https://doi.org/10.1016/j.conbuildmat.2021.123481

Nondestructive Prediction of Rutting Resistance of In-service Middle Asphalt Layer Based on Gene Expression Programing

Linyi Yao[1], Zhen Leng[1,*], Jiwang Jiang[1], Fujian Ni[2], Zili Zhao[2]

[1] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

[2] Department of Roadway Engineering, School of Transportation, Southeast University, Nanjing, Jiangsu, China

*corresponding author. Email: zhen.leng@polyu.edu.hk

Abstract: For a multilayered asphalt pavement, rutting resistance of the mixture in the middle asphalt layer underneath the surface course plays a significant role in the high-tempearture stability of the whole pavement structure. Thus, evaluation of the rutting resistance of the middle asphalt layer using field cores is often necessary for project-level pavement maintenance decision-making. However, the extrusion and tests of field cores are time-consuming and destructive to pavement. To address this problem, developing an empirical model to predict the rutting resistance of the middle asphalt layer at a certain service time from historical test results may be a potential alternative. This study aims to address this challenge by using the gene expression programming (GEP) method and a database composed of a large number of multiple-stress repeated load (MSRL) test results of field cores. By correlating the compound creep rate (CCR) of the middle asphalt layer from the MSRL tests to material properties, environmental and traffic parameters, field rutting depth, and middle layer age, the optimal GEP model was developed, and then compared with the conventional multiple linear regression (MLR) model built on the same database. Uncertainty analysis was also conducted through the Monte Carlo simulation (MCs). It was found that the GEP model outperformed the MLR model, with a 6%-7% higher R-square. The uncertainty analysis allows the transport agencies to estimate the reliability of their prediction and make maintenance and rehabilitation (M&R) plans according to the specified reliability target.

Keywords: In-service asphalt mixture; Rutting resistance; Pavement maintenance; Gene expression programming; Uncertainty analysis.

1 Introduction

The combined effect of traffic and environmental factors result in the accumulation of permanent deformation in asphalt pavement which finally turns into rutting at the

pavement surface [1]. Rutting is one of the common distresses of asphalt pavement, which could significantly degrade the durability of asphalt pavement, reduce the riding quality, and affect road safety [2][3]. Therefore, extensive efforts have been put forward to regularly monitor the field rutting performance of the asphalt pavement and apply pavement management systems (PMS) to systematically manage the collected rutting depth(RD) data[5]. Many transport agencies also consider rutting as one of the decision-making criteria in their maintenance planning tools [4].

Treatments for pavement rutting often require removal and replacement of the problematic layers or other costly M&R activities. For a multilayered asphalt pavement, locating the asphalt layers of insufficient rutting resistance is especially challenging given only the surface rutting depth. This is because measuring the rutting depth at the pavement surface alone is not sufficient to reflect the actual rutting resistance of the asphalt mixtures after long-term field service since the complex maintenance and rehabilitation(M&R) history of the wearing course may keep the rutting depth on the surface within a certain range but can hardly improve the rutting resistance of the underlying asphalt mixtures. It has been reported that for the asphalt pavements with three asphalt layers, such as those in China and many other countries, the middle asphalt layer contributes the most permanent deformation among the three asphalt layers [6]. Therefore, the high-temperature performance of asphalt mixtures in the middle asphalt layer is particularly concerned. However, without taking field cores, the actual rutting resistance of the in-service middle asphalt layer is unclear. Hence, decision-makers often face the dilemma whether the middle asphalt layer should be replaced, especially when the RD is in the mediocre range (8-12 mm). The traditional method is to make M&R desicions according to the pre-specified RD thresholds. The middle asphalt layer will only be rehabilitated when the RD meets the corresponding threshold. While this method is simple to implement, it completely relies on the engineering experience and ignores the specific material properties of the in-service asphalt mixtures. This motivates the current engineering practice to use the field core tests to assist in the evaluation of rutting resistance of in-service asphalt pavements and incorporate the test results into the maintenance decision-making framework.

Various field core tests were developed to evaluate the high-temperature performance of the old pavement materials. The accelerated loading tests (APTs) in small scales, such as the Asphalt Pavement Analyzer (APA) and Hamburg Wheel Tracking Tester (HWTT) tests, have been widely adopted considering the confined geometry and size of the field cores [7][8]. In addition, various uniaxial and tri-axial

loaded tests have been applied to the field cores [9]. To better simulate the actual condition of the field, new procedures have also been proposed to modify the traditional dynamic creep test, including temperature gradient control, application of multiple compressive loading, and confining stress [10]. Although tests on field cores can directly reflect the high-temperature performance of the middle asphalt layer, such repetitive tests are typcially time- and resource-consuming, cost-prohibitive and destructive to pavement structure, while covering only limited testing locations. Therefore, it would lead to considerable savings both in cost and time, if empirical models between the rutting resistance of middle asphalt layer measured from the field cores and the known or readily obtained parameters could be developed.

