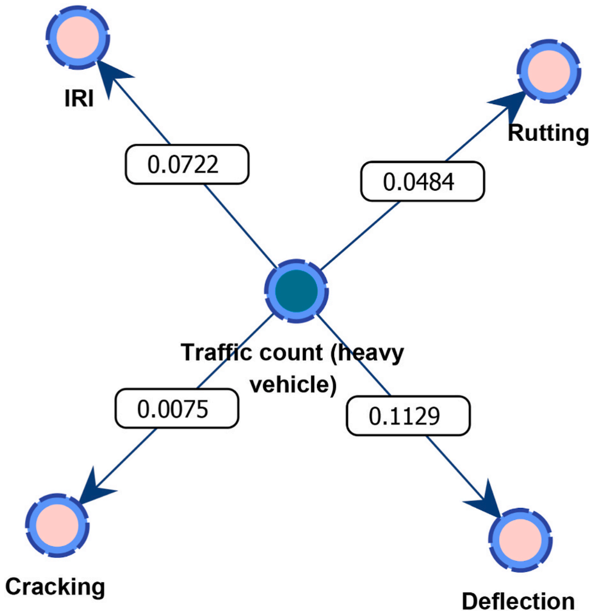

Fig. 6. Impact of traffic count (heavy vehicle) on road distress parameters.

expenses.

Once the current status of the road condition has been estimated in terms of the road distress values, the future behavior of the road distress parameters could be forecasted based on the principles of dynamic BBN. Dynamic BBN which follows the Markovian property and invariant transition probability in combination with the correlation analysis will contribute to the knowledge of future pavement

Table 2
Relationship analysis of traffic count (heavy vehicle) and road distress parameters.

Parent	Child	KL Divergence	Relative Weight	Overall Contribution	Pearson's Correlation
Traffic count (heavy vehicle)	Deflection	0.1129	1.0000	46.8295 %	0.1343
Traffic count (heavy vehicle)	IRI	0.0722	0.6400	29.9731 %	0.0256
Traffic count (heavy vehicle)	Rutting	0.0484	0.4286	20.0693 %	0.1655
Traffic count (heavy vehicle)	Cracking	0.0075	0.0668	3.1280 %	0.0012

deterioration trends. Having the knowledge of the current and future behavior of the road deterioration process, optimal decisions related to the type and time of maintenance activities could be adopted subject to:

1. Intensity of each road distress parameter: Knowing the intensity of each road distress parameter will enable the pavement maintainers to decide the type of maintenance activity. The maintenance activities adopted in the study region are major, surface, and partial. Hence, knowing the intensity of each road distress parameter will enable to select the appropriate type of maintenance. This is beneficial for highway agencies performing routine maintenance programs on a fixed time interval. These highway agencies usually conduct maintenance activities without estimating the values of the road distress parameters which result in the unnecessary usage of the highway budgets.
2. Budget availability: The correlation analysis will imply the significance of each road distress parameter. Hence, the maintenance activities can be delayed if the budgets are limited until the most significant factor crosses the safe limit value. For example, the most important parameter for the region under study is IRI [18]. Hence, maintenance activities can be delayed until the IRI rises above the allowed safe limit, even though other distress parameters cross the safe limit.
3. Significance of the road network: The maintenance decisions can be further optimized based on the significance of the road network itself. Considering the optimal utilization of the highway budgets, the decision-makers could delay the maintenance activity even if one among the prominent road distress parameters has crossed the safe limit in a less significant road. The significance of the road can be inferred based on the traffic flow. Although the most significant parameter crosses the safe limit, the optimal decision could be to delay the maintenance activity if the roads are less significant (for example, local roads are less significant than arterial roads).
4. Progression of other road distress parameters: Optimal maintenance decisions could be taken by considering the rate of progress of different road distress parameters across the years apart from their significance. If the rate of progress is slow, the maintenance activities can be delayed accordingly even when the significant factors have exceeded the allowed limit. On the other hand, maintenance activities should be urgently initiated if the distress parameters are progressing at a fast rate.
5. Environmental impact (emissions from traffic): Reduced greenhouse gas (GHG) emissions are becoming increasingly important in pavement management decision-making. Pavement repair activities, such as resurfacing, contribute millions of tons of GHG emissions in the United States each year. Optimizing pavement resurfacing activities can lower pavement maintenance's carbon footprint [36]. Additionally, in 2015, the transportation sector accounted for 27 % of overall GHG emissions in the United States, with vehicle fuel consumption accounting for 83 % of total GHG emissions [37]. Hence, environmental impacts based on carbon emissions can further improve the decisions in the real-world scenario, which is the future extension of the expert system presented in this study.

