buildings

MDPI

Article

A Bayesian Approach towards Modelling the Interrelationships of Pavement Deterioration Factors

Babitha Philip [1,2] and Hamad Al Jassmi [1,2,*]

[1] Department of Civil and Environmental Engineering, UAE University, Al Ain P.O. Box 15551, United Arab Emirates; 201890119@uaeu.ac.ae
[2] Emirates Center for Mobility Research (ECMR), UAE University, Al Ain P.O. Box 15551, United Arab Emirates
* Correspondence: h.aljasmi@uaeu.ac.ae

Abstract: In this study, Bayesian Belief Networks (BBN) are proposed to model the relationships between factors contributing to pavement deterioration, where their values are probabilistically estimated based on their interdependencies. Such probabilistic inferences are deemed to provide a reasonable alternative over costly data collection campaigns and assist in road condition diagnoses and assessment efforts in cases where data are only partially available. The BBN models examined in this study are based on a vast database of pavement deterioration factors including road distress data, namely cracking, deflection, the International Roughness Index (IRI) and rutting, from major road sections in the United Arab Emirates (UAE) along with the corresponding traffic and climatic factors. The dataset for the analysis consisted of 3272 road sections, each of 10 m length. The test results showed that the most critical parameter representing the whole process of road deterioration is the IRI with the highest nodal force. Additionally, IRI is strongly correlated with rutting and deflection, with mutual information of 0.147 and 0.143, respectively. Furthermore, a Bayesian network structure with a contingency table fit of over 90% illustrates how the road distress parameters change in the presence of external factors, such as traffic and climatic conditions.

Keywords: road distress parameters; correlation analysis; Bayesian belief networks; uncertainty

check for updates

Citation: Philip, B.; Jassmi, H.A. A Bayesian Approach towards Modelling the Interrelationships of Pavement Deterioration Factors. *Buildings* **2022**, *12*, 1039. https://doi.org/10.3390/buildings12071039

Academic Editors: Sara Shirowzhan, Brian Lee and Cynthia Changxin Wang

Received: 13 June 2022
Accepted: 15 July 2022
Published: 18 July 2022

1. Introduction

The ongoing challenge faced by all levels of highway administration is the inadequate funding for highway projects [1]. Hence, collecting road data is an expensive process in terms of cost, time, and effort. Storing and analysing the excessive amount of data collected with a variety of available technologies is a serious problem for road authorities. In addition, the inclusion of irrelevant data in the database reduces the efficiency of road management systems in providing optimal and cost-effective pavement solutions. Identifying the factors that significantly contribute to the deterioration of road conditions would limit the amount and frequency of data collected, thus providing the right amount of data for pavement performance analysis [2]. In addition, studying the correlations between different factors helps in understanding the importance of one factor over another.

Generally, a correlation analysis is carried out to understand the various relationships between factors that represent pavement conditions for efficient pavement management. Since many factors influence pavement condition, two important questions arise: which factors help to represent the overall pavement condition, and how can these factors be modelled efficiently for pavement management? The condition of the roads are analysed in the literature based on the intensity of various road distress factors such as the International Roughness Index (IRI), cracking, rutting, bleeding, deflection and many others [3]. However, it is not yet clear which factor provides the highest information gain, which could be used as an indicator of the general distress condition of the road as well as a predictor for probabilistically estimating the rest of the distress factors. This is indispensable for optimizing road assessment data collection efforts and finding a compromise when it is not possible to obtain data about all factors but only some of them.

These road distress factors are influenced by various explanatory variables, including traffic conditions, environmental factors, design features, and similar [4]. The main factors affecting pavement condition are shown in Figure 1. Due to the distinct nature of the these factors, the pavement conditions on different road sections are heterogeneous, adding uncertainty in assessing road performance [5].

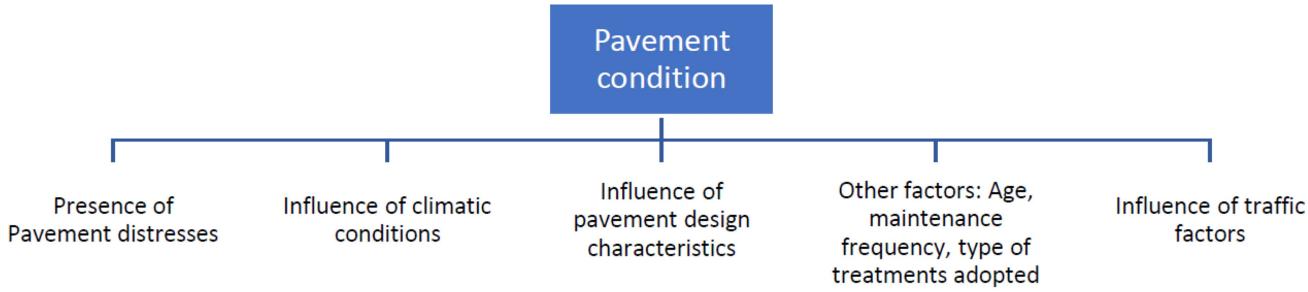

Figure 1. Major factors influencing pavement condition.

Researchers have employed different methods to understand the relationships among road deterioration factors. The most prominent ones are back-propagation neural network [6], random forests regression [7], statistical regression analysis [8], structural equation modelling [9], linear regression and artificial neural networks [10]. The majority of the studies select any one distress parameter such as cracks [11], IRI [6] etc., as an indicator of pavement deterioration and subsequently explore the dependency of this distress parameter with other deterioration factors. This approach may not be appropriate because in certain situations, the selected distress parameter is not a surrogate measure of pavement condition [8,9]. Thus, it is important to identify the central factor which is capable of representing the overall pavement deterioration. In addition, the existing studies do not capture the uncertainty associated with distress parameters [6] and even the uncertainty related to the explanatory variables [7,10].

Bayesian belief networks (BBN) are capable of representing uncertain knowledge about interrelationships existing between variables in a complex system [12]. The BBN structure serves as an inference mechanism which aids in probabilistically estimating the unknown values [13]. In low-data scenarios, BBN allows one to perform inverse modelling without the problem of overfitting to obtain insights on unobserved variables [14]. As far as the authors are aware, there is no study that focuses on the application of BBN to model the correlations between road distress parameters and which factors provide the most information about road condition. Knowing which parameter best reflects road condition is beneficial when prioritizing data collection and maintenance activities in the face of shrinking road budgets. The present work aims to fill this gap by developing different BBN models that incorporate road distress parameters, traffic factors, environmental factors and specific road factors to analyse the correlations between the factors for efficient pavement management.

The manuscript continues as follows: Section 2 mentions several studies in which BBN was used in pavement research. Section 3 explains the methodology of the study and describes the data and the steps followed in the analysis. Section 4 presents the results of the analysis and their implications for developing better management solutions for pavements. In Section 5, we discuss the results in detail and mention certain limitations of the proposed approach and possible future directions. Finally, Section 6 presents the conclusions of the study.

2. Literature Review: Applications of Bayesian Belief Networks in Pavement Studies

Applications of Bayesian networks range from predicting a disease/treatment for a patient [15] to performing profit analysis for a company [16] and to creating genetic maps [17]. The most important properties of the Bayesian network are its ability to provide real-time solutions and its ability to handle missing data under uncertainty [18]. BBN has been used extensively in various areas of pavement management due to its advantages in accounting for uncertainty, capturing unobserved heterogeneity and the like. Therefore, BBN is used in this study to perform a correlation analysis of the factors that cause pavement deterioration.

Mohamed and Tran [19] reported the ability of BBNs to draw causal relationships and their importance in making accurate decisions by investigating the causal relationships between 76 variables grouped under 12 quality assurance inspection activities and Portland cement concrete pavement quality (PCCP). The results show that quality deterioration decreased by 9.7% when the risk of an inspection activity related to moisture in PCCP aggregates was changed from high to low and vice versa. Another study by the same authors found a causal relationship between inspection activities and hot asphalt pavement quality [20].

Ismail et al. [5] proposed a knowledge-based BBN model for prioritizing road sections for efficient management based on several key performance indicators (KPIs) categorized into pavement condition, road safety, environmental impact and capacity. The causal relationships between the KPIs were established based on expert knowledge and previous studies. This shows that BBN can take into account expert judgements (prior knowledge), which is crucial in the evaluation of road transport systems.

Attoh-Okine [13] investigated the ability of BBN to deal with incomplete and insufficient data related to highway construction costs by developing random and associative relationships among cost variables. The probabilistic relationship and the corresponding information flow between nodes are represented by grouping the variables into environmental costs, road design costs, directed labour costs, other labour costs and material costs. Once the variables are identified, logical and dependence relationships are established among the variables. Based on the relationships obtained among the variables, the information flows among the network to generate inferences, both predictive inferences and posterior computations based on available information and historical data. The main advantages of the literature-based BBN method are shown in Figure 2.

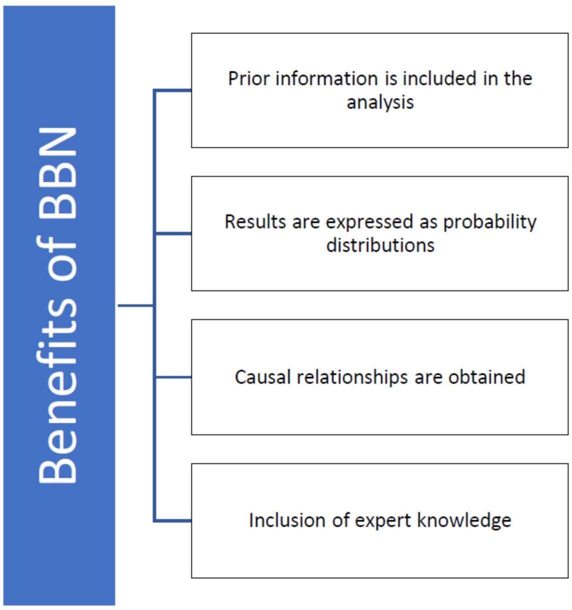

Figure 2. Major advantages of BBN for pavement assessment studies.