There have been some availalbe models which can predict the high-temperature performance of asphalt mixture. However, few of them deal with in-service asphalt mixtures, especially those mixtures in the middle asphalt layer. Instead, most of these models only apply to laboratory prepared specimens or newly built pavement [11][12], serving as a tool to assist in mixture design rather than in-service asphalt mixture evaluation. Hence, it is keenly desired to establish a model which can predict the high-temperature performance of the in-service middle asphalt layer mixtures. To fill such gap, a machine learning (ML) method, named gene expression programming (GEP), has been applied to develop the prediction model for the rutting resistance of middle asphalt layer at a certain service age. The main reason for using GEP model is that GEP has the advantages of generating the predictive equations without assuming the prior form of the underlying relationship. Thus, compared with other ML methods such as neural network, GEP model could provide more insight into the key relationships while still have the ability of modelling complex non-linear relationships. In addition, the generated functional relationship can be easily used by practitioners. GEP is an extension to genetic programming (GP) which has been widely used in the field of civil engineering [13][14]. In pavement engineering area, GEP has gained popularity in modeling the deflection basin of flexible pavements [15], back-calculation of pavement layer thickness [16], predicting the flow number [17], dynamic modulus [18] and fracture energy [19] of asphalt mixture, predicting the calfornia bearing ratio (CBR) of subgrade soil [20], evaluating the permanent deformation of asphalt mixture [21], modeling the expansion ratio and half-life of foamed bitumen [22], etc.

The research presented herein aims to: (a) develop a GEP model that can be applied to predict the rutting resistance of middle asphalt layer after long-term field

service without any destructive field core test; and (b) quantify different sources of uncertainty and investigate their effect on the GEP model output to enhance the reliability of prediction.

2 Gene Expression Programming

GEP is a natural development of the genetic programming invented by Ferreira [23]. It is a method to develop computer programs and mathematical equations based on evolutionary algorithms. GEP encodes complicated expressions as linear strings of fixed length, which are called genomes or chromosomes, and can be converted to nonlinear solutions expressed as Expression Trees (ETs) with different sizes and shapes. The fact that generally more than one genes are contained in a chromosome of GEP enables the evolution of more complex programs. There are five main components in GEP: function set, terminal set, fitness function, control parameters, and termination condition. The function set includes arithmetic operators (+, -, *, /, √), logic operators (if, and, or, not) or any other self-defined operators, while the terminal set consists of variables and constants (x_1, x_2, 1, 2). Genes in GEP are an array of symbols whose elements could be either functions or terminals. For example, an given algebraic expression: $x_1 - \sqrt{x_3} + 2x_2$, can be represented as a tree-shape ET, as shown in Figure 1.

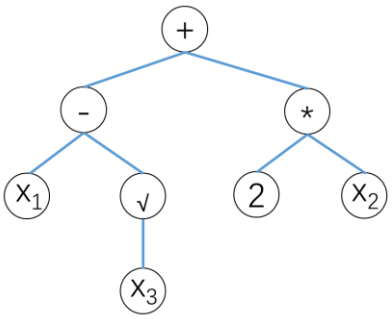


Figure 1 An example of ET.

By straightforwardly reading the ET from left to right and from top to bottom, it can be converted to a *K*-expression gene, as shown below:

$$+.\ -.\ *.\ x_1.\quad \surd.\ 2,\ x_2.\ x_3 \tag{1}$$

where x_1, x_2, and x_3 are variables; and *2* is a constant. GEP starts from the random generation of the chromosome of each individual to form the initial population. Afterward, the chromosomes are expressed, and the fitness of each individual is evaluated according to the predefined fitness function. The individuals with the best

fitness values are then selected to reproduce with modification (transposition/ rotation, mutation, and recombination/ crossover), ensuring the genetic diversification that allows for long-term evolution. The individuals in this new generation are, in their turn, subject to the same developmental process. The process continues until the termination criterion is fulfilled or a satisfactory solution is obtained. More details related to GEP can be found in [23].

3 Experimental Database

3.1 Field Core Source

Field core samples from 59 sections of seven highways in Jiangsu Province of China were extruded from the in-service pavement to build the experimental dataset. In each section, three full-depth field cores with a diameter of 150mm were collected (Figure 2a). In total, 177 field cores with different aggregate gradations, asphalt binders, pavement ages, traffic or environmental conditions were collected. Each road section has a semi-rigid base course and three asphalt layers: 4cm-thick wearing course + 6cm-thick middle layer + 7 to 8cm-thick bottom layer. The middle asphalt layers were sawed from the full-depth field cores as shown in Figure 2b. Finally, as Figure 2c shows, the middle layer materials of the field core samples with a diameter of 150mm and a height of 60mm were used for laboratory tests.

(a) (b) (c)

Figure 2 Preparation of the field cores: (a) original full-depth field core samples; (b) sawing and (c) separated middle layer samples.