Generally, the road deterioration analysis is a complex task due to the influence of road performance by the interaction of a wide range of traffic, environmental, and construction-related factors. The framework of the BDSS proposed in this study combines the principles of BBN techniques with ML algorithms to develop probabilistic graphical models to determine optimal maintenance and rehabilitation strategies. The methodology developed in this study allows for the delineation of complex correlations between variables and the quantification of risk based on the limiting values accepted in the corresponding location. As a result, the suggested framework can be customized and utilized by highway authorities to their conditions for the sustainable improvement of road networks on a global scale.

According to the framework shown in Fig. 2., the practicality of the BDSS, can be enhanced by making optimized maintenance decisions based on budget availability, the importance of the road network, the severity of the road distress parameters, and the environmental impact of traffic emissions. This illustrates the need for a large volume of data during the initial phase for the suggested system to be effective. As a result, gathering sufficient high-quality historical data on the factors involved in the road deterioration process can be time-consuming in the beginning in order to assure the model's accuracy. Apart from this constraint, BBN principles enable the development of a realistic pavement management solutions, as exemplified in this work.

5. Conclusion

The road deterioration process is a unique process as it differs from one place to another. Furthermore, the complexity associated with the factors leading to road deterioration makes it highly uncertain. Thus, developing an accurate and reliable pavement analysis system is challenging. In addition, the insufficient or inconsistent data related to road performance and road deterioration factors further complicate the road analysis. To overcome these hurdles, this study has adopted Bayesian analysis, a powerful and flexible approach that is capable of handling uncertainty, integrating different data sources, addressing missing values, and even incorporating

