ISSN 1822-427X/eISSN 1822-4288
2022 Volume 17 Issue 2: 107–134
https://doi.org/10.7250/bjrbe.2022-17.562

THE BALTIC JOURNAL OF ROAD AND BRIDGE ENGINEERING
2022/17(2)

A CASE-BASED REASONING AND RANDOM FOREST FRAMEWORK FOR SELECTING PREVENTIVE MAINTENANCE OF FLEXIBLE PAVEMENT SECTIONS

SALEH ABU DABOUS*[1,2], KHALED HAMAD[1,2], RAMI AL-RUZOUQ[1,2], WALEED ZEIADA[1,2], MAHER OMAR[1,2], LUBNA OBAID[1,2]

[1]*Department of Civil and Environmental Engineering, College of Engineering, University of Sharjah, Sharjah, United Arab Emirates*

[2]*Sustainable Civil Infrastructure Systems Research Group, Research Institute of Sciences and Engineering, University of Sharjah, Sharjah, United Arab Emirates*

Received 30 October 2021; accepted 7 March 2022

Abstract. Pavement maintenance decision-making is receiving significant attention in recent research, since pavement infrastructure is aging and deteriorating. The decision-making process is mainly related to selecting the most appropriate maintenance intervention for pavement sections to ensure performance and enhance safety. Several preventive maintenance methods have been proposed in the previous studies, yet the potential of implementing Case-Based Reasoning (CBR) in pavement maintenance decision-making has

* Corresponding author. E-mail: sabudabous@sharjah.ac.ae

Saleh ABU DABOUS (ORCID ID 0000-0002-8777-2331)
Khaled HAMAD (ORCID ID 0000-0002-8110-1115)
Rami AL-RUZOUQ (ORCID ID 0000-0001-7111-0061)
Waleed ZEIADA (ORCID ID 0000-0003-2248-5208)
Maher OMAR (ORCID ID 0000-0003-3077-1263)
Lubna OBAID (ORCID ID 0000-0002-2636-7287)

been investigated rarely. The CBR is an artificial intelligence technique, it is knowledge-based on several known cases, which are used to adapt a solution for a new case through retrieving similar cases. This research introduces the CBR to the area of pavement management to select the most appropriate preventive maintenance strategy for flexible pavement sections. The needed database was extracted from maintenance cases at Long-Term Pavement Performance Program. The criteria used to characterize condition of each section were identified based on the common practices in pavement maintenance published in the literature and implemented in the field. To assign weights to the selected criteria, different machine learning techniques were tested, and subsequently, Random Forest (RF) algorithm was selected to be integrated with the proposed CBR method producing the CBR-RF framework. A case study was analyzed to validate the proposed framework and a sensitivity analysis was conducted to assess the influence of each criterion on case retrieval accuracy and overall framework performance. Results indicated that the CBR-RF approach could assist effectively in the preventive maintenance decision-making with regard to new cases by learning from the previous similar cases. Accordingly, several agencies can depend on the proposed framework, while facing similar decision-making problems. Future research can compare the CBR-RF framework with other machine learning algorithms using the same dataset included in this research.

Keywords: case-based reasoning, decision-making, flexible pavement, preventive maintenance, random forest.

Introduction

Pavement management decision-making is critical for transportation agencies to properly plan maintenance, rehabilitation, and construction of roads. Recent reports worldwide have indicated major concerns related to pavement deterioration and road infrastructure maintenance needs. The American Society of Civil Engineers (ASCE) report card assigned Grade D for the pavement, reflecting poor condition associated with high risk, the report also indicated clear decrease in performance (ASCE, 2017). Precisely, the ASCE reported that approximately 19.7% of the federal highways are characterized by poor ride quality, are in a poor condition, and there is an increasing backlog with regard to meeting maintenance and rehabilitation requirements. Similarly, the National Transportation Group reported that over approximately 20% of the national highways had poor pavement conditions in 2016. It was also reported that in 2015, a repair budget of $120.5 billion was required for road maintenance, whereas drivers incurred extra vehicle repair and operating costs valued at $533 per driver (TRIP, 2016).

Governmental agencies and consulting companies in charge of pavement management have to continuously maintain pavements under acceptable conditions throughout their life span. To face this challenge, several agencies in the United States, Canada, and other parts of the world have developed pavement maintenance systems. Among these, the Second Strategic Highway Research Program for high traffic volume roadways that was developed by the State Department of Transportation to preserve and extend the pavement life under high traffic loads and offers a matrix of treatment options (SHRP2, 2015), Transportation Curriculum Coordination Council that provides a full-depth reclamation program for asphalt pavement rehabilitation covering both Hot In-place Recycling and inspector training for Cold In-Place Recycling Programs (AASHTO, 2016), Pavement Maintenance Program of Santa Clara that predicts the proper timing and most cost-effective treatment for each street based on available funding resources, based on the stored current and historical street condition data using a built-up computer model (Erickson, 2015), and many other similar pavement maintenance programs launched in Canada and several European countries to solve this serious problem can be mentioned.

Generally, there are two general strategies of pavement maintenance, namely, preventive and corrective maintenance. The first type, preventive maintenance, is applied when the pavement section is still serviceable to prevent minor deterioration, stop development of failure, and to decrease the need for any later corrective maintenance. At the same time, corrective maintenance is applied after a deficiency has occurred, such as loss of friction, rutting, cracking, or plastic deformation (Hicks et al., 1999; Hicks et al., 1997). Both pavement maintenance strategies are required to implement a full pavement maintenance program, yet more attention is paid to preventive maintenance techniques, so that corrective maintenance is not necessary later. Several previous studies (summarized in Table 1) have proposed diverse models to forecast pavement performance and to choose accordingly an appropriate preventive pavement maintenance intervention, and the appropriate timing for implementing it. Since pavement deterioration is a stochastic process, the Case-Based Reasoning (CBR) approach is preferred over other techniques since it is more efficient in weak-theory domains in comparison with mathematical models.

Typically, recording and presentation of knowledge in weak theory related concepts are case-specific. If many previous cases are available and can be reused for solving new problems, the CBR can be implemented (Huang et al., 2019). On the other hand, mathematical models are usually required to derive explicit associations and generalized relationships between problems and conclusions. The current research investigates the potential of applying the CBR as an artificial intelligence technique in the area of pavement maintenance

Table 1. Summary of the previous models to assess pavement condition and maintenance

Author	Models	Model Description
(Carnahan, 1988)	Markov Decision Process	Modelled the cumulative pavement damage and the selection of appropriate solutions including repair, replacement, prediction, and decision optimization.
(Sundin & Braban-Ledoux, 2001)	Artificial Intelligence	Applied AI in modelling pavement sections deterioration and diagnosis and in the decision-making process. AI used included expert systems, Artificial Neural Networks (ANN), fuzzy logic, genetic algorithms, and hybrid systems.
(Wang et al., 2003)	Integer linear programming model (ILP)	Built an ILP model to choose a group of potential projects from the highway network over a planning period of 5 years to maximize the maintenance and rehabilitation performance and minimize their disturbance cost.
(Chen et al., 2004)	Fuzzy Logic (FL)	Implemented a rule-based FL in developing pavement life-cycle costs analysis system, wherein the system user can set the rules that replicate the agency policies and strategies.
(Wei & Tighe, 2004)	Decision Trees	Designed a DT model that is based on cost-effectiveness analysis to choose treatments, define the strategy level suitable for the same treatment, and select when exactly the treatment should be implemented.
(Herabat & Tangphaisankun, 2005)	Multi-Objective Optimization	Developed a MOO model that utilizes a multi-year decision-making process in highway preventive maintenance management.
(Abo-Hashema & Sharaf, 2009)	Maintenance Unit (MU) Decision	Developed a model that predicts future MU values wherein future pavement maintenance and rehabilitation treatment essentials were determined.
(Bianchini & Bandini, 2010)	Artificial Neural Network (ANN)	Focused on using Neuro-fuzzy reasoning in the prediction of the flexible pavement performance characterized by the multilayer, feedforward neural networks, and a reasoning process using IF-THEN fuzzy rules.
(Li & Wang, 2011)	CBR	Used CBR to reduce the complexity of the pavement rehabilitation problem solving process by relying on solution parameters that were successfully applied in solving the previous similar problem.
(Tabatabaee et al., 2012)	Support Vector Classifier (SVC), Recurrent Neural Network (RNN)	Predicted the performance of a pavement infrastructure system using a two-stage model. In the first stage, SVC is used to group sections having the same characteristics, then RNN used the groups to anticipate sections' performance.
(Gong et al., 2016)	Logistic Regression (LR)	LR model was employed to investigate the influence of pre-treatment roughness condition, pre-treatment surface condition, and several other factors on the performance of Preventive Maintenance treatments.