Better data collection would be possible if the most important factor providing information on the condition of the road was known in advance. The development of a BBN model to represent the various relationships between pavement deterioration factors and the importance of each relationship in the pavement deterioration process would improve pavement management. The objective of this study is to investigate the use of BBN to derive correlations between road distress factors and external factors. The database for the analysis was prepared by collecting road data from 32 road networks in the UAE from 2013 to 2019. The original data was collected from the Ministry of Energy and Infrastructure (MoEI). The major objectives of the study are:

- Estimate the correlations between the road distress parameters: cracking, deflection, IRI and rutting;
- Investigate the influence of external factors related to traffic, environment, and road characteristics on pavement conditions;
- Perform a sensitivity analysis of road distress parameters to understand what type of road distress is prominent on each type of road (arterial, collector, freeway and expressway).

This paper is a first step towards the development of a pavement management system (PMS) capable of making pavement decisions based on a probabilistically estimated right amount of quality data.

3. Methodology

This study was carried out based on road data collected from 32 road networks in the UAE from 2013 to 2019. The major steps followed in this study are given in Figure 3. Details of the data and the method of analysis are presented in the following subsections.

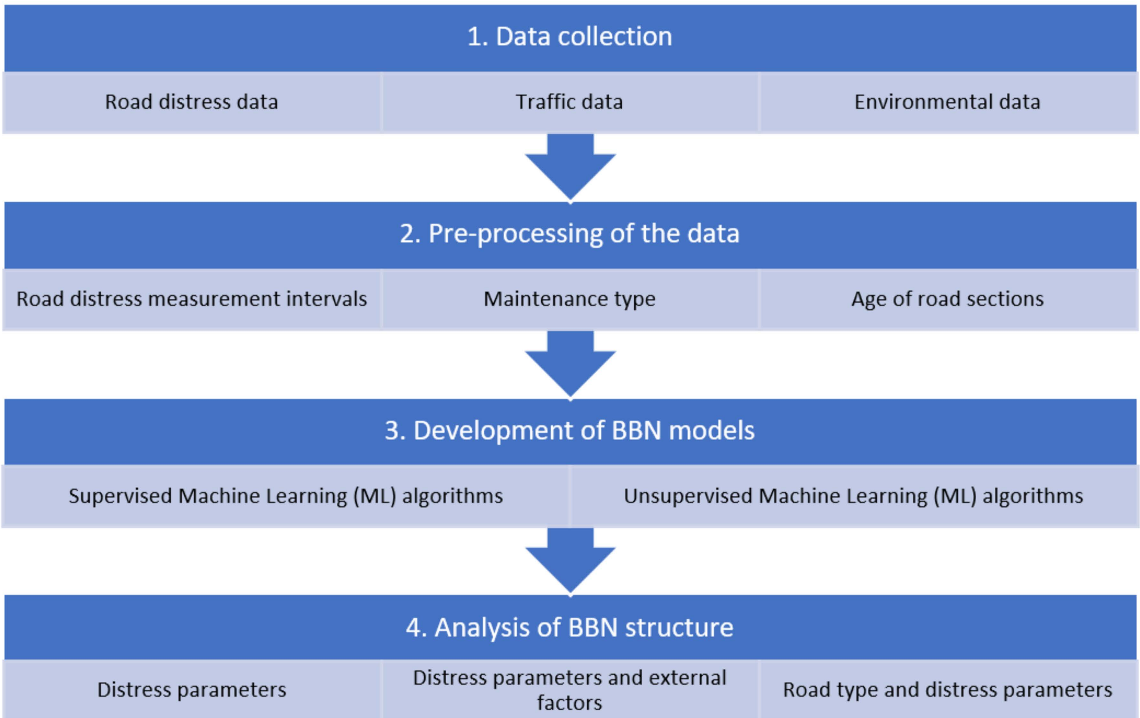

Figure 3. Methodology of the study.

3.1. Data Collection

Road data along 32 road segments in the northern region of the country from 2013 to 2019 managed by the Road Department of UAE Ministry of Energy and Infrastructure (MoEI) were selected for the analysis. Figure 4 provides the satellite view of the studied region where red lines are the roads that are part of the MoEI network. Other roads which are not part of the MoEI network are shown in yellow lines. Thus, data from roads represented by red lines were collected for this study. The dataset considered in this study can be broadly divided into road data, traffic data and environmental data.

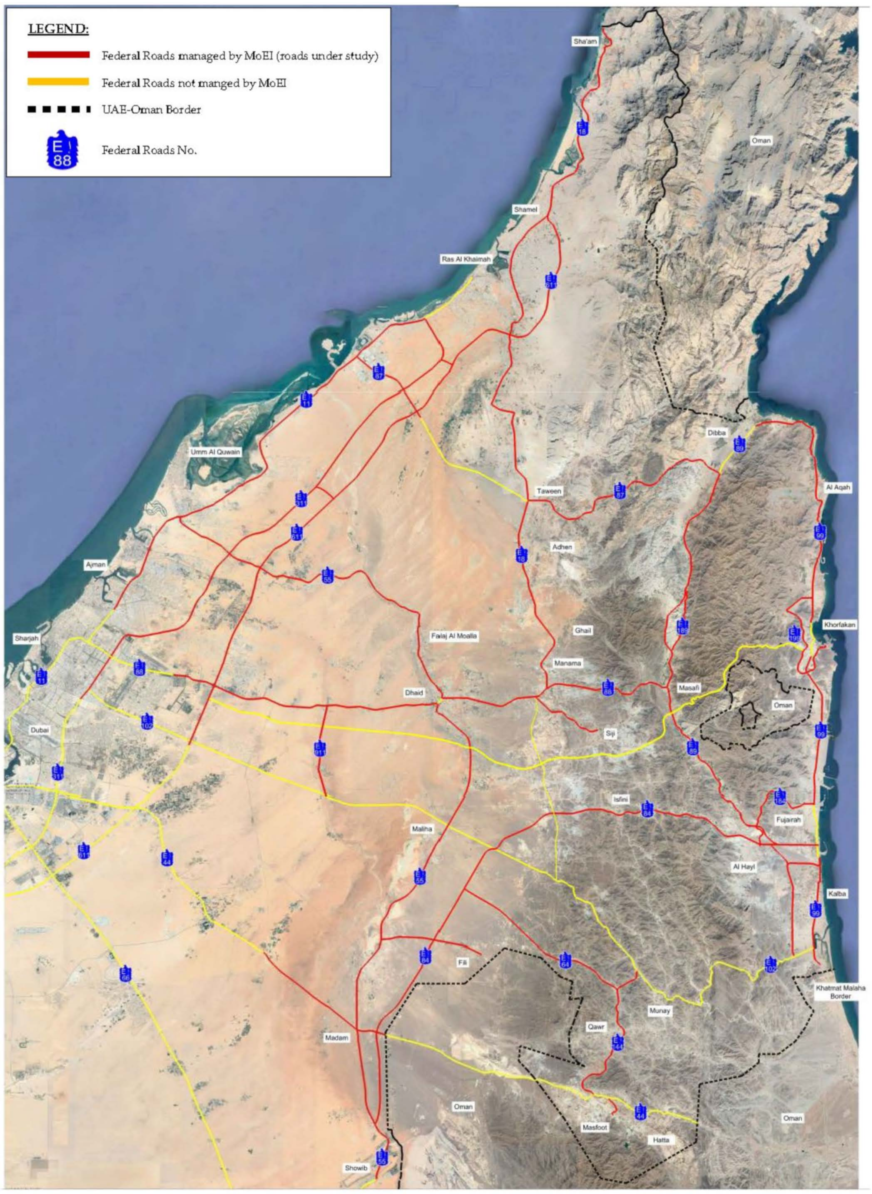

Figure 4. Selected road networks (red lines) for study.

3.1.1. Road Data

The dataset contained values for cracking, deflection, IRI and rutting. The following devices were used by MoEI to measure these road distress parameters.

- Cracking-Laser Crack Measurement System (LCMS);
- Deflection-Falling Weight Deflectometer (FWD);
- IRI-Laser profilometer;
- Rutting-Laser Rutting Measurement System.

Table 1 shows that these values were not observed uniformly along the entire length of the road. For example, cracking was measured every 10 m and deflection every 100 m (approximately). Although cracking, IRI and rutting were measured every 10 m, the sections considered for measurement were different for these parameters.

Table 1. Details of distress parameters.

Distress Factor	Unit	Measuring Interval in Meters	Minimum Value	Maximum Value
Cracking	Percentage (%)	10	0.00	99.87
Deflection	mm/100	100	0.00	374.00
IRI	m/km	10	0.00	52.99
Rutting	Mm	10	0.00	52.65

3.1.2. Traffic Data

Traffic data for the selected road networks were collected from the Road Department. The components of the traffic data were light and heavy vehicle traffic counts and the direction of traffic flow.

3.1.3. Environment Data

The environmental conditions of each road corresponding to its location and date of measurement were collected from weather records available online. The components of the environmental data include temperature (°C), humidity (%) and air pressure (mbar).

3.2. Pre-Processing of the Data

3.2.1. Road Distress Measurement Intervals

The measurement intervals of the individual parameters were different, as shown in Table 1. Therefore, the first task in the pre-processing was to apply different spreadsheet formulas to obtain uniform measurement intervals and intercepts for the parameters. The distress values corresponding to the starting point of the road network to the endpoint were listed in a continuous format of 10 m. Consequently, the road networks in the study were divided into a total of 132,999 road sections, each of 10 m length. An illustration of the pre-processing procedure can be found in Table 2.

Table 2. An example of a pre-processing procedure.

Before Pre-Processing		After Pre-Processing	
Initial Distance	**Final Distance**	**Initial Distance**	**Final Distance**
651	672	660	670

3.2.2. Assumption of Maintenance Treatment Types

No information was available on the frequency of maintenance carried out. However, the type of maintenance was known from the discussions with the officials at the Road Department. In the road survey, the type of maintenance is divided into major maintenance, surface maintenance and partial maintenance. If all the distress values for a road section are corrected, it means that the section in question has undergone major treatment. In the case of surface treatment, maintenance measures for cracking, IRI and rutting are carried out and the deflection of the section is not maintained. Partial treatment, however, involves correction of IRI, ruts and shallow cracks, while deep cracks and deflection remain. The maintenance types are shown in Figure 5.