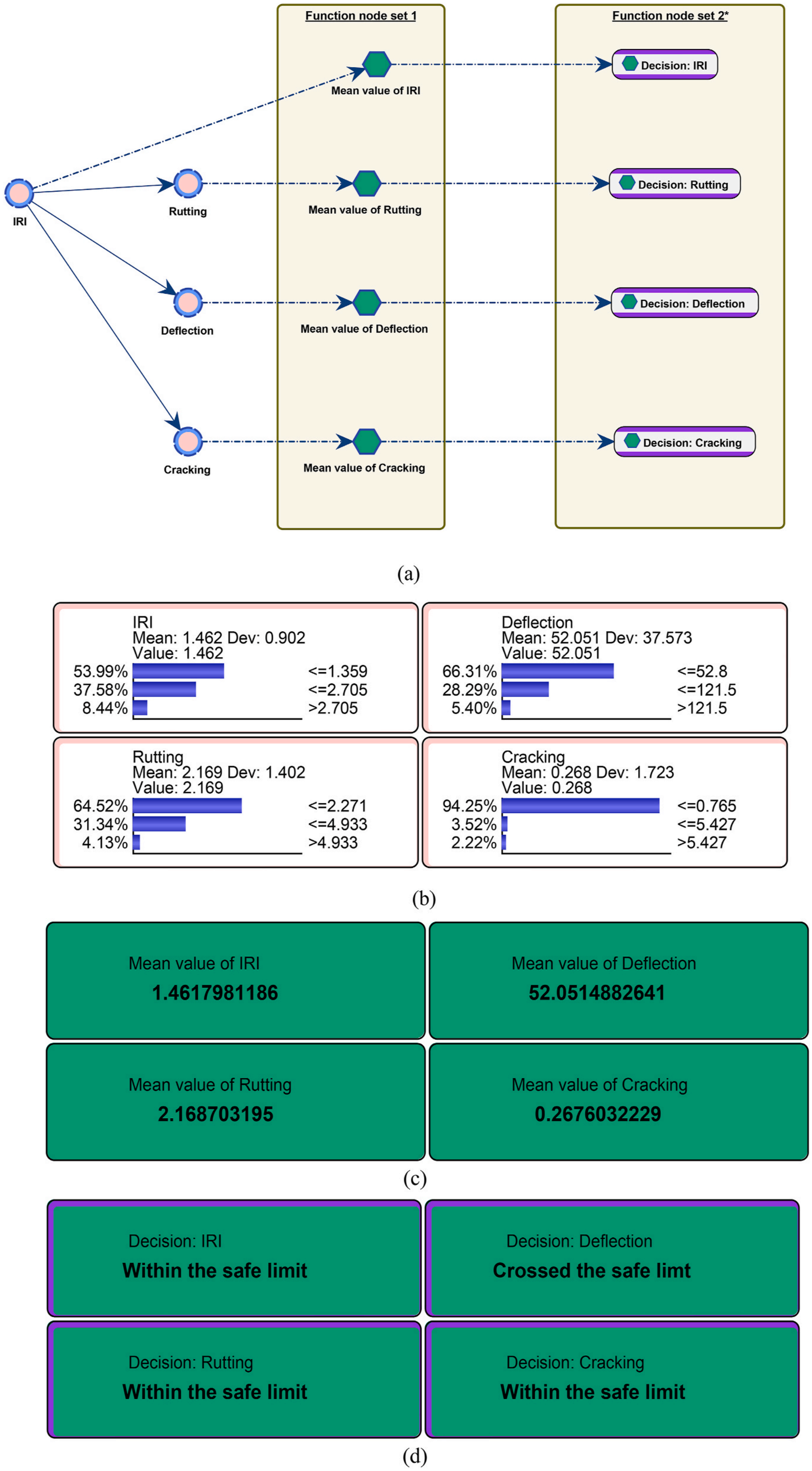

Fig. 7. Example of the decision-making process: (a) Bayesian structure involving function nodes (b) Probability distribution of the road distress parameters (c) Outcome of the function node set 1 (d) Outcome of function node set 2.

expert knowledge.

This study has focused on understanding the complex relationships among the pavement deterioration factors and supporting the decision-making process related to pavement maintenance based on BBN principles. The Bayesian approach paves the way to incorporate prior knowledge such as historical data and update this knowledge when new data become available as in real-time monitoring, improving the accuracy and reliability of the decision-support system. In this study, the road data collected from major road networks in the UAE were analyzed based on the BBN approach to deduce the underlying causal relations among different factors in the form of a probabilistic graphical model. Further, the estimation of the posterior distribution of the factors of interest provides valuable inferences to the pavement maintainers, allowing them to assess the risks associated with different scenarios.

The proposed Bayesian approach has effectively addressed the challenge of reducing uncertainty in road deterioration factors using existing knowledge. However, the approach has certain limitations. The inferences obtained are based on the posterior distribution, which is generated by both the prior distribution (prior knowledge) and likelihood (collected data). Careful selection of the prior distribution is critical, impacting the posterior distribution and subsequent decision-making processes due to mathematical and computational considerations. An additional consideration is that the prior distribution ideally should be a conjugate prior for the likelihood function. In this study, the prior distribution is derived from a large six-year dataset of road data. However, the approach may pose reliability concerns for smaller datasets. Despite the availability of advanced Bayesian Network software packages, ongoing research and development in this field introduces uncertainties, with various concepts still under exploration. Despite these limitations, the BBN approach proves valuable for optimizing road budgets.

Overall, a DSS based on the Bayesian approach can enhance the efficiency and effectiveness of the pavement management process by providing informed decisions to the highway authorities, enabling optimum use of the resources, and improving the service life of the road networks. Future research needs to be conducted to explore further strengths of Bayesian statistics to improve the framework presented in this study. The methodology proposed in this study will aid in reducing the overall cost of the pavement management process by optimizing data collection activities. Moreover, considering additional factors, such as knowing the cost factor of each maintenance activity is extremely important in planning the maintenance programs. In this study, the cost data was unavailable and hence it is not considered. In addition, the impact of carbon emissions from automobiles on the environment, a major concern in the recent decades, needs to be considered in the future research.

Data availability statement

The data used in this study is not publicly available as it is confidential data from the Ministry of Energy and Infrastructure (UAE).

CRediT authorship contribution statement

Babitha Philip: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Hamad AlJassmi:** Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Hamad AlJassmi reports financial support was provided by United Arab Emirates University under the fund code 12R099.

Acknowledgement

The authors are thankful to the Ministry of Energy and Infrastructure in UAE for providing the data related to this research. In addition, the authors appreciate the help of the UAEU library with the language editing services.