Table 1 continued

Author	Models	Model Description
(Milad et al., 2017)	Web-based expert knowledge system	The system takes advantage of the existing web-based expert system technology in pavement problem remediation to emulate a portion of experts' professional reasoning abilities, it can then be used to assist with the maintenance of pavement.
(Chen et al., 2017)	Optimization Method	Optimization was used to compute the dynamic effectiveness/cost-effectiveness and select the optimal performance threshold.
(Yao et al., 2019)	Life-Cycle Cost Analysis (LCCA)	LCCA was employed to evaluate the effectiveness and cost-effectiveness of pavement treatments based on the equivalent area method.
(Marcelino et al., 2019)	Machine Learning (ML)	ML was employed in the prediction of pavement performance.
(Mousa et al., 2020)	Enhanced decision-making tool	Results of the previous comprehensive research addressing pavement maintenance challenges were combined into an enhanced decision-making tool that can be used to select the best maintenance treatment.
(Jia, Dai, et al., 2020; Jia, Wang, et al., 2020)	Statistical Analysis	Scatterplot analysis, cumulative frequency distribution, average effectiveness increment, and pair-samples t-test were used to evaluate the short-term and long-term effectiveness of preventive maintenance treatments.
(Amarasiri et al., 2020)	Regression Models	Regression models were developed to compare the effectiveness of different treatments and select the best treatments for use under varying conditions.
(Abu Dabous et al., 2020)	Utility-based Models	A utility approach was developed for maintenance prioritization purposes based on the condition assessment results of the pavement sections.
(Abu Dabous et al., 2021)	Evidential reasoning	The evidential reasoning theory was introduced to the area of pavement condition assessment and discusses a distress-based condition assessment method using an evidential reasoning approach. This approach utilized a condition matrix to assess the basic probability of pavement condition based on the extent and severity of different distresses identified in a pavement structure.

management. The CBR approach is proposed as an enhanced method to recommend a preventive pavement maintenance solution using knowledge collected from the previous similar cases. The method is capable of learning, adapting, and retrieving information from cases with high similarity to the case under consideration. In addition, the current research aimed at proposing an innovative hybrid approach that integrates the strengths of the random forests (RF) artificial intelligence technique with the CBR approach. The RF is capable of identifying factors with the highest significance in relation to the problem under consideration. Upon identifying the most significant factors, the appropriate maintenance plans for the flexible pavement can be recommended using the CBR approach and based on the previous acquired knowledge. The cases and information used to develop the method are extracted from the Long-Term Pavement Performance (LTPP) database. To achieve the research goal, the following sub-objectives were pursued:

1. To determine the operational, physical, and climate factors that affect the condition of flexible pavement sections;
2. To utilize random forest artificial intelligence technique in extracting factors of relative importance, identifying the weights of these factors and their contribution to the overall preventive treatment decision-making process;
3. To propose a structured CBR method that can store previous cases and then select proper maintenance treatment for a new section based on the experience accumulated from the existing cases stored in the case library;
4. To test the model performance with a sample case study application, perform sensitivity analysis, and discuss the results.

1. Background on modelling preventive pavement maintenance using Case-Based Reasoning

CBR is an innovative approach to problem-solving. The CBR is an artificial intelligence technique that accounts for all earlier similar cases with their critical attributes - “characteristics” and reuses them to answer a new query case (He et al., 2009). It can be simply defined as reasoning from experiences, or a system that solves new problems by using the results obtained from solving the previous similar problems. Besides, the CBR is based on historical data. It allows flexibility in representing the knowledge and simulates the experts in decision-making. The main concept behind the CBR approach is that similar problems have similar solutions (Stéphane & Hector, 2010). Using the CBR for a long period can maintain quality and enhance its solving ability.

The CBR method can follow one of three approaches. The first one is Textual CBR, which is suitable when the existing knowledge is mostly in presented in the textual mode. The second one is Conversational CBR, which includes an addition over the former with enhanced user interaction. The third approach is Structured CBR. It is appropriate for applications that have an intermediate to many cases. In addition, it is the most effective way of representing cases in engineering applications (Waheed & Adeli, 2005). The selected approach should typically be able to minimize the processing time (Yau & Yang, 1998), thus the structured CBR is the most adequate approach for the developed model, since it is suitable for a large number of pavement cases in the library stored in several attributes and since it provides the most effective way of representing attributes in various engineering applications.

The CBR has been applied in bridge infrastructure management to assist in selecting maintenance, rehabilitation, and replacement of bridge structures taking into consideration the budget and time constraints (Morcous et al., 2002). Waheed & Adeli (2005) proposed a CBR system that helps bridge engineers to convert the design ratings for existing bridges to load factor ratings. Another CBR application is construction project management. Chou (2009) developed a prototype system that can compare historical data available in the case library to make a preliminary project cost estimation that helps in the decision-making process (Chou, 2009). Yau & Yang (1998) used the CBR approach to find construction duration and to estimate costs with a minimum input of each project's features.

Leśniak, A., & Zima, K. (2018) proposed a CBR model to estimate the construction costs in the initial investment phase. Similarly, Kwon et al. (2020) and Hyung et al. (2020) developed hybrid models of genetic algorithms and CBR for construction cost estimation and maintenance costs. Moreover, the CBR approach can be also applied in the medical field, business stock market, education, and chemical industrial process (Ahmida & Norwawi, 2008; Chun & Park, 2005; Salem & Voskoglou, 2013; Wang et al., 2012). Some researchers focused their research on enhancing the performance of the CBR approach. For instance, Stéphane & Hector (2010) focused on improving the retrieval step in the CBR for the preliminary design stage and presented a CBR model for the embodiment design devoted to chemical engineering unit operations. As a second instance, Wang (2006) built knowledge management (KM) based on the theories and techniques of Case-Based Reasoning.

The CBR also was applied to predict the costs of pavement maintenance projects at an item level based on the available information. The model utilizes the previous experience of pavement maintenance-related construction in predicting the cost in the conceptual project phase and supports the decision-making process as it was combined with some evaluation criteria. Chou (2008) used the CBR to estimate the pavement maintenance operations costs in the early stages of budgeting. The CBR application in pavement maintenance selection has been rarely studied. Hence, this research explores the use of CBR in the selection process.

2. Pavement maintenance decision-making based on Case-Based Reasoning

The CBR process includes the cycle of the four "REs": REtrive, REuse, REvise, and REtain. It can be described as a four-step process:

(1) *Case Representation*: The process of organizing information on each case according to its contents in the system library.

(2) *Case Retrieval*: The process of looking for similar and most relevant cases in the case library given a target query case, then revising its similarity to the query case using a similarity matrix that is used to compare the degree of similarity between both cases. If the case found matches the query case and returns a value of 1, it will be referenced to, otherwise it will be rejected, and the case library will be searched again for a similar case.

(3) *Case Adaption*: The process of modifying and adapting the retrieved case to the query case.

(4) *Case Accumulation*: The process of addition, removal, and updating the existing cases in the case library to support the CBR model learning capacities.

The standard workflow of the different processes involved in the CBR is illustrated in Figure 1.

The following sections substantiate the employment of the proposed research framework.

3. Proposed CBR-RF Framework

Figure 2 demonstrates the main framework followed in this research to develop the pavement maintenance decision-making process based on the CBR-RF approach. The following sections describe the main steps of the proposed framework.

Four main steps of the proposed method are as follows.

3.1. Data acquisition

As shown in Figure 2, in retrieving old test cases, the data for the pavement sections cases were gathered from the LTPP database with all its available parameters.