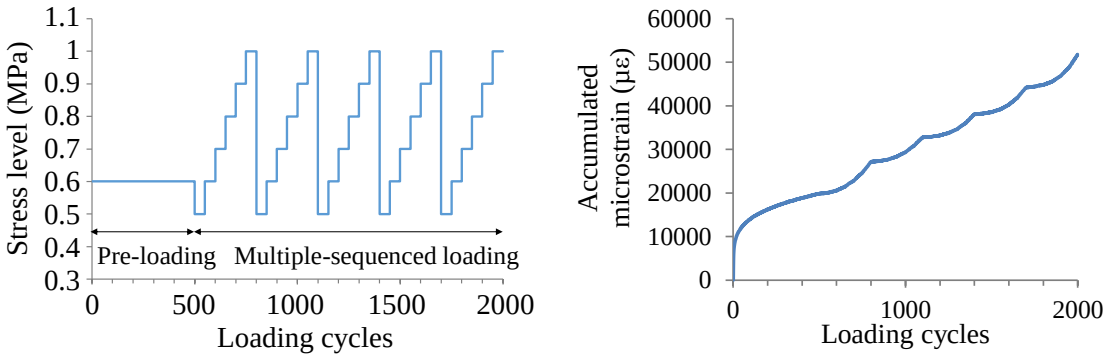
3.2 Multi-Sequenced Repeated Load (MSRL) Test

The multi-sequenced repeated load (MSRL) test was adopted to evaluate the rutting resistance of asphalt mixtures at a certain service time under laboratory conditions. The details of the procedure of the newly developed MSRL test and the feasibility of the MSRL test applied to full-depth or separated layered field cores are discussed in

[24][25]. All samples were preconditioned for more than 4 hours to achieve a constant temperature of 58 °C, which has been proved to be the worst high temperature condition of the middle layer in the asphalt pavement in Jiangsu[26]. During the MRSL test, a 50mm-diameter plate was placed at the top center of the specimen to provide a better boundary condition as shown in Figure 2c. Figure 3(a) illustrates the loading configuration of the MSRL test. A pre-loading sequence with a stress level of 600kPa was first applied for 500 cycles prior to the multi-sequenced repeated loading to warrant the testing materials entering the secondary creep stage. After that, the multi-sequenced repeated loading was repeated five times, with each comprising of six stress levels from 500kPa to 1,000kPa and each stress level continuously loaded for 50 cycles. Finally, the creep rate at each stress level was obtained. Figure 3(b) shows one typical compressive deformation curve of the field core from the MSRL test. A power function could be used to fit the relationship between the measured creep rates and stress levels as Eq. (2) presents.

$$\varepsilon(\sigma)=a\cdot\sigma^{b} \quad (2)$$

where σ is the stress level (kPa), $\varepsilon(\sigma)$ is the corresponding creep rate under stress level σ (με), and a, b are the regression coefficients.



(a) (b)

Figure 3 The MSRL test. (a) loading configuration and (b) one typical deformation curve.

3.3 Definition of the Compound Creep Rate (CCR)

The average axle load spectrum over the recent three years was extracted from the PMS of Jiangsu Province, China, from which the loading frequency of each loading interval could be calculated, as shown in

Table 1. According to Eq. (3), the tire pressure, also termed the stress level in Eq. (2), corresponding to each loading interval was computed by taking the mid-value as the representative axial load. The creep rates at the 13 stress levels could be inferred from Eq. (2). The compound creep rate (CCR) was then obtained by getting the weighted avarge of the 13 creep rates according to their corresponding loading frequencies, as presented in Eq. (4). Compared with other rutting indexes, CCR has the advantage of representing the rutting resistance of middle asphalt layer under the actual load spectrum. In addition, lower values of CCR representt better high-temperature stability of the testing material.

Table 1 Parameters of the axle load spectrum

Number, i	Intervals(KN)	Mid-value (KN)	Frequency(%)	Stress levels, σ_i (kPa)
1	20-40	30	p_1	274
2	40-60	50	p_2	382
3	60-80	70	p_3	476
4	80-100	90	p_4	560
5	100-120	110	p_5	638
6	120-140	130	p_6	712
7	140-160	150	p_7	781
8	160-180	170	p_8	847
9	180-200	190	p_9	911
10	200-220	210	p_{10}	972
11	220-240	230	p_{11}	1031
12	240-260	250	p_{12}	1088
13	260-300	280	p_{13}	1172

Note: p_i denotes the frequency of the *i*-th loading interval, differing from section to section.

$$\sigma_i/\sigma_s = (q_i/q_s)^{0.65} \tag{3}$$

$$CCR = \sum_{i=1}^{13} \varepsilon(\sigma_i) \cdot p_i \tag{4}$$

where *CCR* is the compound creep rate (με), *i* is the number of axle-load intervals, σ_i is the representative stress level of interval *i* (kPa), q_i is the representative axial load of interval *I* (KN), σ_s is the standard tire pressure for the middle layer (600kPa), q_s is

the standard axial load for the middle layer (100KN), $\varepsilon(\sigma_i)$ is the corresponding creep rate under stress level σ_i (με), and p_i is the loading frequency of axle-load interval *i*.