The type of maintenance (major, surface and partial) was decided for each road section by comparing the values of distress parameters in the present year with the values in the previous year. Subsequently, in addition to the road sections which have undergone the above-mentioned treatment types, there were road sections which were not maintained at all and road sections with maintenance status as “unknown”, due to non-availability of distress data in the consecutive years. Hence, filtering out the road sections which were not maintained and with “unknown” status for maintenance resulted in a dataset of 3272 road sections.

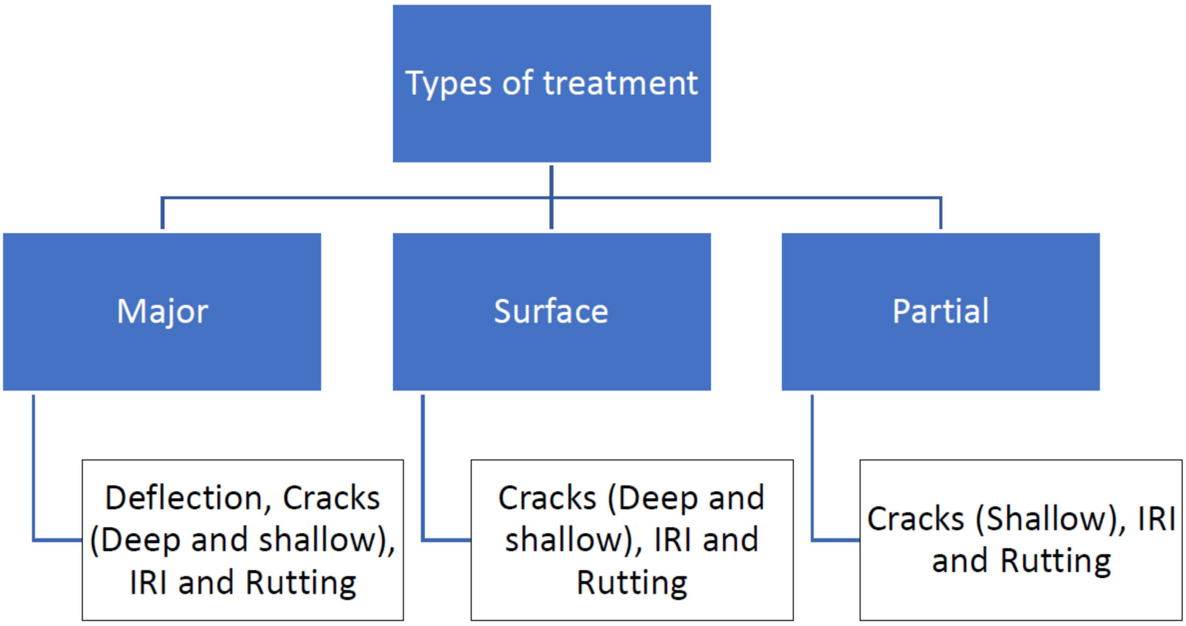

Figure 5. Types of maintenance treatment.

3.2.3. Assumption of the Age of the Road Section

The Road Department indicated the year in which the roads were paved. Therefore, the age could be easily calculated based on the year of construction. However, it is known that the life span increases when a road section is maintained [21]. The Road Department in UAE adopts several initiatives to improve the quality of infrastructure, and the road networks generally undergo major treatments to improve the road conditions and to attain well-maintained roads. Due to this fact, this study assumes a maintained road section equivalent to a newly constructed road. Therefore, the age of the road was calculated from the time of maintenance. In addition, the road name, road type and location were also recorded.

3.2.4. Prepared Data for Analysis

After filtering the data, the final dataset for analysis comprised 3272 data points. Each data point represents a road segment of 10 m in length. A data point from the database prepared for analysis is shown in Table 3.

Table 3. An example of a data point in the dataset.

Road Data						Distress Parameters				Environment Factors			Traffic Factors		
Road	Road Type	Initial Distance	Final Distance	Maintenance Type	Age	Cracking	Deflection	IRI	Rutting	Temperature	Humidity	Atm. Pressure	Traffic Count (Light Vehicle)	Traffic Count (Heavy Vehicle)	Traffic Direction
E11	Arterial	8540	8550	Partia	1	0	51	1.016	1.75375	26	49	1019	4,203,429	487,421	Forward

3.3. Development of BBN Model

The proposed methodology is primarily based on the principles of BBN and algorithms from artificial intelligence (AI). In this research, Bayesian network structures are modelled by machine learning of the data, making probabilistic inferences considering the collected evidence, which reduces the associated uncertainty. It is used to gain a deep understanding of complex problems and to draw conclusions and to predict the consequences of different scenarios.

Bayesian networks are directed acyclic graphs. Events are represented as nodes, a series of arrows connect the nodes, and conditional probability distribution tables represent the probabilistic relationships between the nodes. Bayesian networks have various applications in modelling uncertainty with causal and probabilistic relationships. Uncertainty arises primarily from incomplete observations, noisy information, and incomplete knowledge, which are common in real-life data collection. Bayesian networks are based on Bayes' theorem, which is represented by Equation (1).

$$P(H|D) = \frac{P(H)P(D|H)}{P(D)} \tag{1}$$

where $P(H \mid D)$ is the conditional probability of the event H given the event D, and P(D) and P(H) are the probabilities of events D and H correspondingly [22].

In Bayesian statistics, the uncertainty associated with a parameter is first quantified by a probability distribution known as the prior distribution $P(\theta)$, which represents the total prior knowledge. Further, the newly collected evidence is represented by the likelihood $P(Y \mid \theta)$. According to Bayes' theorem, the conditional distribution $P(\theta \mid Y)$ is calculated as in Equation (2).

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{P(Y)} \tag{2}$$

where P(Y) is the marginal distribution of Y and is a normalizing constant. Thus,

$$P(\theta|Y) \infty\, P(\theta)P(Y|\theta) \tag{3}$$

This is called posterior distribution. It summarizes the knowledge θ we have after observing data Y. Thus, the posterior distribution is proportional to the likelihood and the prior distribution. Bayesian causal inferences are made based on the posterior distribution, which contains all possible information concerning θ [23]. Compared to a frequentist approach, the Bayesian approach can not only capture prior knowledge in addition to the available data to make decisions, but can also deal with missing data, datasets with outliers, and non-linear relationships, as well as present probabilistic results in a visual way that is easy to interpret [24].

Missing values are inevitable in real-world data collection, especially when continuously recording the values of parameters. Missing values can be divided into three categories: missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR). The approaches to deal with these data can in turn be divided into inference restricted to complete data, imputation-based approaches and likelihood-based approaches [25]. Bayesian networks can handle the different approaches mentioned, and the appropriate approach is chosen depending on the type of data. Categorical, continuous, and discrete variables can be included, with continuous variables being discretized using the appropriate discretization methods. Discretization aims to find a set of thresholds that can be used to divide the data into finite intervals. BBN adopts several univariate, bivariate and multivariate algorithms for discretizing continuous variables [26].

3.3.1. Supervised and Unsupervised Bayesian Learning

The most important ML algorithms in Bayesian networks are supervised learning and unsupervised learning algorithms. Supervised ML algorithms are used when a "target node" (final output) is available. It is highly recommended to analyse the impact of each factor involved in the study on the target. Unsupervised ML algorithms, however, are used when there is no "target node", and it is recommended to analyse all direct probabilistic relationships between factors [27]. Different Bayesian networks obtained against different ML algorithms are further analysed to select the best Bayesian network structure for making probabilistic inferences.

The minimum description length (MDL) score is used to select the optimal Bayesian network model. The MDL score is a combination of the complexity of the structure and the information obtained from the structure. According to the principles of the MDL score, any dataset can be learned with a model determined by the degree to which the model can compress the data. Here, the model and the data are considered as codes, and the code length represents the generalization ability of the model. The model with the shortest MDL score is the best [28]. The MDL score is given by Equation (4).

$$\mathrm{MDL} = \mathrm{L}(\mathrm{Data}|\mathrm{Model}) + \mathrm{L}\ (\mathrm{Model}) \tag{4}$$

where, L(Data | Model) is the description length of the data given the model, and L(Model) is the description length of the model. The framework proposed in this study for developing a BBN model to estimate the correlation between pavement deterioration factors is presented below.

3.3.2. Proposed BBN Framework

The framework presented in this study evaluates the use of BBN techniques in modelling road data to help highway authorities make wise decisions. The design of the framework was carried out using BayesiaLab 10.2, a software tool based on BBN that is particularly suitable for realizing causal relationships between the variables under study [29]. The framework mainly comprises three phases: the data cleaning and pre-processing phase, the Bayesian model building phase and the probabilistic inference phase. The first phase is the most important, as the whole analysis is carried out on the processed data obtained from this phase. The collected data should undergo appropriate pre-processing before being fed into the ML tool. In the data cleaning and pre-processing phase, the variable type, the techniques for processing missing values and the discretization method are determined.

First, the "type" of data must be mentioned. Most of the variables used in this study are continuous, except for road type, maintenance type and direction of traffic flow. The continuous variables are further discretized using different algorithms (univariate, bivariate and multivariate) and intervals. The different discretization algorithms considered in this study are Tree, Perturbed Tree, Supervised Multivariate, R2-GenOpt*, R2-GenOpt, K-Means, Density Approximation, Normalized Equal Distance, Equal Distance, Equal Frequency and Unsupervised Multivariate. The missing data, either MCAR, MAR or MNAR, are treated either with the filter method, static imputation, or dynamic imputation. The framework used in this study is shown in Figure 6.