References

[1] K.P. George, A.S. Rajagopal, L.K. Lim, Models for predicting pavement deterioration, Transport. Res. Rec. 1215 (1989) 1–7.
[2] A. K. A., A. S. A., and A.-K. I. A., Integrated pavement management system with a markovian prediction model, J. Transport. Eng. 130 (1) (Jan. 2004) 24–33, 10.1061/(ASCE)0733-947X(2004)130:1(24).
[3] T. García-Segura, L. Montalbán-Domingo, D. Llopis-Castelló, M.D. Lepech, M.A. Sanz, E. Pellicer, Incorporating pavement deterioration uncertainty into pavement management optimization, Int. J. Pavement Eng. 23 (6) (May 2022) 2062–2073, https://doi.org/10.1080/10298436.2020.1837827.
[4] C. Han, M. Fang, T. Ma, H. Cao, H. Peng, An intelligent decision-making framework for asphalt pavement maintenance using the clustering-PageRank algorithm, Eng. Optim. 52 (11) (2020) 1829–1847, https://doi.org/10.1080/0305215X.2019.1677636.
[5] M. Hafez, K. Ksaibati, R.A. Atadero, Optimizing expert-based decision-making of pavement maintenance using artificial neural networks with pattern-recognition algorithms, Transport. Res. Rec. 2673 (11) (2019) 90–100, https://doi.org/10.1177/0361198119851085.
[6] S. P. Hadiwardoyo Hamdi, A.G. Correia, P. Pereira, Pavement maintenance optimization strategies for national road network in Indonesia applying genetic algorithm, Procedia Eng. 210 (2017) 253–260, https://doi.org/10.1016/j.proeng.2017.11.074.
[7] M. Ahmed, Kawther, Al-Khateeb, Belal, Mahmood, Application of chaos discrete particle swarm optimization algorithm on pavement maintenance scheduling problem, Cluster Comput. 22 (2019).
[8] L.N.O. Mills, N.O. Attoh-Okine, S. McNeil, Developing pavement performance models for Delaware, Transport. Res. Rec. 2304 (2012) 97–103, https://doi.org/10.3141/2304-11.
[9] S. Madeh Piryonesi, Tamer E. El-Diraby, Data analytics in asset management: cost-effective prediction of the pavement condition index, J. Infrastruct. Syst. 26 (1) (2020).