4 Model Development

4.1 Analysis of Influencing Factors

The factors that are believed to have an impact on the rutting resistance of the in-service middle asphalt layer are determined according to the relevant literature and the available data. Since it is difficult to collect a very large number of field cores, most of the previous models acquired data from laboratory specimens or newly built pavement using only material property indexes as input variables, without considering the actual traffic and weather conditions, nor making good use of the annually collected pavement condition data. This research, therefore, has filled in this gap by employing the following five groups of explanatory variables.

4.1.1 Variables Related to the Material Properties

Based on the available data in PMS, variables related to the properties of middle layer materials in this study include aggregate gradation, penetration and softening point of asphalt binder, and binder aggregate ratio.

4.1.2 Variables Related to the Weather

Temperature is always considered as a significant rutting factor as the asphalt binder is a time-temperature dependent material [27]. It has been found that mean temperature exerts a great influence on the permanent deformation of in-service asphalt pavement [28]. As for precipitation, Smith et al.[29] found that precipitation changes significantly affected the level of in-service asphalt mixture deformation because of moisture infiltrating. Some researchers also considered annual precipitation as one of the input variables when developing rutting models [30][31]. Therefore, two weather-related variables, namely mean temperature and proportion of rainy days, were considered in this research, which were calculated annually based on the available weather data.

4.1.3 Variables Related to the Traffic

Rutting develops when the accumulative traffic loading leads to layer densification or even shear deformation within the asphalt mixtures. Hence, the accumulative axle load terms, such as equivalent single-axle load (ESAL), are frequently used for developing rutting models. Since the middle layer samples involved all stay intact, the ESAL was calculated from the date of opening to traffic to the date of coring. Besides, overloading of varying degrees is widespread in China despite various traffic regulations, from which pavement suffered severe damage or even collapse. Rusbintardjo [32] showed that 150% overloading of single-, dual-, and triple-axle trucks will bring about 500%, 135%, and 122% level of damage, respectively. Consequently, the overloading rate, which is defined as the ratio of the number of axle loads exceeding the axle-load limit to the number of the total axle loads, was employed to describe the severity of overloading on a particular pavement segment. The equivalent single-axle dual-wheel load spectrum converted from different axle-wheel types with a loading limit of 10 tons was utilized to calculate this index [33].

4.1.4 Variables Related to the Field Rutting Depth

Field rutting depth and compound creep rate represent the high-temperature performance of asphalt mixtures on two different scales, measured by pavement condition survey and field core test, respectively. However, they are not always in consistence with each other, especially in road sections where the wearing course has been rehabilitated, but the underlying courses remain intact. Therefore, three RD-related variables were considered in this research, namely current RD, pre-treatment RD, and RD area, which refer to the rutting depth collected closest to the coring time, rutting depth before the last treatment and the area bounded by the RD curve and the age, respectively, as illustrated in Figure 4.

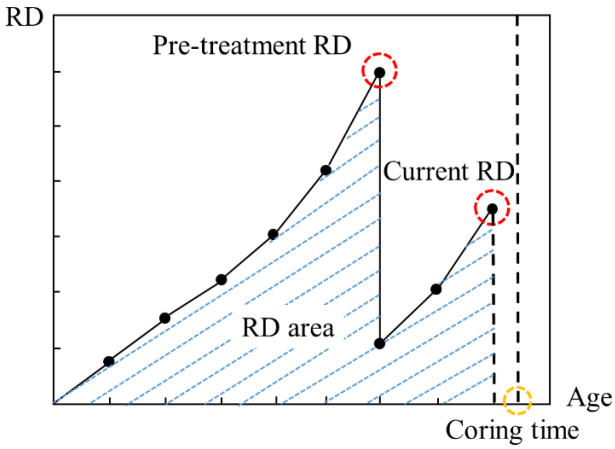


Figure 4 Schematic diagram of the adopted RD-related indicators.

4.1.5 *Variable Related to Service Age*

The service age of the middle layer samples was also considered in the prediction model as it affects the aging degree of the materials.

4.1.6 *Summary of the Variables*

Table 2 presents the descriptive statistics of the variables used in model development. It should be noted that aggregate gradation is transformed from a nominal variable to a dummy variable, with "1" representing AC-20 and "2" representing SUP-20 (AC and SUP are two different dense-graded gradations and the number 20 denotes the nominal maximum aggregate size in mm). Arguably, better model performance is expected for cases within the ranges given in Table 2.