The second stage involves Bayesian analysis, where ML algorithms are applied to the imported data based on the requirements of the decision makers to build the Bayesian model. Then, supervised and unsupervised ML algorithms are applied to define the causal relationships between road distress parameters, traffic factors and environmental conditions. The types of supervised algorithms considered in this study include Naïve Bayes, Augmented Naïve Bayes, Tree Augmented Naïve Bayes, Sons & Spouses, Markov Blanket, Tree Augmented Markov Blanket, and Minimal Augmented Markov Blanket. The unsupervised learning algorithms used to determine the probabilistic relationships between the variables under study are Maximum spanning Tree, Taboo, EQ, SopLEQ and Taboo order. In this stage, several optimization attempts are made to obtain the best possible network structure, which is determined based on the MDL score.

Finally, the last stage contributes to the probabilistic inference phase, where the results obtained from the analysis are used to generate a Bayesian inference for optimal decision making. The optimal model thus developed is used to test different scenarios related to road management. Depending on the framework conditions, the Bayesian inferences obtained are then used to prioritize maintenance measures. In this way, a tailor-made model is available based on the road conditions specified by the operator. The term "tailor-made" is used here to indicate that the thresholds accepted by countries and/or authorities may be different, which influences the decision on determining the critical road sections.

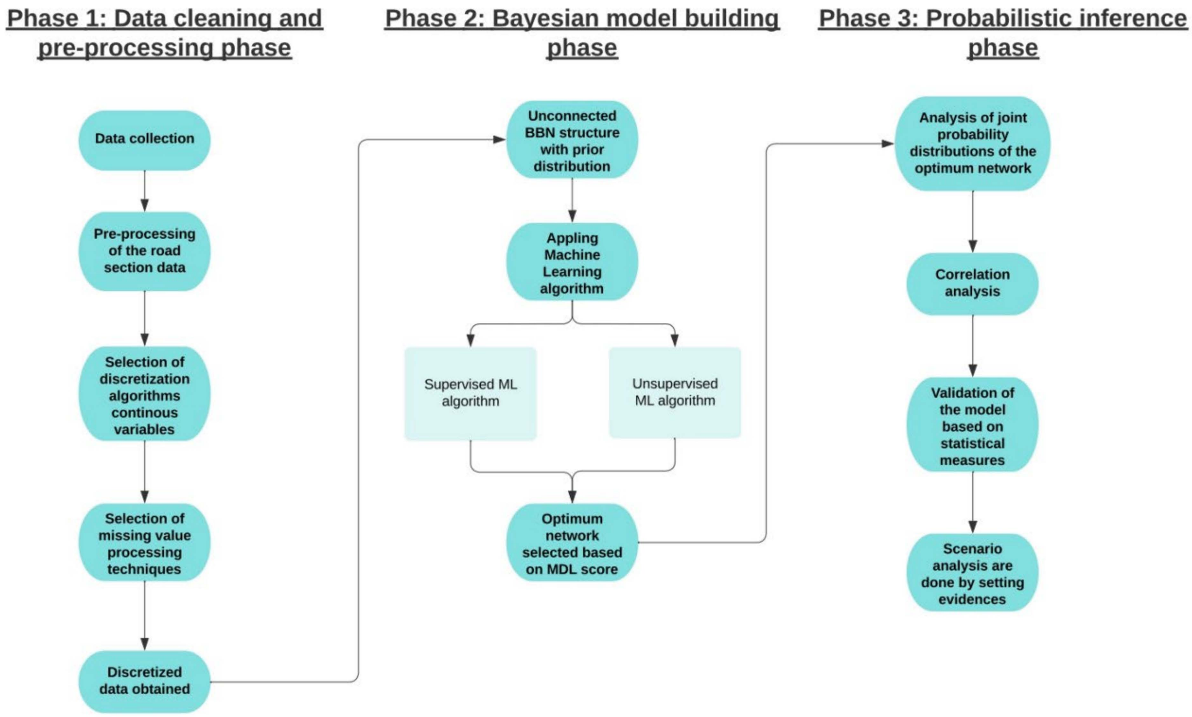

Figure 6. The framework of the proposed BBN model.

3.4. Analysis of BBN Structure

3.4.1. Bayesian Correlation Analysis

The BBN model represents causal relationships between nodes. This is primarily estimated by the arc force, which measures the strength of the relationships between the nodes. The Kullback–Leibler divergence (relative entropy) of the two probability distributions is measured to estimate the similarity between the two probability distribution functions. Let p and q be the joint probability distributions of a discrete random variable x, and their Kullback–Leibler divergence (DKL) is shown in Equation (5) [30]. Furthermore, mutual information, which is a measure of the dependence or information between two variables, is closely related to the entropy of a variable. The mutual information I(x,y) between x and y corresponds to the Kullback–Leibler divergence as shown in Equation (6). The estimate of the Pearson correlation coefficient indicates the strength and correlation between two variables. The Pearson correlation coefficient plays an important role in the statistical analysis of the relationship between two variables and ranges from −1 to 1. The equation of the Pearson correlation coefficient r for a sample size of n is given in Equation (7).

$$\mathrm{DKL}\ (\mathrm{p}\ ||\ \mathrm{q}) = \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{q(x)} dx \quad (5)$$

$$\mathrm{I}(\mathrm{x},\mathrm{y}) = \mathrm{DKL}\ (\mathrm{p}(\mathrm{x},\mathrm{y})\ ||\ \mathrm{p}(\mathrm{x})\mathrm{p}(\mathrm{y})) \quad (6)$$

$$\mathrm{r} = \frac{\mathrm{n}(\sum \mathrm{xy}) - (\sum \mathrm{x})(\sum \mathrm{y})}{\sqrt{[\mathrm{n}\sum x^2 - (\sum x)^2][\mathrm{n}\sum y^2 - (\sum y)^2]}} \quad (7)$$

3.4.2. Optimum Model Selection

The final Bayesian structure is confirmed by measuring the "entropy" of the model. Entropy (H) is the measure of the "uncertainty" inherent in the possible states of the event. For example, if the possible outcomes of an event X are $x_1, x_2, \ldots, x_n$ and the corresponding probabilities for the occurrence of each outcome are $P(x_1), P(x_2), \ldots, P(x_n)$, then the entropy is as in Equation (8). The value of entropy depends on the number of possible states of the event [31]. The maximum value of entropy increases logarithmically with the number of states of the event. The maximum entropy is expressed in Equation (9). In this study, a BBN model is developed with nodes (road deterioration factors) that have different numbers of states. Therefore, a normalized measure of entropy is used to represent the data. Normalized entropy (HN) represents a normalized measure of the uncertainty of an event, independent of the number of states associated with the event [32]. It is the entropy ratio to the maximum entropy as shown in Equation (10).

$$\mathrm{H(X)} = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i) \tag{8}$$

$$\mathrm{Hmax(X)} = \log_2(n) \tag{9}$$

$$\mathrm{HN\ (X)} = \frac{H(X)}{Hmax(X)} = -\sum_{i=1}^{n} \frac{P(x_i)\log_2 P(x_i)}{\log_2(n)} \tag{10}$$

The reliability of the model is compared using the values of entropy and other statistical measures such as the contingency table fit and the Person's correlation coefficient. The contingency table represents the frequency distribution of the variables in a matrix format. Contingency tables are widely used in various research fields to represent the correlation between variables. Contingency is measured by various statistical tests, including the G-test, Pearson's Chi-square test, etc. The value of contingency implies the dependence between variables [33].

4. Results

This section presents the results of the correlation analysis, interprets the results, and draws Bayesian conclusions.

4.1. Correlation Analysis among Road Distress Parameters

By estimating the relationships between the road distress parameters, it may be possible to prioritize areas that need urgent attention and optimize the effort of data collection. This approach will reduce the huge sums of money highway authorities spend on monitoring roads. If you are certain that a specific parameter may occur in a road section, you no longer need to monitor that parameter. Furthermore, the relationships between road distress parameters can be used to estimate the value of another related parameter without actual measurement.

Learning algorithm: The nodes representing the road distress parameters are included in this analysis. Unsupervised learning algorithms are used here, as each parameter is given the same importance at the beginning. Different unsupervised learning algorithms are applied until a network structure with the best performance is reached. After several learning cycles, the lowest MDL score was obtained for the algorithms Unsupervised Learning-EQ, Discretization-R2 GenOpt* for interval 3 and Missing Value Processing-Dynamic Imputation. The Bayesian structure is shown in Figure 7. To illustrate the choice of the optimal model based on the MDL, Table 4 lists the MDL scores for different ML algorithms under the R2 GenOpt* discretization method for discretization intervals 3 and 4 with Missing Value Processing-Dynamic Imputation. As mentioned in Section 3.3.2, different discretization options are also tested to obtain the final model. The algorithm with the lowest MDL value is highlighted.

This Bayesian model presented here shows that about 94% of the road sections have a crack value close to 0, which means that the roads in the UAE are well maintained or that the road sections considered in this study have been recently inspected for cracks.

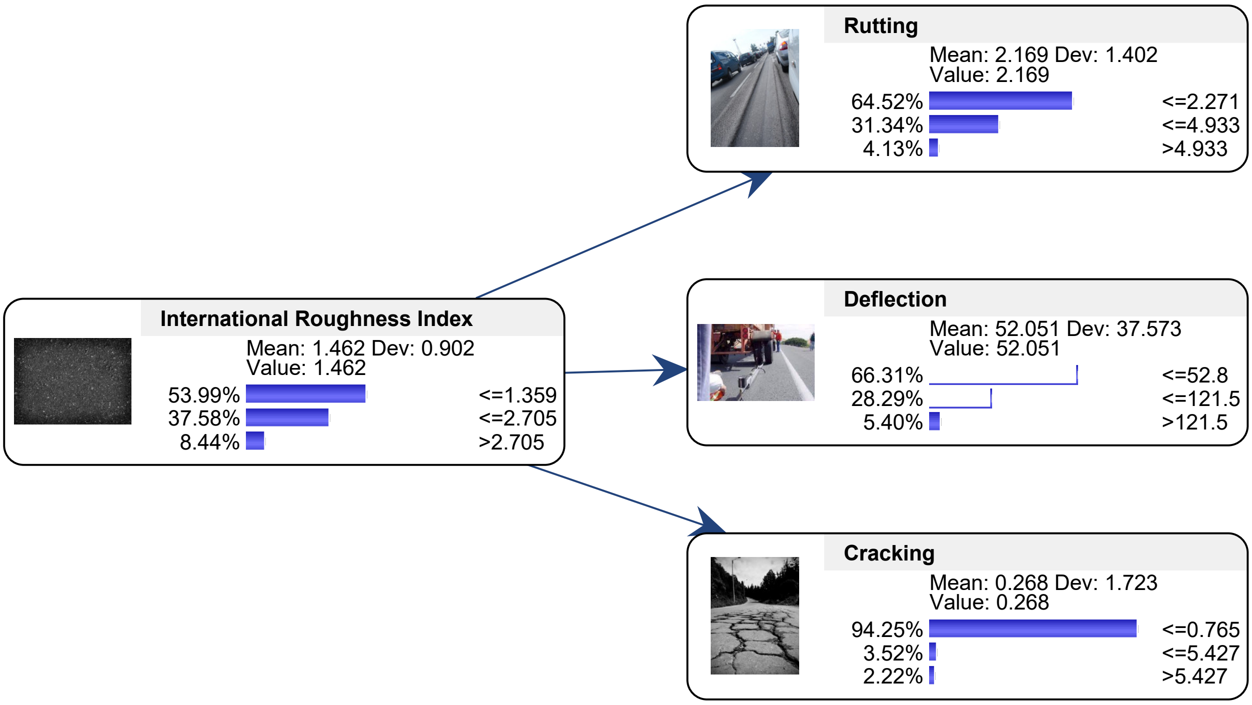

Figure 7. Bayesian network structure.