[10] V. Vyas, A. Pratap Singh, Anshuman, Modeling asphalt pavement condition using artificial neural networks, Mater. Today: Proc. (2022), https://doi.org/10.1016/j.matpr.2022.05.050.
[11] K. Kobayashi, K. Kaito, N. Lethanh, A statistical deterioration forecasting method using hidden Markov model for infrastructure management, Transp. Res. Part B Methodol. 46 (4) (2012) 544–561, https://doi.org/10.1016/j.trb.2011.11.008.
[12] A. Pantuso, G.W. Flintsch, S.W. Katicha, G. Loprencipe, Development of network-level pavement deterioration curves using the linear empirical Bayes approach, Int. J. Pavement Eng. 22 (6) (2021) 780–793, https://doi.org/10.1080/10298436.2019.1646912.
[13] F. Xiao, X. Chen, J. Cheng, S. Yang, Y. Ma, Establishment of probabilistic prediction models for pavement deterioration based on Bayesian neural network, Int. J. Pavement Eng. (May) (2022) 1–16, https://doi.org/10.1080/10298436.2022.2076854.
[14] J. Liu, F. Liu, H. Gong, E.O. Fanijo, L. Wang, Improving asphalt mix design by predicting alligator cracking and longitudinal cracking based on machine learning and dimensionality reduction techniques, Construct. Build. Mater. 354 (2022) 129162, https://doi.org/10.1016/j.conbuildmat.2022.129162.
[15] A. Osorio-Lird, A. Chamorro, C. Videla, S. Tighe, C. Torres-Machi, Application of Markov chains and Monte Carlo simulations for developing pavement performance models for urban network management, Struct. Infrastr. Eng. 14 (9) (2018) 1169–1181, https://doi.org/10.1080/15732479.2017.1402064.
[16] D. Han, K. Kaito, K. Kobayashi, Application of Bayesian estimation method with Markov hazard model to improve deterioration forecasts for infrastructure asset management, KSCE J. Civ. Eng. 18 (7) (2014) 2107–2119, https://doi.org/10.1007/s12205-012-0070-6.
[17] E.S. Park, R.E. Smith, T.J. Freeman, C.H. Spiegelman, A Bayesian approach for improved pavement performance prediction, J. Appl. Stat. 35 (11) (2008) 1219–1238, https://doi.org/10.1080/02664760802318651.
[18] B. Philip, H.A. Jassmi, A bayesian approach towards modelling the interrelationships of pavement deterioration factors, Buildings 12 (7) (2022), https://doi.org/10.3390/buildings12071039.
[19] M. Sharafeldin, O. Albatayneh, A. Farid, K. Ksaibati, A bayesian approach to examine the impact of pavement friction on intersection safety, Sustainability 14 (19) (2022), https://doi.org/10.3390/su141912495.
[20] H.-C. Dan, C. Ling, Z. Zhu, L. Gao, X. Zeng, A data fusion approach for estimating traffic distribution characteristics of expressway: a case study of guangdong province, China, Front. Mater. 9 (2022), https://doi.org/10.3389/fmats.2022.939579.
[21] B.-H. Woo, D.-H. Yoo, I.-K. Jeon, J.-S. Kim, K.-S. Yoo, H.G. Kim, Estimation of the thermal conductivity of cement composites using bayesian statistical approach, Compos. B Eng. 243 (2022), https://doi.org/10.1016/j.compositesb.2022.110073.
[22] H. Liang, et al., Investigating long-term performance of flexible pavement using Bayesian multilevel models, Road Mater. Pavement Des. (2022), https://doi.org/10.1080/14680629.2022.2117062.
[23] T. Blumenfeld, et al., A methodology for probabilistic pavement condition forecast based on Bayesian filters, Struct. Infrastruct. Eng. 0 (0) (2022) 1–14, https://doi.org/10.1080/15732479.2022.2077769.
[24] A. Heba, G.J. Assaf, Bayesian linear regression performance model for the libyan national road network without an initial database on its condition based on expert's knowledge, in: Joint Annual Conference of Canadian Transportation Research Forum and US Transportation Research Forum: North American Transport Challenges in an Era of Change, CTRF-TRF, 2016, p. 2020.
[25] A. Wolters, K. Zimmerman, K. Schattler, A. Rietgraf, Implementing Pavement Management Systems for Local Agencies, 2011.
[26] G. Kabir, N.B.C. Balek, S. Tesfamariam, Consequence-based framework for buried infrastructure systems: a Bayesian belief network model, Reliab. Eng. Syst. Saf. 180 (August) (2018) 290–301, https://doi.org/10.1016/j.ress.2018.07.037.
[27] N.O. Attoh-Okine, Probabilistic analysis of factors affecting highway construction costs: a belief network approach, Can. J. Civ. Eng. 29 (3) (2002) 369–374, https://doi.org/10.1139/l02-003.
[28] O. Starkova, A.I. Gagani, C.W. Karl, I.B.C.M. Rocha, J. Burlakovs, A.E. Krauklis, Modelling of environmental ageing of polymers and polymer composites—durability prediction methods, Polymers 14 (5) (2022), https://doi.org/10.3390/polym14050907.
[29] J. Pearl, Department of Statistics Papers, in: UCLA Department of Statistics Papers, 1994, p. 36.
[30] W.M. Bolstad, J.M. Curran, Introduction to Bayesian Statistics, John Wiley & Sons, 2016.
[31] N.J. Garber, L.A. Hoel, Traffic and Highway Engineering, Cengage Learning, 2018.
[32] R. Van de Schoot, S. Depaoli, Bayesian analyses: where to start and what to report, Eur. Health Psychol. 16 (2) (2014) 75–84.
[33] X. Neapolitan, R. Jiang, Decision analysis fundamentals, in: Probabilistic Methods for Financial and Marketing Informatics, 2007, pp. 177–228, https://doi.org/10.1016/b978-012370477-1/50022-0.
[34] B. S.A.S., "BayesiaLab," Available online: https://www.bayesialab.com/..
[35] N. Veyrat-Charvillon, F.-X. Standaert, Mutual information analysis: how, when and why?, in: Cryptographic Hardware and Embedded Systems-CHES 2009: 11th International Workshop Lausanne Springer, Switzerland, 2009, pp. 429–443. September 6-9, *2009 Proceedings*.
[36] A. Ogwang, S. Madanat, A. Horvath, Optimal cracking threshold resurfacing policies in asphalt pavement management to minimize costs and emissions, J. Infrastruct. Syst. 25 (2) (2019) 4019003.
[37] H. Wang, I. Al-Saadi, P. Lu, A. Jasim, Quantifying greenhouse gas emission of asphalt pavement preservation at construction and use stages using life-cycle assessment, Int. J. Sustain. Transp. 14 (1) (2020) 25–34.