Table 2 Descriptive statistics of the variables

Variables	Mean	Std	Min.	Max.	Median	Unit
Aggregate gradation	/	/	/	/	/	/
Penetration	69.57	3.89	55.70	74.00	70.20	0.1mm
Softening point	64.34	13.41	47.00	81.90	73.00	℃
Binder aggregate ratio	4.41	0.09	4.18	4.60	4.41	/
Mean temperature	17.53	0.55	16.25	18.72	17.87	℃
Proportion of rainy days	30.57	2.35	26.58	37.29	29.49	%
ESAL($\times 10^6$)	15.34	10.52	0.14	39.47	13.26	/
Overloading rate	43.97	10.16	20.05	56.32	48.88	%
Current RD	7.98	3.42	2.38	21.20	7.55	mm

Pre-treatment RD	5.12	5.38	0.00	16.88	4.77	mm
RD area	103.03	33.22	48.94	221.71	97.59	mm*year
Age	14.91	2.25	4.37	16.97	15.89	year
CCR	6.95	11.47	0.08	55.13	0.65	/

4.2 Model Development Using GEP

Table 3 shows the main parameters that need to be carefully tuned in the GEP algorithm. The number of chromosomes (population size) determines the number of programs that GEP will evolve. Larger population size will make the algorithm run longer. Considering the complexity of the problem, three levels of population size (1,000, 2,000, and 3,000) were tried out. The architectures of different models evolved by GEP depend on the head size and the number of genes. The head size decides the complexity of each ET, while the number of genes indicates the number of sub-ETs in the model. In this research, three levels of head size (8, 10 and 15) and gene number (3, 4 and 5) were considered. Consequently, the overall number of runs was about 3 (levels of generation number) × 3 (levels of chromosome number) × 3 (levels of gene number) × 3 (levels of head size) × 3 (replications)= 243. The fitness function was set to be the mean squared error (MSE) function.

Table 3 Parameter settings for the GEP algorithm

Parameters	Settings
Number of generations	1,000, 2,000, 3,000
Number of chromosomes	400, 500, 600
Number of genes	3, 4, 5
Head size	8, 10, 15
Linking function	+
Fitness function error type	MSE
Mutation rate	0.2
Inversion rate	0.1
IS transposition rate	0.1
RIS transposition rate	0.1
Gene transposition rate	0.1
One-point recombination rate	0.2
Two-point recombination rate	0.3

Gene recombination rate	0.1
Function set	+, -, x , /, √ , ^2

Note: IS is insertion sequence and RIS is root insertion sequence.

Figure 5 illustrates the flowchart of GEP modelling. A computational framework named geepy developed in Python was used to implement the GEP algorithm.