Table 4. An example of selecting the optimum Bayesian model.

Machine Learning Algorithm	Discretization Algorithm	Discretization Interval	Missing Value Processing Method	MDL Score
Maximum Spanning Tree	R2 GenOpt*	3	Dynamic Imputation	20,547.234
Taboo	R2 GenOpt*	3	Dynamic Imputation	20,493.881
EQ	R2 GenOpt*	3	Dynamic Imputation	18,311.781
TabooEQ	R2 GenOpt*	3	Dynamic Imputation	20,493.881
SopLEQ	R2 GenOpt*	3	Dynamic Imputation	20,493.881
Taboo Order	R2 GenOpt*	3	Dynamic Imputation	20,493.881
Maximum Spanning Tree	R2 GenOpt*	4	Dynamic Imputation	24,668.868
Taboo	R2 GenOpt*	4	Dynamic Imputation	24,615.515
EQ	R2 GenOpt*	4	Dynamic Imputation	22,9140.438
TabooEQ	R2 GenOpt*	4	Dynamic Imputation	24,615.515
SopLEQ	R2 GenOpt*	4	Dynamic Imputation	24,615.515
Taboo Order	R2 GenOpt*	4	Dynamic Imputation	24,615.515

Validation: The model was further validated to obtain the entropy and contingency table fit (CTF). The values obtained for the network created are given below:

- Entropy (H) = 3.7651;
- Normalized entropy (Hn) = 59.3873%;
- Hn(Complete) = 58.8520%;
- Hn(Unconnected) = 63.2849%;
- Contingency table fit = 87.9245%.

Here, we have obtained three values for normalized entropy: Hn, Hn(Complete) and Hn(Unconnected). Hn(Complete) is the normalized entropy of the model in which all variables are connected. Hn(Complete) thus corresponds to the normalized entropy value of the best possible representation of the model. In contrast, Hn(Unconnected) is the value of the model in which all variables are independent and there are no correlations. Hn(Unconnected) thus corresponds to the normalized entropy value of the worst possible representation of the model. Here, the normalized entropy (Hn) of the model obtained is closer to the value of Hn(Complete). Therefore, the model developed in this study can be considered a good representation of the pavement deterioration factors. Moreover, a higher CTF (above 87% in this case) also favours the excellent representation of the joint probability distribution. Normally, the network needs to be restructured or relearned when the CTF falls below 70% [34].

Bayesian causal inference: The Bayesian network structure (see Figure 7) has revealed several relationships between road distress parameters. Estimating the overall contribution (significance of the relationship concerning other identified relationships), KL divergence, mutual information and Pearson correlation helps in identifying the level of influence of the relationships. Table 5 thus shows how knowledge of one parameter can reduce the uncertainty of the other parameters. In this structure, the KL divergence and mutual information are the same for all relationships because the child nodes have a single parent.

Table 5. Relationship analysis among road distress parameters.

Parent	Child	KL Divergence	Overall Contribution	Mutual Information	Pearson's Correlation
International Roughness Index	Rutting	0.1470	49.2141%	0.1470	0.4249
International Roughness Index	Deflection	0.1429	47.8553%	0.1429	0.4218
International Roughness Index	Cracking	0.0088	2.9306%	0.0088	0.1025

Hence, it is clear that the IRI is related to ruts, deflection and cracks, with a higher correlation with ruts and deflection. In other words, in this situation, observing the IRI can provide important information about ruts and deflection without actually measuring them and vice versa. As mentioned earlier, this type of inference can significantly reduce the cost of retrofitting and data collection. Based on the values obtained in Table 5, the importance of each parameter in reducing the uncertainty associated with other parameters involved is estimated using the nodal force. The nodal force is derived from the arc force, which is equal to the KL divergence. The sum of the arc forces in the outgoing and incoming arcs is called the outgoing force and incoming force and gives the total force, as shown in Table 6.

Table 6. Ranking of distress parameters based on Node force.

Node	Outgoing Force	Incoming Force	Total Force
International Roughness Index	0.2986	0.0000	0.2986
Rutting	0.0000	0.1470	0.1470
Deflection	0.0000	0.1429	0.1429
Cracking	0.0000	0.0088	0.0088

For the road sections considered in this study, the ranking of distress parameters based on nodal strength suggests that knowledge of IRI relative to other variables is important in understanding the overall condition of the road and that it is, therefore, beneficial to prioritize IRI in data collection to optimize data collection efforts. In contrast, the observation of cracks is not recommended, as it provides the least information about the overall condition of the road. The IRI correlates more strongly with rutting and deflection than with cracking. In addition, monitoring the IRI can provide evidence for other parameters. It can also be inferred that a low IRI value indicates lower overall stress on the road.

Figure 8 illustrates the combination of the prior distribution and observed data to attain posterior distribution for decision making.

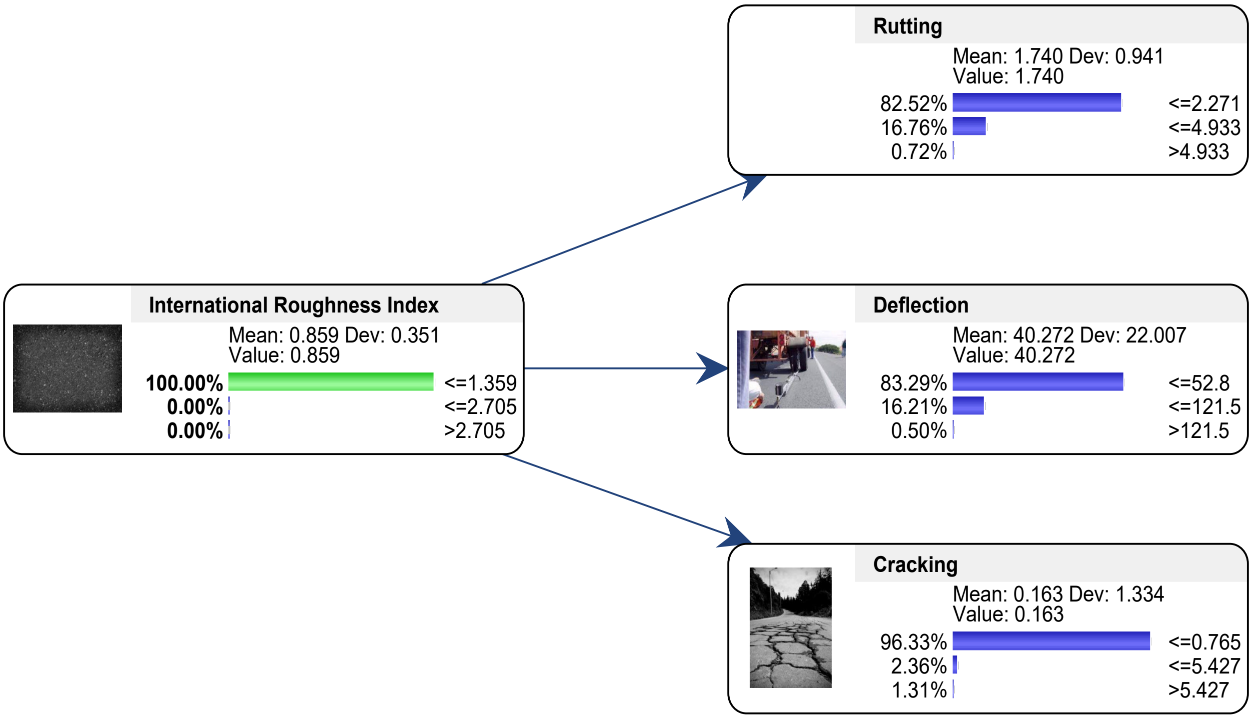

Figure 8. Posterior probability distribution based on likelihood and prior distribution.

As can be seen in Figure 8, the probability distribution associated with road distress parameters changes when new evidence (observed data) is available. Further conclusions are drawn based on the updated probability distributions (i.e., posterior distribution). It can be observed that the probability of occurrence of 100% for the class <=1.359, the probability of rutting for class <=2.271 increases from 64.52% to 82.52%, the probability of deflection for class <=52.8 increases from 66.31% to 83.29% and the probability of cracking for class <=0.765 slightly increases from 94.25% to 96.33% when the evidence is based on the availability of data. This again illustrates the correlation between the distress parameters described above.

4.2. Correlation between Road Distress Parameters and External Factors

This analysis aims to understand the behaviour of road distress parameters under the influence of dependent factors, including road factors, traffic factors and environmental factors. This allows one to know the deterioration rate of the road section based on various external conditions to plan repair schedules.