As a part of the Strategic Highway Research Program (SHRP) research, the LTPP program started in 1987 under the supervision of the Federal Highway Administration in the United States. The program mainly aimed at developing a methodology for new and rehabilitated pavements, assessing pavement condition, developing methodologies to improve design and maintenance of the pavement, and determining the impacts of the construction processes, environmental conditions, traffic loads, and properties of construction material on the structural performance of flexible and concrete pavements (Elkins et al., 2003). The LTPP management system is aimed to be a comprehensive pavement database storing historical performance and condition data for around

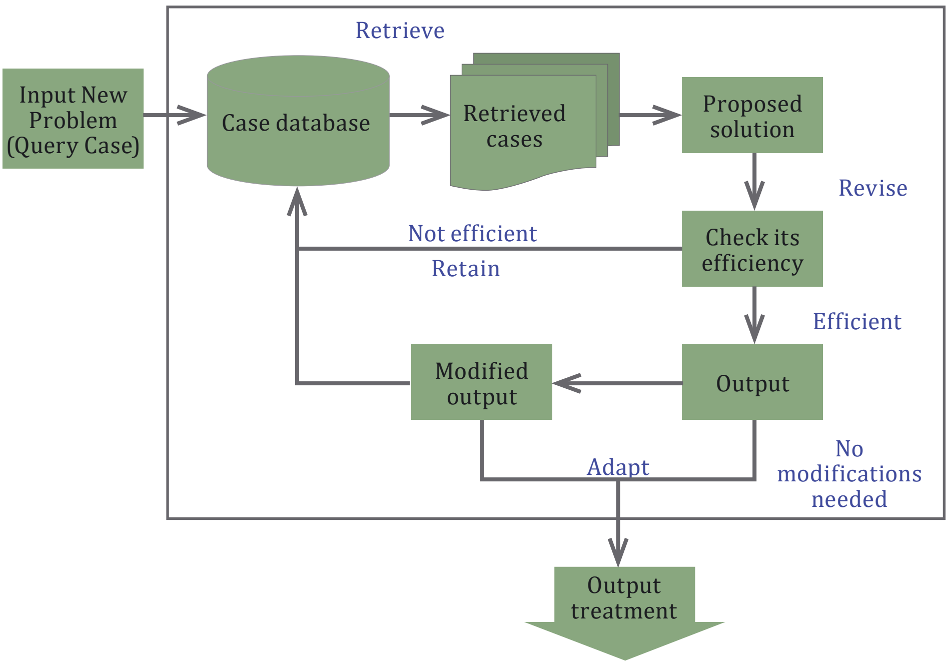

Figure 1. CBR workflow

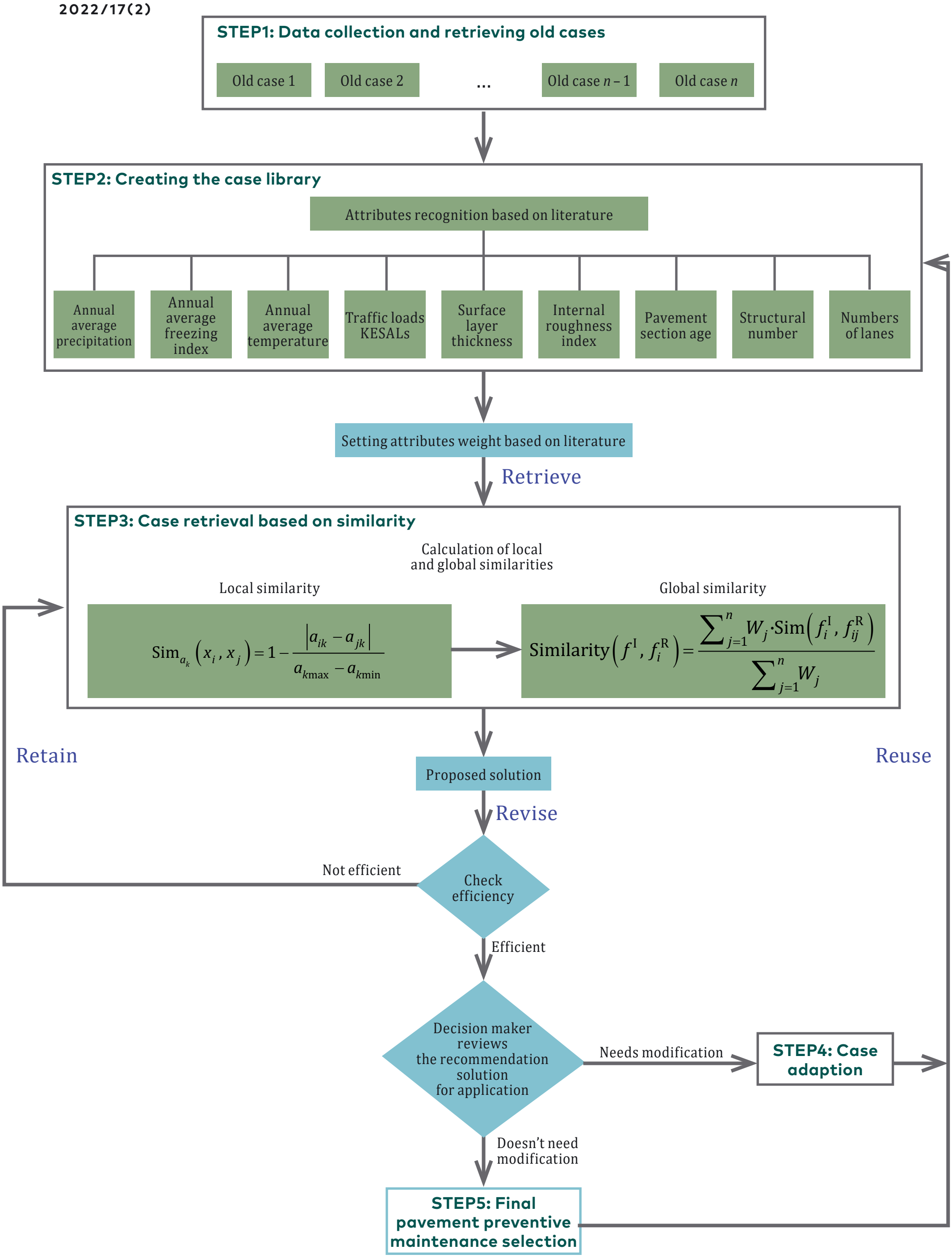

Figure 2. The proposed CBR-RF framework

2500 in-service and monitored test pavement sections from across the United States and Canada.

The datasets collected for this study included 100 pavement sections extracted randomly from the LTPP data specifically for New York, Alaska, Arizona, and Arkansas, in the United States, and for Alberta, British Columbia, Manitoba, Newfoundland, Ontario, Quebec, New Brunswick and Prince Edward Island in Canada. From the extracted data, flexible pavement sections were selected and analyzed.

Three types of data were collected for each pavement section:

1. Inventory data, which is composed of the *climate parameters*: Annual Average Precipitation, mm, Annual Average Freeze Index, degree-days, Annual Average Temperature, °C, and the Annual Average Humidity range (Min-Max, %); *cumulative traffic loads attribute*: cumulative Kelo equivalent single axle load (KESALs), the age of pavement section; *material property parameters*: Asphalt Layer and its thickness, mm; *pavement structure design parameters*: Road Functional Class, and its number of lanes.
2. Inspection data, which is composed of inspection data of the structure at specific time intervals (e.g., pavement performance measured in terms of the International Roughness Index "IRI" Section Average, m/km).
3. Maintenance and rehabilitation data, which is composed of main maintenance treatments that are possible future actions for the current problem. (e.g., AC Overlay, Crack Sealing, Crack, and seat/Rubblization, Grinding, Joint Load Transfer Restoration, Joint Sealing, PCC Overlay, Patching, Shoulder Restoration, Subdrains, Subsealing, and Thin Seal/Surface Treatment).

The standard preventive maintenance treatments to be implemented within the proposed method were identified based on the literature review (Eltahan, Daleiden, & Simpson, 1999; Hicks et al., 1999; Hicks et al., 1997). Seven preventive treatments were considered:

1. *Crack Seal*: this technique is applied to prevent water from accessing the pavement cracks. It is recommended that the entire crack be cleaned before applying the sealing material.
2. *Fog Seal*: this method is applied to improve the pavement surface, as well as to avoid raveling and oxidation processes. Through this treatment, a material, typically a diluted asphalt emulsion, is spread right on the pavement surface.
3. *Chip Seal*: this technique helps waterproof the pavement surface. It enhances friction and seals existing small cracks.
4. *Thin cold-mix seals*: this includes all treatment methods that are implemented to fill cracks, enhancing both friction and riding

quality. Slurry, cap seals, and micros facing are broad subclasses of these treatments.