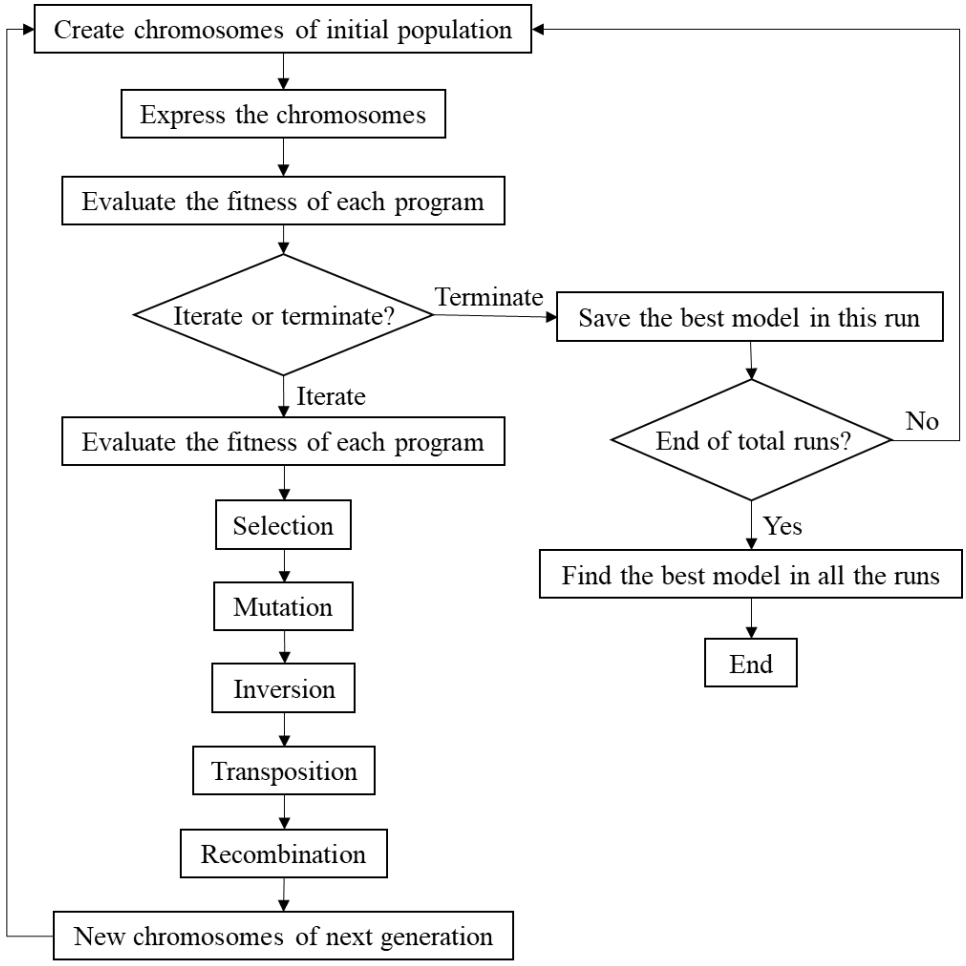


Figure 5 The flowchart of GEP modelling

The data samples were randomly divided into training (80%) and testing subsets (20%). Figure 6 presents the changes of fitness with generations during the training process. The best fitness was found at the 940th generation. The optimal GEP model for CCR is shown in Eq.(5). It was found that the GEP technique can automatically select the variables that contribute the most to construct the model.

$$Log(CCR) = SP \cdot \left[P_{rainy} - \left(\sqrt{R_{B-A}} + \mathrm{Pre} + AG\right)\right]/(Overl - ESAL)^2 + RD + Overl \cdot \left(1 + \sqrt{\left|P_{rainy} - R_{B-A} + RD\right|}\right) + (R_{B-A} \cdot T - SP)^2 \quad (5)$$

where *Log* is natural logarithm, *CCR, RD, SP,* P_{rainy}, R_{B-A}, Pre, *AG,* $Overl$, *ESAL,*

and T denote compound creep rate, current RD, softening point, proportion of rainy days, binder aggregate ratio, pre-treatment RD, aggregate gradation, overloading rate, equivalent single-axle loads, mean temperature, respectively.

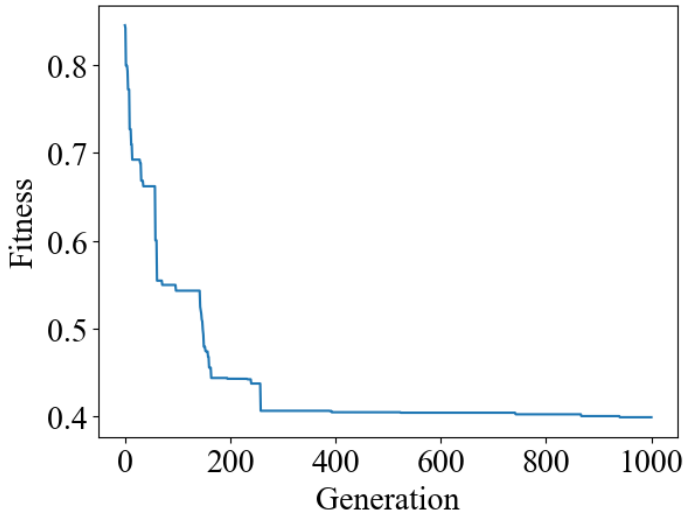
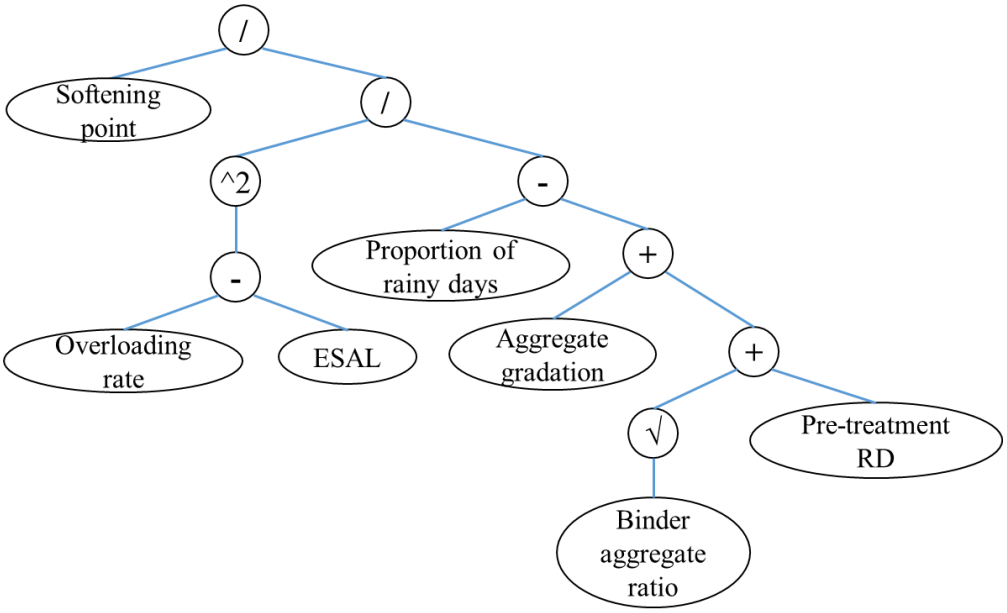


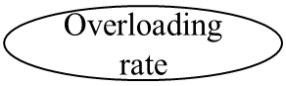
Figure 6 Fitness variations against generations.