Learning algorithm: All data collected in relation to this study are included in this analysis. Various unsupervised learning algorithms are applied until a network with the best performance is obtained. Taboo learning, discretisation-R2 GenOpt* for interval 3 and processing of missing values via Dynamic imputation resulted in the network with the lowest MDL score.

Validation: The structure was further validated to obtain the entropy and contingency table fit (CTF). The values obtained for the network created are shown below:

- Entropy (H) = 8.0357;
- Normalized entropy (Hn) = 33.8716%;
- Hn(Complete) = 30.3863%;
- Hn(Unconnected) = 65.4378%;
- Contingency table adjustment = 90.0567%.

The BBN structure showing the correlations between various road distress parameters, road factors, traffic factors and environmental factors is shown in Figure 9.

Bayesian causal inference: This Bayesian structure illustrates that knowledge/observation of the interrelated factors (road, traffic and climatic factors) reduces the uncertainty associated with the road distress parameters. A data point from the dataset is randomly selected to test the model and to verify the efficiency of the BBN approach in managing road pavements. Table 7 lists the characteristics of the selected data point.

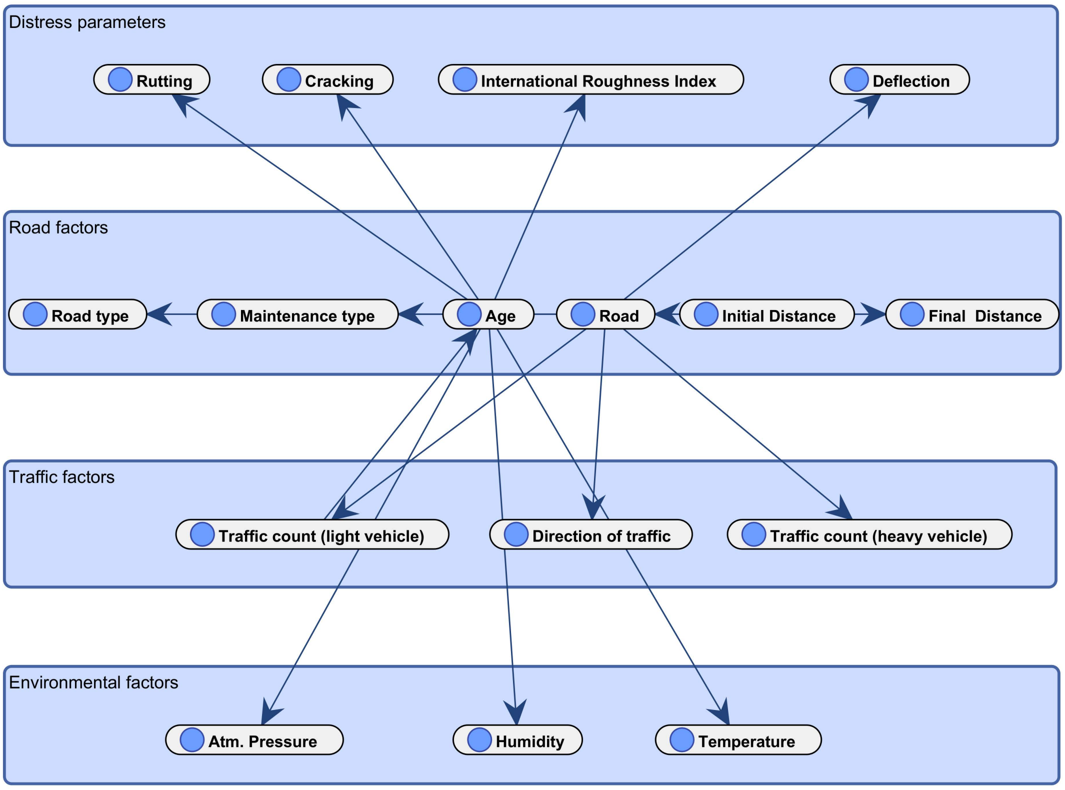

Figure 9. Bayesian network for road distress parameters, road factors, traffic factors and environmental factors.

Table 7. Datapoint selected for testing.

Features	Values
Road type	Arterial
Direction of traffic	Forward
Temperature	26
Humidity	49
Atm. Pressure	1019
Traffic count (light vehicle)	4,203,429
Traffic count (heavy vehicle)	487,421
Maintenance type	Partial
Age from last maintenance	1
Road	E11
Initial distance	8540
Final distance	8550
Cracking	0
Deflection	51
International Roughness Index	1.016
Rutting	1.75375

The features are entered into the BBN model, and the corresponding values of the distress parameters are observed in Figure 10.

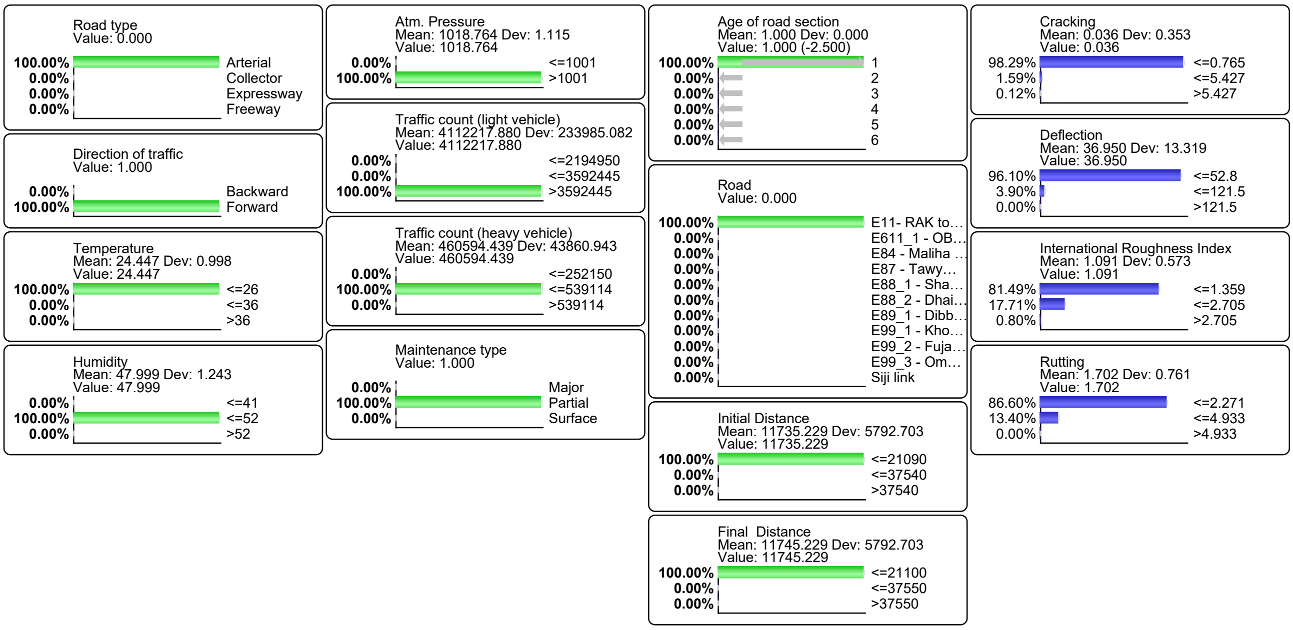

Figure 10. Testing of the BBN model.

The BBN model shows the probability distribution corresponding to each of the distress parameters. Table 8 compares the actual value and the Bayesian value for the test components.

Table 8. Actual value and Bayesian value for the test datapoint.

Variable	Actual Value	Bayesian Inference	Probability of Occurrence
Cracking	0	<=0.765	98.29%
Deflection	51	<=52.8	96.10%
International Roughness Index	1.016	<=1.359	81.49%
Rutting	1.75375	<=2.271	86.60%

This indicates that the proposed approach has a high probability of producing positive results. The value of each distress parameter was estimated at a probability of over 80%. Moreover, the model was tested in a scenario where there was no knowledge of the distress parameters, and only the data related to the explanatory factors were available. Similarly, the model can be applied to different scenarios based on the available knowledge. The strength of the approach in reducing uncertainty in unobserved variables is thus clearly defined in the case study.

4.3. Correlation between Road Type and Road Distress Parameters: Sensitivity Analysis

A sensitivity analysis was conducted to examine which distress parameter had the greatest impact on each road type. The widely used method for conducting a sensitivity analysis is to vary the probability distribution of each input factor and to observe the associated change in the target variable [19]. Here, the distress parameter is the input factor, and road type is the target variable. This approach was followed in this study to examine the variations in road type for two cases for each distress parameter.

- **Case 1:** The probability of the distress parameter is set high, e.g., P(IRI > 2.705) = 100%, while the probability of the other three parameters is not changed (no evidence). This case is shown in Table 9 as IRI = high.

Table 9. Sensitivity analysis.

States (Road Types)	IRI		Rutting		Deflection		Cracking	
	Low	High	Low	High	Low	High	Low	High
Arterial	56.9683%	76.5180%	72.5815%	82.6908%	70.8836%	91.7979%	70.1623%	66.6197%
Collector	3.2154%	4.4249%	1.4362%	4.9982%	1.6350%	1.6369%	2.8507%	10.0693%
Expressway	4.7211%	3.5550%	1.9244%	1.5604%	11.1251%	0.0000%	7.3121%	8.6105%
Freeway	35.0952%	15.5021%	24.0580%	10.7505%	16.3563%	6.5651%	19.6748%	14.7006%

- **Case 2:** The probability of the distress parameter is set to a lower value. For example: P(IRI <= 1.359) = 100%, while the probability of the other three parameters is not changed (no evidence). This case is shown in Table 9 as IRI = low. Figure 11 represents the scenarios for IRI.