5. *Thin hot-mix seals*: this method includes treatments that help in enhancing friction and riding quality with correct surface irregularities such as dense, gap, and open-graded mixes. It is similar to the conventional overlay, with the only difference being that the overlay thickness is less than 37.5 mm.
6. *Thin Hot-Mix Overlays*: this treatment is applied to remedy surface irregularities that cannot be tackled by executing other treatments. Overlays are done using dense, open, and gap-graded mixes and sometimes include modified binders. Dense- and gap-graded mixes are beneficial in sealing the pavement surface, improving ride quality, and skid resistance. On the other hand, open-graded mixes are beneficial in improving ride quality, surface friction, and enhancing the ability of water to drain off the pavement.
7. *Other treatments* such as AC shoulder replacement, full-depth patching of AC pavement, Machine Premix Patch, Manual Premix Spot Patch, etc.

3.2. Case library collection

The preliminary set of important parameters was defined to form a case library, and the parameters with the lowest importance that have small or no impact on the overall pavement treatment decision were neglected. All decisions were made based on the previous literature. When the final set of parameters was recognized, the weights and the relative significance of each attribute were set based on their relative importance collected from the previous research. Once the case library was created, it was used to recommend the most appropriate pavement treatments for any new query case.

Choosing the most appropriate parameters for the Case Library

In the previous studies, the parameters summarized in Table 2 mostly represent the pavement condition in the year of the treatment application. Based on the literature, it is acceptable to utilize the parameters as input variables for performance evaluation and preventive maintenance decision-making, since all these parameters affect the deterioration of the pavement.

The parameters depend on the preliminarily designed criteria that represent the overall condition of the pavement. The criteria were determined from the previous studies collected from literature and using expert knowledge, which mainly focuses on the following aspects:

traffic loads applied to pavement sections, climatic conditions, structural design, and pavement performance.

To form the case library, the parameters were chosen first. Table 3 shows the case criteria considered while forming the case library and factors description. A sample of the final library is illustrated in Table 4.

Table 2. Input parameters based on the previous studies

Reference	Factors Highlighted
(Abaza & Ashur, 1999)	Project location, average daily traffic, percentage of trucks, traffic projections, environmental conditions
(Hicks et al., 1997)	Roughness, the extent of distress, traffic level, existing pavement type
(Eltahan et al., 1999)	Treatment age
(Al-Mansour et al., 1994)	Pavement age, traffic load, climate
(Abu-Samra et al., 2017)	Climatic condition, pavement age, material properties "surface layer depth", average daily traffic, roughness measurements, transverse cracking amount, rutting amount

Table 3. Parameter description and attributes

Parameter Category		Parameter Description	Parameter Type
Pavement section Age, years		The number of years from the day of construction till the day of treatment application	Numeric
Climatic Conditions	Annual Average Precipitation, mm	If the climate data on the treatment date is missing, the climatic conditions of the previous year are used	Numeric
	Annual Average Temperature, °C		Numeric
	Annual Average Freeze Index		Numeric
Traffic loads (Cumulative KESALs)		The traffic loads of ADT are taken from the State of Highway Administration. The traffic is measured in the unit of KESALs cumulatively from the year of construction till the year of treatment or in terms of two consecutive treatments. If any traffic load in the study period is missing, it is calculated by interpolating/ extrapolating the known traffic loads	Numeric
Pavement Structure (No of lanes)		Indicates the structural capacity of the section	Numeric
Material Property (Asphalt Surface Thickness, mm)		The surface material is chosen to be studied: Asphalt Concrete Pavement	Numeric
Pavement Performance (IRI Section Average, m/km)		The performance of the pavement is presented just before the treatment application in terms of the Longitudinal Profile (IRI) Section Average, m/km	Numeric
Structural Number		Structure related factors were integrated into a single structural indicator	Numeric

Table 4. Sample pavement sections (library inputs)

Ref.	Section No.	Climate			Traffic	Pavement Structure		Material Properties			Preventive Pavement Maintenance Treatment
		Annual Average Precipitation, mm	Annual Average Temp., °C	Annual Avg. Freeze Index	Traffic (Cumulative KESALs)	No. of lanes	Structural Number	Material Property (Thickness, mm)	Performance (IRI, m/km)	Age of Section, years	
81	507	0.154	0.269	0.439	0.274	1.000	0.233	0.338	0.291	0.560	Skin Patching
81	508	0.108	0.264	0.534	0.298	1.000	0.400	0.649	0.229	0.600	Skin Patching
81	509	0.152	0.251	0.494	0.243	1.000	0.300	0.493	0.371	0.520	Skin Patching
81	1803	0.155	0.159	0.777	0.373	0.000	0.125	0.236	0.291	0.520	Crack Sealing
81	1804	0.119	0.203	0.691	0.312	0.000	0.317	0.419	0.345	0.880	Crack Sealing
81	1805	0.147	0.247	0.439	0.416	0.000	0.500	1.000	0.352	0.840	Skin Patching
81	2812	0.108	0.300	0.424	0.007	0.000	0.150	0.304	0.230	0.120	Aggregate Seal Coat
81	8529	0.156	0.339	0.298	0.544	1.000	0.142	0.446	0.201	1.000	Asphalt Concrete Overlay
82	1005	0.107	0.419	0.167	0.088	1.000	0.275	0.358	0.030	0.280	Aggregate Seal Coat
82	6006	0.408	0.559	0.010	1.000	1.000	0.333	0.284	0.065	0.840	Asphalt Concrete Overlay
82	6007	0.515	0.599	0.012	0.167	1.000	0.183	0.439	0.038	0.200	Asphalt Concrete Overlay
82	9017	0.088	0.374	0.229	0.032	1.000	0.217	0.020	0.159	0.160	Crack Sealing
83	AA01	0.143	0.304	0.650	0.004	0.000	0.192	0.514	0.298	0.000	Asphalt Concrete Overlay
83	AA02	0.143	0.304	0.650	0.004	0.000	0.208	0.534	0.263	0.000	Asphalt Concrete Overlay
83	AA03	0.143	0.304	0.650	0.004	0.000	0.133	0.459	0.342	0.000	Asphalt Concrete Overlay
83	AA61	0.143	0.304	0.650	0.004	0.000	0.167	0.547	0.376	0.000	Asphalt Concrete Overlay
84	1684	0.367	0.441	0.282	0.182	1.000	0.250	0.453	0.063	0.320	AC Shoulder Replacement
84	1802	0.408	0.410	0.264	0.164	0.000	0.392	0.635	0.403	0.360	Manual Premix Spot Patch
84	6804	0.342	0.410	0.287	0.432	0.000	0.358	0.324	0.152	0.360	Mill Off AC and Overlay with AC
85	1803	0.606	0.256	0.502	0.008	1.000	0.075	0.115	0.191	0.040	Crack Sealing
87	903	0.244	0.344	0.444	0.347	0.000	0.167	0.270	0.588	0.680	Asphalt Concrete Overlay
87	960	0.244	0.344	0.444	0.347	0.000	0.108	0.243	0.780	0.680	Asphalt Concrete Overlay
87	961	0.244	0.344	0.444	0.347	0.000	0.100	0.236	0.317	0.680	Asphalt Concrete Overlay
87	1620	0.256	0.427	0.227	0.240	1.000	0.308	0.811	0.478	0.920	Asphalt Concrete Overlay
87	1622	0.412	0.449	0.478	0.005	1.000	0.233	0.412	0.392	0.000	Mill Off AC and Overlay with AC

Choosing the most appropriate parameters for the Case Library

Attribute weights were assigned according to their significance based on both previous experiences obtained from relevant studies and weights obtained by utilizing RapidMiner software. Firstly, the designated weights were normalized to the weighted average retrieved from literature were shown in (Abu-Samra et al., 2017; Hicks et al., 1999). The weighted average was computed as illustrated in Abu Dabous (2018). Then, machine learning techniques were used to extract attribute weights. To do so, RapidMiner was used to develop workflows that assign factor weights based on their relative contribution to the output (optimum treatment), and the process accuracy was recorded in terms of % accuracy and kappa. RapidMiner software was used to perform this step, where it is an open-source software platform with an integrated environment for machine learning, data mining, text mining, predictive, and business analytics.