Figure 7 presents the expression trees of the optimal GEP model. This model consists of five individual sub-models linked by an addition operation. A comparison of the measured CCR values with the predicted CCR values is shown in Figure 8. It can be observed that the determination coefficients (R^2) for training and testing dataset are 0.879 and 0.912, respectively. The points are densely around the ideal 45-degree angle line, indicating that the developed model has good performance. Besides, it seems the model encounters more difficulties in predicting smaller CCR values. One reason for this may be the usage of material property indicators tested at the construction stage instead of those tested at the current stage. However, it should be noted that the purpose of this research is to develop a prediction model without conducting any destructive field core test and the predicted CCR is expected to serve as an evaluation index. Thus, the level of CCR instead of the specific value is more important. Furthermore, the accuracy of the GEP model could be improved by accumulating more data samples in the future.

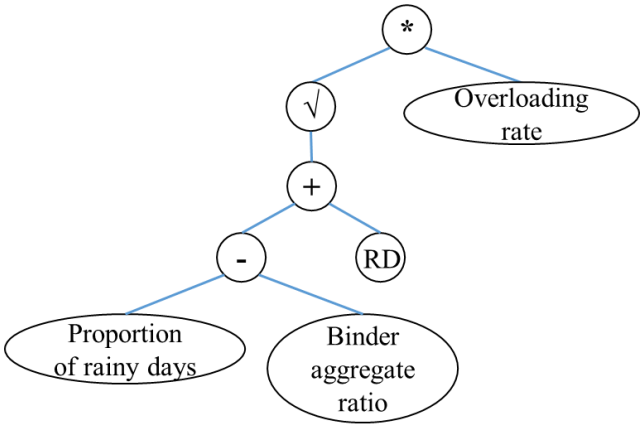


(a)

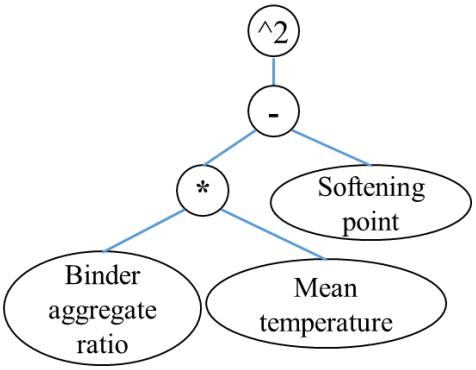
(b)



(c)

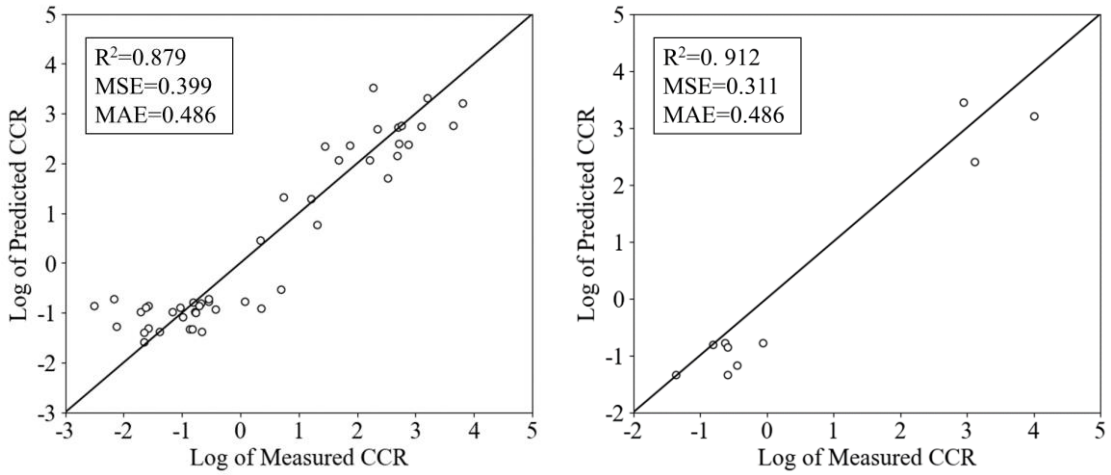


(d)



(e)

1 Figure 7 Expression trees for the best GEP model ($ET = \sum Sub\text{-}ET_i$). (a) Sub-$ET_1$; (b)
2 Sub-ET_2; (c) Sub-ET_3; (d) Sub-ET_4 and (e) Sub-ET_5.