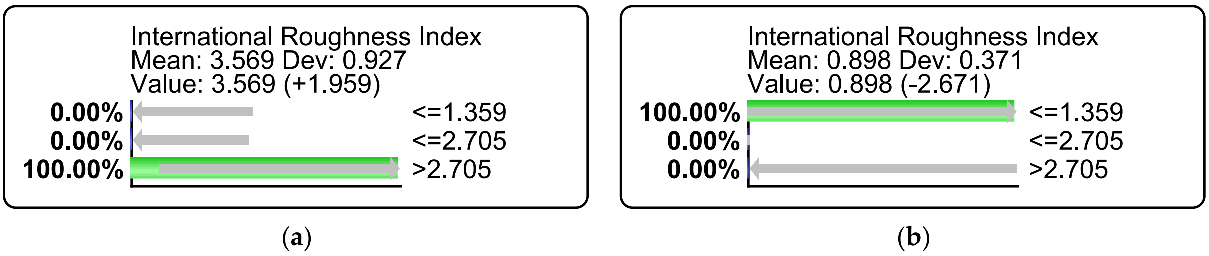

Figure 11. Scenarios tested for IRI: (**a**) case 1; (**b**) case 2.

The road type is more sensitive to a particular distress parameter if the difference between the two cases is more significant than for the other three parameters. For example, P(arterial) increased from 56.9683% to 76.518% when the IRI was changed from "low" to "high", which contributed to a 19.55% difference in the probability of arterial state with respect to IRI. Table 9 summarizes the results of the sensitivity analysis for four distress parameters.

Findings from this analysis:

- For arterial roads, deflection and IRI are the prominent distresses with a difference of 20.91% and 19.55%, respectively.
- For collector roads, cracking is the most important with a difference of 7.22%.
- For expressways, deflection is the most important with a difference of 11.13%.
- For freeways, IRI is particularly pronounced with a difference of 19.59%.

The results of this analysis suggest that deflection and IRI should take priority over other parameters when inspecting arterial roads. Knowing the main distress parameter for each type of road can save time and effort in data collection.

5. Discussion and Future Scope

In this study, Bayesian Belief Networks (BBN) are used to develop inferences for pavement management. Bayesian models effectively captured the historical data of roads with missing values to estimate (1) the interrelationship between the distress parameters, (2) the behaviour of the distress parameters concerning traffic and environmental conditions, and (3) the road distresses prominent on different road types. Mutual information is a central concept in learning and analysing the relationships between two variables and therefore has paramount importance in prioritizing and optimizing activities. It is a measure that indicates the amount of information gained by observing another variable. In this study, mutual information is emphasized as the basis for optimizing data collection.

The correlation analysis between road distress parameters showed that the highest mutual information is between IRI and rutting (see Table 5), indicating that IRI and rutting are strongly related, as shown in previous studies [33,34]. However, the mutual information between IRI and cracking is the lowest, suggesting that knowledge of IRI conveys less knowledge of cracking, and therefore, they are less interdependent. A regression analysis conducted on road networks in Saudi Arabia found that IRI is more strongly related to cracking than to rutting [8]. A similar trend was observed for rural roads in India where cracking and potholes attained higher correlation coefficients compared to other distress parameters [35]. This shows that the behaviour of the distress parameters varies depending on the road section, location and other factors. In Table 6, cracking and deflection have a lower nodal force, which means that these two parameters contribute less to the overall deterioration of the road network. The lower influence of cracking could be due to the well-maintained road network in the UAE. Another possible reason could be that the road sections investigated in this study might have been recently repaired for cracking [36]. In addition, if loads are distributed equally over an area, the deflection will be distributed evenly and consequently, and deflection could become less prominent. This could be a possible case in this study; however, it needs further investigation. The correlations identified in this study are expressed as probability distributions that represent the weightage of each relationship, considering various uncertainties arising from missing and complex data. This leads to better pavement management decisions.

Effective pavement management to improve the service life of the pavements is not possible unless the contributions made by external factors in pavement performance are becoming known [37]. This study examined the influence of various factors on the occurrence of road distresses. The results indicated that variations in the road and traffic and environmental factors influence the pavement performance as found in previous studies [38]. Increased traffic load and hot weather result in rutting even on newly constructed roads [39,40]. Another study which focused on various climatic conditions on the performance of pavement found that pavement performance is highly sensitive to temperature and less sensitive to humidity and precipitation [41]. The environmental data used in this study were collected based on the date of measurement of the road distress parameter from the road sections. The region under study does not have much climatic variation throughout the year. Hence, the influence of environment is not completely captured in this study. The impact of climatic variations will be more prominent for regions experiencing differences in climatic conditions throughout the year. Minhoto et al. [42] compared the road costs of overloaded vehicles and identical vehicles with legal loads. For overloaded vehicles, pavement costs increased by more than 100%. Although the present study agrees with the previous studies in confirming the influence of external factors on the pavement deterioration process, the conclusions obtained in both cases are based upon particular datasets. However, the application of Bayesian analysis in this study helps to incorporate expert knowledge on unknown factors, thus extending the applicability of the results. For example, ruts are the longitudinal depressions created in the roadway by heavy axle loads. When ruts increase, there is a possibility that water will accumulate in the depressions when it rains, leading to the formation of potholes and making it difficult to steer, which in turn affects road safety [43]. This type of knowledge (such as possibility of occurrence of an event) can be captured in the Bayesian model, leading to more realistic road network management.

In addition, this study has investigated the distress parameters that occur with different types of roads. Although previous studies have focused intensively on the influence of structural, traffic and environmental factors on pavement deterioration, less attention has been paid to variation according to road type. The results suggest that the effort and time required to monitor roads can be optimized if the road distress parameter prominent on each type of road is known. Bayesian belief networks are advantageous over other methods when dealing with complex data, as they facilitate the interpretation of results, offer better computational efficiency and provide more practical results [24].

The proposed approach has proven that it can effectively reduce the uncertainty of the road deterioration factors with the existing knowledge. However, the approach itself has many drawbacks. Bayesian networks discretize the continuous variables because their ability to deal with continuous values is limited. Moreover, finding the best discretization algorithm is a tedious task. Table 8 shows the importance of discretization intervals to obtain more complete results. For example, the actual value of the crack is 0, and Bayesian inference corresponds to a value <=0.765. This value (interval) is determined based on the chosen discretization interval, which in this case is 3. Tighter values can be obtained by choosing a different interval. However, the most suitable discretization algorithm and the criteria for choosing the interval are still being developed. Incorporating expert knowledge in the form of probability distributions is another challenge. Various efforts are being made to address these issues [44]. Inferences are made based on the posterior distribution, which is the result of the prior distribution (prior knowledge) and likelihood (collected data). For mathematical and computational reasons, the choice of the prior distribution should be made carefully, as it affects the posterior distribution and ultimately the decisions made. Furthermore, the prior distribution should be chosen to be a conjugate prior distribution for the likelihood function [45].

In this study, the prior distribution is based on road data collected over six years, which is a large dataset. For small datasets, the results may not be reliable. Although advanced BN software packages are available, research and development in this area are still ongoing, and various concepts are still being investigated [24]. Despite these limitations, it is evident that the BBN approach will help in optimizing road budgets. Data on road monitoring costs and other maintenance strategies should be considered in future studies to fully understand the relevance of this approach.

6. Conclusions

This study investigated the relationships between different factors contributing to pavement deterioration using Bayesian belief networks. The relationships between significant parameters of pavement deterioration, namely cracking, International Roughness Index, deflection, and rutting, were shown. The behaviour of the road distress parameters under the influence of other damage parameters, various road factors, traffic factors and environmental factors was shown. Furthermore, road distresses prominent on different road types were quantified through sensitivity analyses. The major conclusions of the study are summarized hereafter:

- Correlation analysis among road distress parameters revealed that IRI is strongly correlated with rutting and deflection, and it has a less significant correlation with cracking. IRI was found to be the central factor to represent overall pavement condition, which could be used as an indicator of the presence of rutting and deflection that generally progress at similar rates within the analysed UAE dataset. Further studies are needed to verify whether a similar trend is also generalizable for different countries
- The Bayesian analysis of road distress parameters and external factors unveiled the role of external factors in the development of road distress parameters. The BBN model developed in this paper indicated that the knowledge of external factors such as traffic, environment and road characteristics are capable to probabilistically estimating the values of road distress parameters at a higher accuracy without actually measuring them in the field.
- Sensitivity analyses revealed that the road distresses are sensitive to each road type (i.e., arterial, expressway, freeway and collector). Deflection and IRI were found to be prominent in arterial roads. Cracking was prominent in collector roads, deflection in expressways and IRI in freeways.
- Practically, these results provide the basis for optimizing road assessment data collection efforts and finding a compromise when it is not possible to obtain data about

all distress factors. Since pavement deterioration factors are expressed as probability distributions in the developed Bayesian belief network models, road maintainers could adopt the proposed BBN approach to approximate the values of unknown road distress parameters even with the minimum available data.

This study provides an insight into the application of Bayesian belief networks to optimize efforts for pavement management. The ability of BBN to capture the uncertainties associated with natural processes such as road deterioration and to incorporate the judgement of experts would lead to efficient pavement management. Prior knowledge of the behaviour of road distress parameters under the influence of external factors would enable decision makers to apply economic management strategies in terms of data collection, monitoring and maintenance. Future research should investigate the application of BBN at different levels of pavement management. Such analyses would extend the applicability of BBN to other infrastructure facilities such as bridges, tunnels and other structures, thus extending its application to nationwide structural health monitoring (SHM) of infrastructures, where carefully selected data are collected automatically to infer the most needed information required for maintenance prioritization and optimization.

Author Contributions: Conceptualization, H.A.J. and B.P.; methodology, H.A.J. and B.P.; software, H.A.J. and B.P.; validation, H.A.J. and B.P.; formal analysis, B.P.; investigation, B.P.; resources, H.A.J.; data curation, B.P.; writing—original draft preparation, B.P.; writing—review and editing, H.A.J. and B.P.; visualization, B.P.; supervision, H.A.J.; project administration, H.A.J.; funding acquisition, H.A.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research is financially supported by the office of Associate Provost for Research at UAE university under fund code 31R291.