Multiple machine learning algorithms were tested throughout the analysis, namely, Naive Bayes, Generalized Linear Model, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Trees, as well as Support Vector Machine ML techniques. As shown in Table 5, the result accuracy was measured by the relative number of correctly classified predictions. Also, the kappa statistic for the classification was utilized as an accuracy measure, which is a more robust measure than simple percentage, since it considers the correct prediction occurring by chance. The kappa statistic values were computed using the following relation (Costache & Tien Bui, 2020; Foody, 2020)indeed it is not an index of overall agreement but one of agreement beyond chance. Chance agreement is, however, irrelevant in an accuracy assessment and is anyway inappropriately modelled in the calculation of a kappa coefficient for typical remote sensing applications. The magnitude of a kappa coefficient is also difficult to interpret. Values that span the full range of widely used interpretation scales, indicating a level of agreement that equates to that estimated to arise from chance alone all the way through to almost perfect agreement, can be obtained from classifications that satisfy demanding accuracy targets (e.g. for a classification with overall accuracy of 95% the range of possible values of the kappa coefficient is −0.026 to 0.900:

$$k = \frac{p_0 - p_e}{1 - p_e}, \tag{1}$$

where k is the kappa coefficient, p_0 is the observed class, and p_e is the estimated class.

Table 5 indicates that the random forest algorithm (RF) fits best the used pavement sections in the case library. The RF algorithm uses an ensemble of unpruned DTs grown to the maximum size using a different set of the bootstrap sample made of two-thirds of the original training data, and the remaining set was held back as out-of-bag (OOB) samples used to estimate the error in the overall classification accuracy. In many cases, the OOB error was found unbiased (Breiman, 2001). The RF algorithm combines the outcomes from a relatively large number of fully grown (unpruned) Classification and Regression Trees to find the final classification of a specific data point. The internal accuracy was developed by applying the decision rules defined by the 'in the bag' ITB data to the set of the training data to the OOB data for those trees. Consequently, the final set of parameter weights optimized from both literature and random forest results considered to be adequate for the preventive pavement treatment decision-making problem solving were listed in Table 6.

Table 5. The accuracy of several machine learning techniques tested

Machine learning algorithm	Accuracy, %	Kappa
Naive Bayes	59	0.535
Generalized Linear Model	54	0.401
Deep Learning	68	0.606
Decision Tree	67	0.596
Random Forest	95	0.94
Gradient Boosted Trees	89	0.866
Support Vector Machine	91	0.892

Table 6. Final parameters weights extracted from RF

Parameter	Weight, %
Annual Average Temperature	17
Pavement Structure Design (No of lanes)	5
Material Property (Thickness, mm)	10
Performance (IRI)	13
Annual Average Freeze Index	14
Annual Average Precipitation, mm	12
Age, years	10
Traffic (cumulative KESALs)	11
Structural Number	8
Sum	**100**

The final set of parameters consisted of a set of numerical attributes that were related to the decision-making process. To model the pavement condition and predict future maintenance based on the known attributes, two sets of cases were required to build the case library:

1. *The testing set*: it was used in refining model parameters, such as attribute weights and degree of similarities;
2. *The validation set*: it was used in evaluating the refined model.

The case library was formed using Microsoft Access sheets containing both sets.

3.3. Case retrieval "Similarity Measure"

To retrieve similar cases, the appropriate analysis starts with the calculation of similarity in order to retrieve the cases with the maximum similarity index. The new query case was compared in pairs with each case stored in the database. This comparison aimed to seek the most similar case. By calculating the overall similarity, the cases with the highest similarity were chosen and later implemented in the query case.

The pair similarity of each attribute between the query case and the retrieved case were computed using different formulas according to the type of each attribute. There were three categories of attributes: numerical, ordered, and disordered attributes. All selected attributes in the case library were categorized as numerical attributes. Hence, numerical local similarity calculations were performed for each pair of attributes of the query and old cases.

- The similarity of numerical attributes a_k, $k = 1, 2, \ldots, m$;
- The similarity of case x_i and case x_j on attribute a_k was calculated as follows (Yamin et al., 2017):

$$\mathrm{Sim}_{a_k}\left(x_i, x_j\right) = 1 - \frac{\left|a_{ik} - a_{jk}\right|}{a_{k\max} - a_{k\min}}, \tag{2}$$

where $a_{k\min}$, $a_{k\max}$ are minimum and maximum values of attributes a_k;

- In the absence of any attribute data:

$$\mathrm{Sim}_{ak}\left(x_i, x_j\right) = 0. \tag{3}$$

The overall (global) similarity was based on the weighted similarity sum of the local similarities of attribute pairs, and it was calculated as follows:

$$\mathrm{Similarity}\left(f^{\mathrm{I}}, f_i^{\mathrm{R}}\right) = \frac{\sum_{j=1}^{n} W_j \cdot \mathrm{Sim}\left(f_i^{\mathrm{I}}, f_{ij}^{\mathrm{R}}\right)}{\sum_{j=1}^{n} W_j}, \tag{4}$$

where $\text{Similarity}\left(f^{\mathrm{I}}, f_i^{\mathrm{R}}\right)$ stands for global similarity measure between the input case, f^{I}, and the retrieved i^{th} case of case-base f_i^{R}; $\text{Sim}\left(f_i^{\mathrm{I}}, f_{ij}^{\mathrm{R}}\right)$ stands for local similarity for the i^{th} attribute; w_j – the weight of the case attribute a_k; and $\sum_{j=1}^{n} W_j$ stands for the sum of weights of n attributes in each case.

The assumptions that were taken into consideration throughout the CBR method are:

The cases from the library with the highest global similarity $\text{Similarity}\left(f^{\mathrm{I}}, f_i^{\mathrm{R}}\right)$ were chosen to decide on the appropriate treatment.

For the chosen case, the minimum value of the global similarity $\text{Similarity}\left(f^{\mathrm{I}}, f_i^{\mathrm{R}}\right)$ should be above 80% global similarity.

After proposing the solution with the highest global similarity, an efficiency check was performed on the retrieved case to ensure whether it satisfied the required assumptions. If it failed to satisfy the criteria, the cases were checked again for another suitable solution.

3.4. Case adaption and treatment selection

If the retrieved solution satisfied the required criteria, then it should have been verified if it required further modifications before implementation. If any verifications were needed, the case retained through case adaption to adapt the chosen pavement maintenance treatment according to the decision maker's required specifications. Corrections were made individually for each pavement treatment chosen in the analysis process to meet the conditions of assumption and specifications.

Lastly, the pavement treatment of the query case selected from the case library after adaption was accepted as the final treatment for that specific pavement section. The verified result was then accumulated in the case library database for later use.

4. Application of CBR- based method to in-service flexible pavement sections

4.1. Case study

A case study is presented to illustrate and evaluate the proposed system. 100 pavement sections extracted from the LTPP database were used by comparing the treatments recommended by the model and treatments implemented in real life. The decision-making process was performed following the CBR algorithm described above. A detailed case study example of the first pavement section with the following attributes summarized in Table 7 was presented showing step-by-step calculation.

The local and global similarity equations were used for choosing the maximum resemblance. The following steps show the implementation of the proposed CBR in preventive pavement maintenance treatment decision-making on the specified pavement section.

Step 1: Measuring the local similarity of each attribute between the query and the chosen case from the library: The local similarities of paired attributes were calculated by applying local Equations (1)–(3). The minimum and maximum values of each attribute in the case library were determined. The final set of attribute ranges is shown in Table 8.

Table 7. Query case parameters

Variable	Section Value
Annual Average Precipitation, mm	803.8
Annual Average Temperature, °C	2.7
Annual Average Freeze Index	1546
Traffic cumulative KESALs	4835
Asphalt Surface layer thickness, mm	208.3
Number of lanes	2
Performance in: IRI Section Average, m/km	2.434
Age, years	13
Structural Number	5.3

Table 8. Attribute range in the case library

Parameter	Minimum Values	Maximum Values	$a_{j\max} - a_{j\min}$
Annual Average Precipitation, mm	$a_{j\min} = 121$	$a_{j\max} = 2792.1$	2670.4
Annual Average Temperature, °C	$a_{j\min} = -2.7$	$a_{j\max} = 20$	22.7
Annual Average Freeze Index, °C °days	$a_{j\min} = 0$	$a_{j\max} = 2223$	2223
Traffic cumulative KESALs	$a_{j\min} = 45$	$a_{j\max} = 20\,143$	20 098
Asphalt Surface layer thickness, mm	$a_{j\min} = 38.1$	$a_{j\max} = 414$	375.9
Number of lanes	$a_{j\min} = 1$	$a_{j\max} = 2$	1
Performance in: IRI Section Average, m/km	$a_{j\min} = 0.6$	$a_{j\max} = 3.885$	3.279
Age, years	$a_{j\min} = 1$	$a_{j\max} = 26$	25
Structural Number	$aj_{\min} = 1$	$aj_{\max} = 13.6$	12

The values of these attributes change when the case library was modified by the addition of new cases or deletion of invalid cases. An example of calculating the local similarity of attributes between the query case and one of the cases stored in the database is shown below:

– Annual average precipitation similarity =

$$\mathrm{Sim}_{ak}\left(x_j, x_j\right) = \left(0.12\left(1 - \frac{|\text{query} - \text{database case}|}{2670.3}\right)\right) =$$

$$= \left(0.12\left(1 - \frac{|803.8 - 532.1|}{2670.3}\right)\right) \cong 0.107$$

Step 2: The global similarity of that section to the query section was calculated using Equation (4).

– Overall Similarity = $\sum \text{weight} \cdot \mathrm{Sim}\left(x_i, x_j\right)$ =

$$\sum 0.107 + 0.165 + 0.104 + 0.106 + 0.089 + 0.05 + 0.095 + 0.092 + 0.074 =$$

$$= 0.882 = 88.2\%.$$

Step 3: The same similarity check steps were repeated on all the cases in the library finding the highest three similarities, as shown in Table 9.

Since the global similarities were higher than 70%, the cases were accepted and the case with the highest adequate similarity was chosen. After processing the steps in all cases, the highest similarity recorded in this case was 98.52% with the recommended treatment of Mill Off AC and Overlay with AC.

Table 9. Pavement sections with the maximum similarity to the query case

Similarity	Case 1	Case 2	Case 3
Annual Average Precipitation, mm	100	100	100
Annual Average Temperature, °C	100	100	100
Annual Average Freeze Index	100	100	100
Traffic cumulative KESALs	99.84	99.98	99.82
Asphalt Surface layer Thickness, mm	98	98.6	95.3
Number of lanes	100	100	100
Performance in: IRI Section Average, m/km	90.82	91.52	88.14
Age, years	100	100	100
Structural Number	99.17	68.33	78.33
Overall global similarity	98.52	96.23	96.22
Recommended treatment	Mill Off AC and Overlay with AC	Mill Off AC and Overlay with AC	Mill Off AC and Overlay with AC

Step 4: The decision-maker "Revises" if the retrieved case required further modification and adaption. A valid treatment process was then chosen to be applied.

Step 5: After selecting and revising the appropriate prevention treatment (Mill Off AC and Overlay with AC) for the given query case specifications, the query case was added as "Retained" to the case database to enhance the quality of the database and learn from its outputs.

4.2. Sensitivity analysis (results and discussion)

To test the developed CBR framework for decision-making in preventive pavement maintenance, 70 cases of pavement sections were used to build the CBR library, and 30 cases were used to test the method. The treatments recommended by the pavement inspectors and implemented on 30 sections were provided in the LTPP database. These treatments were compared to the recommendation produced by the CBR method. The method produced identical maintenance treatments like the one recommended by the inspectors for 26 cases (around 87% of the validation set cases were efficient) and produced different recommendations for three cases only. This could be attributed to the fact that the case library has not yet developed fully to include similar

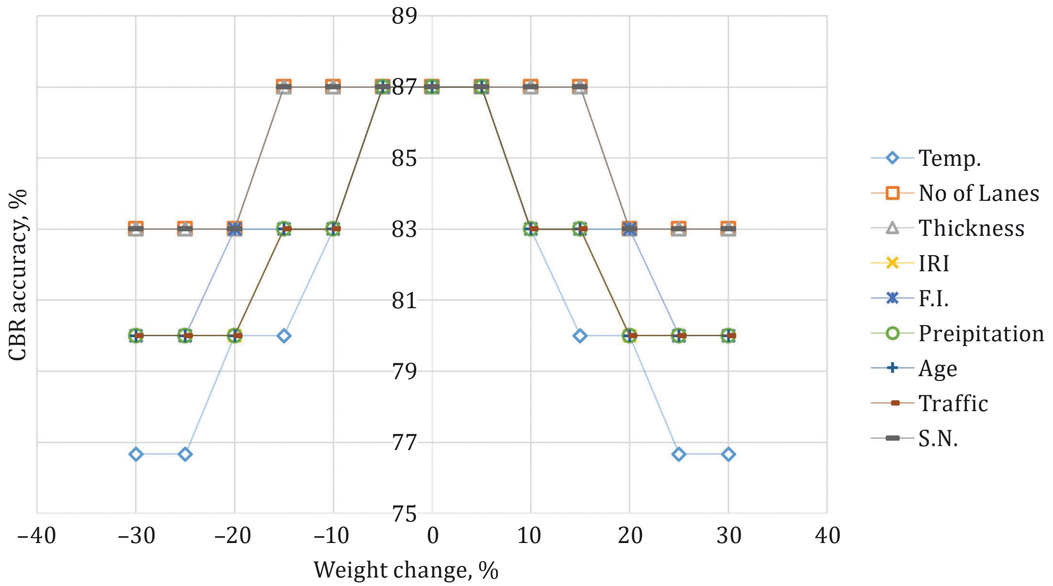

Figure 3. Results of sensitivity analysis

Table 10. Model validation results

Ref. No.	Cases / Attributes	Precipitation mm	Temperature °C	Freeze Index	Cumulative KESALs	No of lanes	Structural Number	Section Thickness	IRI, m/km	Age, years	Applied Treatments	Treatment Ref.	Proposed treatment
1	Quebec	0.255	0.238	0.695	0.238	1.000	0.308	0.453	0.557	0.480	Mill Off AC and Overlay with AC	8	8
2	New York	0.449	0.401	0.291	0.003	0.000	0.133	0.162	0.171	0.000	Aggregate Seal Coat	3	3
3	New York	0.449	0.401	0.291	0.003	0.000	0.108	0.088	0.185	0.000	Aggregate Seal Coat	3	3
4	New York	0.469	0.537	0.133	0.085	0.000	0.200	0.128	0.326	0.000	Aggregate Seal Coat	3	3
5	B. Columbia	0.408	0.559	0.010	1.000	1.000	0.333	0.284	0.065	0.840	Asphalt Concrete Overlay	2	2
6	Manitoba	0.143	0.304	0.650	0.004	0.000	0.208	0.534	0.263	0.000	Asphalt Concrete Overlay	2	2
7	Manitoba	0.143	0.304	0.650	0.004	0.000	0.133	0.459	0.342	0.000	Asphalt Concrete Overlay	2	2
8	Ontario	0.244	0.344	0.444	0.347	0.000	0.108	0.243	0.780	0.680	Asphalt Concrete Overlay	2	2
9	Ontario	0.341	0.441	0.239	0.018	1.000	0.167	0.324	0.618	0.040	Asphalt Concrete Overlay	2	2
10	Ontario	0.381	0.370	0.331	0.020	1.000	0.208	0.372	0.302	0.040	Asphalt Concrete Overlay	2	2
11	Quebec	0.435	0.339	0.483	0.141	1.000	0.242	0.378	1.000	0.240	Asphalt Concrete Overlay	2	2
12	Alberta	0.119	0.203	0.691	0.312	0.000	0.317	0.419	0.345	0.880	Crack Sealing	4	4
13	Ontario	0.341	0.441	0.239	0.018	1.000	0.167	0.236	0.735	0.040	Crack Sealing	4	3
14	Quebec	0.341	0.339	0.502	0.086	1.000	0.542	0.101	0.131	0.040	Crack Sealing	4	10
15	Alaska	0.105	0.044	1.000	0.040	0.000	0.042	0.108	0.215	0.440	Crack Sealing	4	4
16	Arizona	0.000	0.965	0.000	0.167	1.000	0.317	0.399	0.035	0.320	Crack Sealing	4	4
17	Arizona	0.013	1.000	0.000	0.140	1.000	0.508	0.351	0.000	0.280	Crack Sealing	4	4
18	Arkansas	0.474	0.806	0.039	0.179	1.000	0.308	0.236	0.117	0.400	Crack Sealing	4	4
19	Arkansas	0.474	0.806	0.039	0.179	1.000	0.217	0.196	0.107	0.400	Crack Sealing	4	4
20	Arkansas	0.474	0.806	0.039	0.179	1.000	0.267	0.196	0.161	0.400	Crack Sealing	4	4
21	Arkansas	0.286	0.855	0.016	0.098	1.000	0.258	0.358	0.056	0.200	Full Depth Patch of AC Pavement	5	5
22	New York	0.279	0.493	0.251	0.165	0.000	0.225	0.115	0.188	0.200	Machine Premix Patch	6	6
23	Ontario	0.381	0.370	0.331	0.090	1.000	0.208	0.277	0.548	0.320	Manual Premix Spot Patch	7	7
24	Quebec	0.255	0.238	0.695	0.237	1.000	0.350	0.453	0.422	0.480	Mill Off AC and Overlay with AC	8	8
25	Quebec	0.255	0.238	0.695	0.238	1.000	1.000	0.439	0.402	0.480	Mill Off AC and Overlay with AC	8	8
26	Ontario	0.381	0.370	0.331	0.020	1.000	0.233	0.338	0.393	0.040	Slurry Seal Coat	10	2
27	Arizona	0.023	0.956	0.000	0.167	1.000	0.342	0.392	0.077	0.320	Slurry Seal Coat	10	4
28	Arizona	0.000	0.965	0.000	0.167	1.000	0.267	0.203	0.140	0.320	Slurry Seal Coat	10	10
29	Arizona	0.000	0.965	0.000	0.167	1.000	0.342	0.216	0.131	0.320	Slurry Seal Coat	10	10
30	Alberta	0.141	0.242	0.522	0.216	1.000	0.225	0.223	0.256	0.600	Surface Treatment, Single Layer	12	12