Acknowledgments: The authors are thankful to the Ministry of Energy and Infrastructure at UAE for their support in providing the information and data related this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Martin Rogers, B.E. *Highway Engineering*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016.
2. Bennett, C.R.; de Solminihac, H.; Chamorro, A. *Data Collection Technologies for Road Management*; World Bank: Washington, DC, USA, 2006; pp. 1–8. Available online: https://openknowledge.worldbank.org/handle/10986/11776 (accessed on 21 April 2022).
3. Mcdaniel, R.; Shah, A. Asphalt additives to control cracking and rutting. *Joint Transportation Research Program*. January 2003. Available online: https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1495&context=jtrp (accessed on 21 April 2022).
4. Visintine, B.A.; Hicks, G.R.; Cheng, D.X.; Elkins, G.E.; Groeger, J. Factors Affecting the Performance of Pavement Preservation Treatments. In Proceedings of the 9th International Conference on Managing Pavement Assets, Washington, DC, USA, 18–21 May 2015; p. 23. Available online: https://vtechworks.lib.vt.edu/bitstream/handle/10919/56449/ICMPA9-000121.PDF (accessed on 21 April 2022).
5. Ismail, M.A.; Sadiq, R.; Soleymani, H.R.; Tesfamariam, S. Developing a road performance index using a Bayesian belief network model. *J. Franklin Inst.* **2011**, *348*, 2539–2555. [CrossRef]
6. Lin, J.; Yau, J.-T.; Hsiao, L.-H. Correlation Analysis Between International Roughness Index (IRI) by neural network. In Proceedings of the 82nd Annual Meeting of the Transportation Research Board, Washington, DC, USA, 12–16 January 2013; pp. 1–21.
7. Gong, H.; Sun, Y.; Shu, X.; Huang, B. Use of random forests regression for predicting IRI of asphalt pavements. *Constr. Build. Mater.* **2018**, *189*, 890–897. [CrossRef]
8. Mubaraki, M. Highway subsurface assessment using pavement surface distress and roughness data. *Int. J. Pavement Res. Technol.* **2016**, *9*, 393–402. [CrossRef]
9. Chen, X.; Dong, Q.; Zhu, H.; Huang, B. Development of distress condition index of asphalt pavements using LTPP data through structural equation modeling. *Transp. Res. Part C Emerg. Technol.* **2016**, *68*, 58–69. [CrossRef]
10. Fakhri, M.; Karimi, S.M.; Barzegaran, J. Predicting international roughness index based on surface distresses in various climate and traffic conditions using laser crack measurement system. *Transp. Res. Rec.* **2021**, *2675*, 397–412. [CrossRef]
11. Hsieh, Y.-A.; Tsai, Y.J. Machine Learning for Crack Detection: Review and Model Performance Comparison. *J. Comput. Civ. Eng.* **2020**, *34*, 04020038. [CrossRef]
12. Kabir, G.; Balek, N.B.C.; Tesfamariam, S. Consequence-based framework for buried infrastructure systems: A Bayesian belief network model. *Reliab. Eng. Syst. Saf.* **2018**, *180*, 290–301. [CrossRef]

13. Attoh-Okine, N.O. Probabilistic analysis of factors affecting highway construction costs: A belief network approach. *Can. J. Civ. Eng.* **2002**, *29*, 369–374. [CrossRef]
14. Starkova, O.; Gagani, A.I.; Karl, C.W.; Rocha, I.B.C.M.; Burlakovs, J.; Krauklis, A.E. Modelling of environmental ageing of polymers and polymer composites—durability prediction methods. *Polymers* **2022**, *14*, 907. [CrossRef]
15. Park, E.; Chang, H.J.; Nam, H.S. A Bayesian network model for predicting post-stroke outcomes with available risk factors. *Front. Neurol.* **2018**, *9*, 699. [CrossRef]
16. Gemela, J. Financial analysis using Bayesian networks. *Appl. Stoch. Model. Bus. Ind.* **2001**, *17*, 57–67. [CrossRef]
17. De Campos, L.M.; Cano, A.; Castellano, J.G.; Moral, S. Bayesian networks classifiers for gene-expression data. In Proceedings of the 2011 11th International Conference on Intelligent Systems Design and Applications, Cordoba, Spain, 22–24 November 2011; pp. 1200–1206. [CrossRef]
18. Tosun, A.; Bener, A.B.; Akbarinasaji, S. A systematic literature review on the applications of Bayesian networks to predict software quality. *Softw. Qual. J.* **2017**, *25*, 273–305. [CrossRef]
19. Mohamed, M.; Tran, D.Q. Risk-based inspection for concrete pavement construction using fuzzy sets and bayesian networks. *Autom. Constr.* **2021**, *128*, 103761. [CrossRef]
20. Mohamed, M.; Tran, D.Q. Risk-Based Inspection Model for Hot Mix Asphalt Pavement Construction Projects. *J. Constr. Eng. Manag.* **2021**, *147*, 04021045. [CrossRef]
21. Al-Saadi, I.; Wang, H.; Chen, X.; Lu, P.; Jasim, A. Multi-objective optimization of pavement preservation strategy considering agency cost and environmental impact. *Int. J. Sustain. Transp.* **2021**, *15*, 826–836. [CrossRef]
22. Van Koten, C.; Gray, A.R. An application of Bayesian network for predicting object-oriented software maintainability. *Inf. Softw. Technol.* **2006**, *48*, 59–67. [CrossRef]
23. Kotu, V.; Deshpande, B. (Eds.) Chapter 4—Classification. In *Data Science*, 2nd ed.; Morgan Kaufmann, Elsevier: Cambridge, MA, USA, 2019; pp. 65–163. [CrossRef]
24. Huelsenbeck, J.P.; Larget, B.; Miller, R.E.; Ronquist, F. Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst. Biol.* **2002**, *51*, 673–688. [CrossRef]
25. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
26. Fang, H.; Xu, H.; Yuan, H.; Zhou, Y. Discretization of Continuous Variables in Bayesian Networks Based on Matrix Decomposition. In Proceedings of the 2017 International Conference on Computing Intelligence and Information System (CIIS), Nanjing, China, 21–23 April 2017; Volume 2018, pp. 184–187. [CrossRef]
27. Furuya, T.; Ohbuchi, R. Transcoding across 3D shape representations for unsupervised learning of 3D shape feature. *Pattern Recognit. Lett.* **2020**, *138*, 146–154. [CrossRef]
28. Myung, I.J. Computational Approaches to Model Evaluation. *Int. Encycl. Soc. Behav. Sci.* **2001**, 2453–2457. [CrossRef]
29. Bayesia. S.A.S. BayesiaLab. Available online: https://www.bayesia.com/ (accessed on 31 December 2021).
30. Theodoridis, S. *Probability and Stochastic Processes*; Academic Press: Cambridge, MA, USA, 2020.
31. Zhao, J.; Liang, J.M.; Dong, Z.N.; Tang, D.Y.; Liu, Z. Accelerating information entropy-based feature selection using rough set theory with classified nested equivalence classes. *Pattern Recognit.* **2020**, *107*, 107517. [CrossRef]
32. Qiao, J.; Li, Y. Resource leveling using normalized entropy and relative entropy. *Autom. Constr.* **2018**, *87*, 263–272. [CrossRef]
33. Clogg, C.C.; Rudas, T.; Matthews, S. Analysis of Contingency Tables Using Graphical Displays Based on the Mixture Index of Fit. In *Visualization of Categorical Data*; Academic Press: Cambridge, MA, USA, 1998; pp. 425–439. [CrossRef]
34. Gerassis, S.; Albuquerque, M.; García, J.; Boente, C.; Giráldez, E.; Taboada, J.; Martín, J. Understanding complex blasting operations: A structural equation model combining Bayesian networks and latent class clustering. *Reliab. Eng. Syst. Saf.* **2019**, *188*, 195–204. [CrossRef]
35. Prasad, J.R.; Kanuganti, S.; Bhanegaonkar, P.N.; Sarkar, A.K.; Arkatkar, S. Development of Relationship between Roughness (IRI) and Visible Surface Distresses: A Study on PMGSY Roads. *Procedia-Soc. Behav. Sci.* **2013**, *104*, 322–331. [CrossRef]
36. Colombier, G. Cracking in pavements: Nature and origin of cracks. In *Prevention of Reflective Cracking in Pavements*; CRC Press: London, UK, 2004; pp. 14–29.
37. Zavagna, P.; Khanal, A.; Souliman, M. LTTP Data Analysis: Factors Affecting Pavement Roughness for the State of California. *J. Mater. Eng. Struct.* **2018**, *5*, 319–332.
38. Bhandari, S.; Luo, X.; Wang, F. Understanding the effects of structural factors and traffic loading on flexible pavement performance. *Int. J. Transp. Sci. Technol.* **2022**, *in press*. [CrossRef]
39. Alkaissi, Z.A. Effect of high temperature and traffic loading on rutting performance of flexible pavement. *J. King Saud Univ.-Eng. Sci.* **2020**, *32*, 1–4. [CrossRef]
40. Von Quintus, H.L.; Simpson, A.L. *Structural Factors for Flexible Pavements: Initial Evaluation of the SPS-1 Experiment Final Report*; No. FHWA-RD-01-166; Federal Highway Administration: Washington, DC, USA, 2003. Available online: https://rosap.ntl.bts.gov/view/dot/41562 (accessed on 21 April 2022).
41. Li, R.; Schwartz, C.W.; Forman, B. Sensitivity of Predicted Pavement Performance to Climate Characteristics. In Proceedings of the Airfield and Highway Pavement 2013: Sustainable and Efficient Pavements, Los Angeles, CA, USA, 9–12 June 2013. [CrossRef]
42. Pais, J.C.; Amorim, S.I.; Minhoto, M.J. Impact of Traffic Overload on Road Pavement Performance. *J. Transp. Eng.* **2013**, *139*, 873–879. [CrossRef]

43. Mamlouk, M.; Vinayakamurthy, M.; Underwood, B.S.; Kaloush, K.E. Effects of the International Roughness Index and Rut Depth on Crash Rates. *Transp. Res. Rec.* **2018**, *2672*, 418–429. [CrossRef]
44. Uusitalo, L. Advantages and challenges of Bayesian networks in environmental modelling. *Ecol. Modell.* **2007**, *203*, 312–318. [CrossRef]
45. Wasserman, L. Bayesian Inference. In *All of Statistics*; Springer Texts in Statistics; Springer: New York, NY, USA, 2004. [CrossRef]