cases to these three cases, so the recommendations were not consistent with what was implemented for these sections. The model validation results are illustrated in Table 10.

To implement further validation of the proposed framework, a sensitivity analysis over the different factor weights was applied, where each factor in the criteria set varied in the range of (–30% to +30%) and the weighted average of all factors was utilized to test the framework. The aim of performing sensitivity analysis is to support the understanding of the impact that a change in the decision factor weights can have on the predictions related to a decision-making process. Sensitivity test results are illustrated in Figure 3. It is indicated that the model was more sensitive to the factors with high weights, namely, temperature, freezing index, and IRI. The highest accuracy was recorded at the original weights derived using the RF machine learning approach.

Summary and Conclusions

In this paper, the CBR-RF approach was introduced to the area of pavement management decision-making to select preventive maintenance interventions for flexible pavement. The proposed framework utilizes cases exported from the LTPP database to build a knowledge base library, then the pavement section treatment was selected based on retrieving similar cases stored in the library. Within this framework, the critical factors for the decision-making process were chosen based on the previous studies and common practices.

The CBR-RF system depends on the ranked similarity scores for the stored cases using the significance weights. Different artificial intelligence machine learning techniques were assessed with regard to their efficiency in extracting factors weights. It was found that the RF technique demonstrated the highest accuracy in the weight extraction process with an overall accuracy of 95% and a kappa measurement of 0.94. Hence, the RF technique was adopted in finding the weights of different factors and ranking them accordingly. The factors with the highest weights were the annual average temperature, annual average freeze index, performance in terms of IRI, annual average precipitation, as well as pavement section age, and material properties.

For each new case, the general similarity of the cases to the query case was computed based on a numerical local similarity function assigned to each attribute. The case that held the highest similarity was considered the best match in the database, then it was retained in the library to learn from it for future cases.

The proposed method was elucidated with a case study using the LTPP database. The developed framework was trained with one part of the library stock and tested with the other part. Results showed that the proposed methodology demonstrated around an 87% efficiency rate in selecting the appropriate treatment method. Sensitivity analysis was conducted to assess the impact of varying the weights within a specific range on the overall results.

This study has some limitations that may be overcome in future research. While 100 cases were utilized to test the developed framework, a larger database can be used to improve performance if more high similarity cases are included. Thus, a more detailed investigation can be conducted to enhance the learning process and develop a full-scale system with comprehensive dataset, including several cases. Furthermore, additional variables that may affect the preventive pavement maintenance could be added to the analysis criteria. To further enhance the efficiency of the model, further investigation and testing are needed. The enhancements include taking full advantage of the CBR system in the decision-making process of pavement preventive treatment using a larger case database, which includes the maximum number of pavement sections from the LTPP database with different conditions and various proposed treatments.

REFERENCES

AASHTO. (2016). The Asphalt Pavement Rehabilitation Series. https://tc3.transportation.org/

Abaza, K. A., & Ashur, S. A. (1999). Optimum decision policy for management of pavement maintenance and rehabilitation. *Transportation Research Record*, *1655*(1), 8–15. https://doi.org/10.3141/1655-02

Abo-Hashema, M. A., & Sharaf, E. A. (2009). Development of maintenance decision model for flexible pavements. *International Journal of Pavement Engineering*, *10*(3), 173–187. https://doi.org/10.1080/10298430802169457

Abu-Samra, S., Zayed, T., & Tabra, W. (2017). Pavement Condition Rating Using Multiattribute Utility Theory. *Journal of Transportation Engineering, Part B: Pavements*, *143*(3), Article 04017011. https://doi.org/10.1061/JPEODX.0000011

Abu Dabous, S., & Al-Khayyat, G. (2018). A flexible bridge rating method based on analytical evidential reasoning and Monte Carlo simulation. *Advances in Civil Engineering*, *2018*, Article 1290632. https://doi.org/https://doi.org/10.1155/2018/1290632

Abu Dabous, S., Al-Khayyat, G., & Feroz, S. (2020). Utility-based road maintenance prioritization method using pavement overall condition rating. *Baltic Journal of Road and Bridge Engineering*, *15*(1), 126–146. https://doi.org/10.7250/bjrbe.2020-15.464

Abu Dabous, S., Zeiada, W., Al-Ruzouq, R., Hamad, K., & Al-Khayyat, G. (2021). Distress-based evidential reasoning method for pavement infrastructure condition assessment and rating. *International Journal of Pavement Engineering*, *22*(4), 455–466. https://doi.org/10.1080/10298436.2019.1622012

Ahmida, A. A., & Norwawi, N. M. (2008). Mobile case-based reasoning for reservoir gate operation decision recommendation. *3rd International Conference on Information and Communication Technologies: From Theory to Applications*, Damascus, Syria, 1–6. https://doi.org/10.1109/ICTTA.2008.4530323

Al-Mansour, A. I., Sinha, K. C., & Kuczek, T. (1994). Effects of routine maintenance on flexible pavement condition. *Journal of Transportation Engineering*, *120*(1), 65–73. https://doi.org/10.1061/(ASCE)0733-947X(1994)120:1(65)

Amarasiri, S., & Muhunthan, B. (2020). Evaluating the effectiveness of pavement preventive-maintenance treatments in mitigating longitudinal cracks in wet-freeze climatic zones. *Journal of Transportation Engineering, Part B: Pavements*, *146*(2), 1–9. https://doi.org/10.1061/JPEODX.0000158

ASCE. (2017). *A Comprehensive Assessment of America's Infrastructure*.

Carnahan, J. V. (1988). Analytical framework for optimizing pavement maintenance. *Journal of Transportation Engineering*, *114*(3), 307–322. https://doi.org/10.1061/(ASCE)0733-947X(1988)114:3(307)

Chen, C., Flintsch, G. W., & Al-Qadi, I. L. (2004). Fuzzy logic-based life-cycle costs analysis model for pavement and asset management. *6th International Conference on Managing Pavements*, Australia.

Chen, X., Zhu, H., Dong, Q., & Huang, B. (2017). Optimal thresholds for pavement preventive maintenance treatments using LTPP data. *Journal of Transportation Engineering, Part A: Systems*, *143*(6), 1–9. https://doi.org/10.1061/JTEPBS.0000044

Chou, J. (2008). Applying AHP-based CBR to estimate pavement maintenance cost. *Tsinghua Science and Technology*, *13*(S1), 114–120. https://doi.org/10.1016/S1007-0214(08)70136-6

Chou, J. (2009). Web-based CBR system applied to early cost budgeting for pavement maintenance project. *Expert Systems with Applications*, *36*(2/2), 2947–2960. https://doi.org/10.1016/j.eswa.2008.01.025

Chun, S., & Park, Y. (2005). Dynamic adaptive ensemble case-based reasoning: application to stock market prediction. *Expert Systems with Applications*, *28*(3), 435–443. https://doi.org/10.1016/j.eswa.2004.12.004

Costache, R., & Tien Bui, D. (2020). Identification of areas prone to flash-flood phenomena using multiple-criteria decision-making, bivariate statistics, machine learning and their ensembles. *Science of the Total Environment*, *712*, Article 136492. https://doi.org/10.1016/j.scitotenv.2019.136492

Elkins, G. E., Schmalzer, P. N., Thompson, T., & Simpson, A. (2003). *Long-term pavement performance information management system: Pavement performance database user reference guide* (Report No. FHWA-RD-03-088). Turner-Fairbank Highway Research Center.

Eltahan, A. A., Daleiden, J. F., & Simpson, A. L. (1999). Effectiveness of maintenance treatments of flexible pavements. *Transportation Research Record*, *1680*(1), 18–25. https://doi.org/10.3141/1680-03

Erickson, S. W. (2015). *Street pavement maintenance: Road condition is deteriorating due to insufficient funding* (Report 15-02).

Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, *239*, Article 111630. https://doi.org/10.1016/j.rse.2019.111630

Gong, H., Dong, Q., Huang, B., & Jia, X. (2016). Effectiveness analyses of flexible pavement preventive maintenance treatments with LTPP SPS-3 experiment data. *Journal of Transportation Engineering*, *142*(2). https://doi.org/10.1061/(ASCE)TE.1943-5436.0000818

He, W., Wang, F. K., Means, T., & Xu, L. D. (2009). Insight into interface design of web-based case-based reasoning retrieval systems. *Expert Systems with Applications*, *36*(3/2), 7280–7287. https://doi.org/10.1016/j.eswa.2008.09.043

Herabat, P., & Tangphaisankun, A. (2005). Multi-objective optimization model using constraint-based genetic algorithms for Thailand pavement management. *Journal of the Eastern Asia Society for Transportation Studies*, *6*, 1137–1152. https://doi.org/10.11175/easts.6.1137

Hicks, R. G., Moulthrop, J. S., & Daleiden, J. (1999). Selecting a preventive maintenance treatment for flexible pavements. *Transportation Research Record*, *1680*(1), 1–12. https://doi.org/10.3141/1680-01

Hicks, R G, Dunn, K., & Moulthrop, J. S. (1997). Framework for selecting effective preventive maintenance treatments for flexible pavements. *Transportation Research Record*, *1597*(1), 1–10. https://doi.org/10.3141/1597-01

Huang, Z., Fan, H., & Shen, L. (2019). Case-based reasoning for selection of the best practices in low-carbon city development. *Frontiers of Engineering Management*, *6*(3), 416–432. https://doi.org/10.1007/s42524-019-0036-1

Hyung, W., Kim, S., & Jo, J. (2020). Improved similarity measure in case-based reasoning : a case study of construction cost estimation. *Construction and Architectural Management*, *27*(2), 561–578. https://doi.org/10.1108/ECAM-01-2019-0035

Jia, Y., Dai, X., Wang, S., Gao, Y., Wang, J., & Zhou, W. (2020). Evaluation of long-term effectiveness of preventive maintenance treatments using LTPP SPS-3 experiment data. *Construction and Building Materials*, *247*, Article 118585. https://doi.org/10.1016/j.conbuildmat.2020.118585

Jia, Y., Wang, J., Gao, Y., Yang, M., & Zhou, W. (2020). Assessment of short-term improvement effectiveness of preventive maintenance treatments on pavement performance using LTPP data. *Journal of Transportation Engineering, Part B: Pavements*, *146*(3), 1–10. https://doi.org/10.1061/JPEODX.0000208

Kwon, N., Song, K., Ahn, Y., Park, M., & Jang, Y. (2020). Maintenance cost prediction for aging residential buildings based on case-based reasoning and genetic algorithm. *Journal of Building Engineering, 28*, 101006. https://doi.org/10.1016/j.jobe.2019.101006

Leśniak, A., & Zima, K. (2018). Cost Calculation of Construction Projects Including Sustainability Factors Using the Case Based Reasoning (CBR) Method. *Sustainability*, *10*(5), 1608. https://doi.org/10.3390/su10051608

Li, L., & Wang, K. C. P. (2011). Strategies for flexible pavement rehabilitation based on case-based reasoning. *T&DI Congress 2011: Integrated Transportation and Development for a Better Tomorrow – Proceedings of the 1st Congress of the Transportation and Development Institute of ASCE*, *479*, 32–39. https://doi.org/10.1061/41167(398)4

Marcelino, P., Antunes, M. de L., Fortunato, E., & Gomes, M. C. (2019). Machine learning approach for pavement performance prediction. *International Journal of Pavement Engineering*, *22*(3), 341–354. https://doi.org/10.1080/10298436.2019.1609673

Milad, A., Basri, N. E. A., & M., H. (2017). Prototype web-based expert system for flexible pavement maintenance. *Journal of Engineering Science and Technology*, *12*(11), 2909–2921. https://www.researchgate.net/publication/316915090_Prototype_web-based_expert_system_for_flexible_pavement_maintenance

Morcous, G., Rivard, H., & Hanna, A. M. (2002). Modeling bridge deterioration using case-based reasoning. *Journal of Infrastructure Systems*, *8*(3), 86–95. https://doi.org/10.1061/(ASCE)1076-0342(2002)8:3(86)

Mousa, M. R., Elseifi, M. A., Zhang, Z., & Gaspard, K. (2020). Development of a decision-making tool to select optimum preventive maintenance treatments in a hot and humid climate. *Transportation Research Record*, *2674*(1), 44–56. https://doi.org/10.1177/0361198119898397

Salem, A. M., & Voskoglou, M. G. (2013). Applications of the CBR methodology to medicine. *Egyptian Computer Science Journal*, *37*(7), 68–78.

SHRP2. (2015). *Project inception through December 2015*. Annual Report.

Stéphane, N., & Hector, R. (2010). Effective retrieval and new indexing method for case based reasoning: Application in chemical process design. *Engineering Applications of Artificial Intelligence*, *23*(6), 880–894. https://doi.org/10.1016/j.engappai.2010.03.005

Sundin, S., & Braban-Ledoux, C. (2001). Artificial intelligence-based decision support technologies in pavement management. *Computer-Aided Civil and Infrastructure Engineering*, *16*(2), 143–157. https://doi.org/10.1111/0885-9507.00220

Tabatabaee, N., Ziyadi, M., & Shafahi, Y. (2012). Two-stage support vector classifier and recurrent neural network predictor for pavement performance modeling. *Journal of Infrastructure Systems*, *19*(3), 266–274. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000132

TRIP. (2016). *The Interstate Highway System turns 60*.

Waheed, A., & Adeli, H. (2005). Case-based reasoning in steel bridge engineering. *Knowledge-Based Systems*, *18*(1), 37–46. https://doi.org/10.1016/j.knosys.2004.06.001

Wang, F. K. (2006). Applying case-based reasoning in knowledge management to support organizational performance. *Performance Improvement Quarterly*, *19*(2), 173–188. https://doi.org/10.1111/j.1937-8327.2006.tb00371.x

Wang, F., Zhang, Z., & Machemehl, R. B. (2003). Decision-making problem for managing pavement maintenance and rehabilitation projects. *Transportation Research Record*, *1853*(1), 21–28. https://doi.org/10.3141/1853-03

Wang, W., Wang, Y., & Gong, W. (2012). Case-based reasoning application in e-learning. *9th International Conference on Fuzzy Systems and Knowledge Discovery*, Chongqing, China, 930–933. https://doi.org/10.1109/FSKD.2012.6234117

Wei, C., & Tighe, S. (2004). Development of preventive maintenance decision trees based on cost-effectiveness analysis: an Ontario case study. *Transportation Research Record*, *1866*(1), 9–19. https://doi.org/10.3141/1866-02

Yamin, Z., Mengmeng, Z., Xiaomin, G., Zhiwei, Z., & Jianhua, Z. (2017). Research on matching method for case retrieval process in CBR based on FCM. *Procedia Engineering*, *174*, 267–274. https://doi.org/10.1016/j.proeng.2017.01.134

Yao, L., Dong, Q., Ni, F., Jiang, J., Lu, X., & Du, Y. (2019). Effectiveness and cost-effectiveness evaluation of pavement treatments using life-cycle cost analysis. *Journal of Transportation Engineering, Part B: Pavements*, *145*(2). https://doi.org/10.1061/JPEODX.0000106

Yau, N., & Yang, J. (1998). Case-based reasoning in construction management. *Computer-Aided Civil and Infrastructure Engineering*, *13*(2), 143–150. https://doi.org/10.1111/0885-9507.00094