(a) (b)

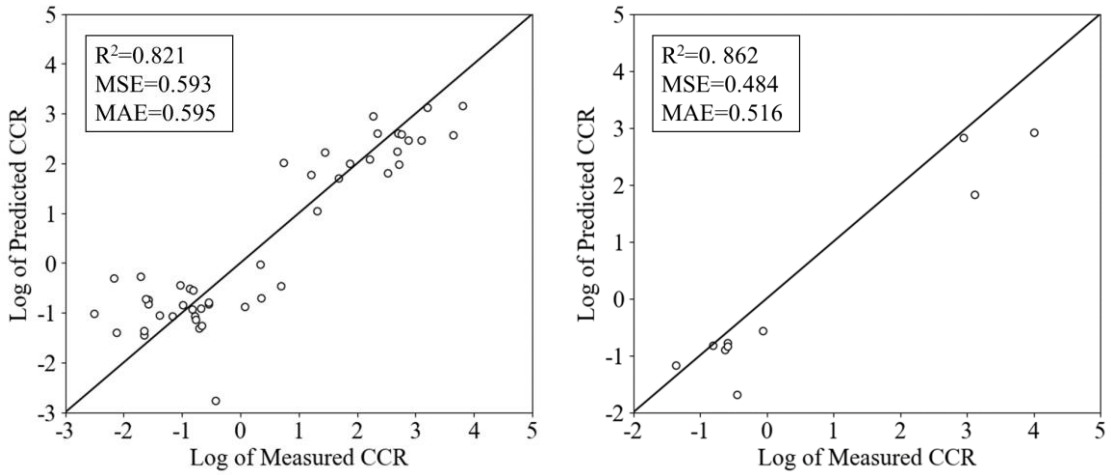
Figure 8 Measured against the predicted log of CCR using GEP model. (a) Training data and (b) Testing data.

4.3 Model Development Using MLR

For comparison purpose, multiple linear regression (MLR), a conventional tool for building a model, was also applied for the same prediction purpose. The same input variables as the GEP model were used to develop the MLR model. The formulation of the MLR model can be expressed as follows:

$$\begin{aligned} Log(CCR) = {} & -23.89 - 0.34AG - 0.02P - 0.14SP + 3.82R_{B-A} \\ & +0.04Age + 1.12T - 0.04P_{rainy} - 0.04ESAL \\ & +0.03Overl - 0.02\mathrm{Pre} - 0.01RD_{area} + 0.01RD \end{aligned} \quad (6)$$

where P, Age, RD_{area} denote penetration, middle layer service age, and RD area, respectively; Log, CCR, AG, SP, R_{B-A}, T, P_{rainy}, $ESAL$, $Overl$, Pre, RD have the same meanings as those in Eq.(5). From Eq. (6), it is found that the effects of some independent variables are not consistent with the practical obervations. For example, the predicted CCR decreases with the increase of service age and ESAL, which is thought to be problematic. Figure 9 shows the relationship between the measured and predicted CCR values for both training and testing datasets. It can be seen that the R^2 for the training and the testing dataset are 0.821 and 0.862, respectively, lower than those of the GEP model. In addition, the points around the ideal line are not as dense as those in the GEP model. All these imply that the GEP model provides better predction performance than the MLR model.



(a) (b)

Figure 9 Measured against the predicted log of CCR using MLR model. (a) Training data and (b) Testing data.

5 Performance Evaluation

5.1 Model Validity

Frank and Todeschini [34] indicated that the ratio of the number of data samples over the number of input variables should be at least three and preferrably five. In this study, the ratio is 59/9=6.6 for the GEP model and 59/12=4.9 for the MLR model. Smith [35] suggested that there is a strong relationship between the two variables if the R-value between them is higher than 0.8. Besides, more new criteria recommended by Golbraikh and Tropsha [36] and Roy and Roy [37] for external verification of the GEP model and MLR model are shown in Table 4. It is evident that the GEP model satisfies all the criteria, while the MLR model fails one criterion. Thus, it can be inferred again that the GEP model outperforms the MLR model.

Table 4 Statistical parameters of GEP and MLR model for external validation

Item	Formula	Condition	GEP	MLR
1	$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\tilde{y}_i - \bar{\tilde{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\tilde{y}_i - \bar{\tilde{y}})^2}}$	$R > 0.8$	0.938	0.906
2	$k = \frac{\sum_{i=1}^{n} y_i \tilde{y}_i}{\sum_{i=1}^{n} \tilde{y}_i^2}$	$0.85 \leq k \leq 1.15$	1.000	1.000
3	$k' = \frac{\sum_{i=1}^{n} y_i \tilde{y}_i}{\sum_{i=1}^{n} y_i^2}$	$0.85 \leq k' \leq 1.15$	0.884	0.828
4	$R_0^2 = 1 - \frac{\sum_{i=0}^{n}(\tilde{y}_i - y_i^{r_0})^2}{\sum_{i=0}^{n}(\tilde{y}_i - \bar{\tilde{y}})^2}$	Close to 1	1.000	1.000
5	$R'^2_0 = 1 - \frac{\sum_{i=0}^{n}(y_i - \tilde{y}_i^{r_0})^2}{\sum_{i=0}^{n}(y_i - \bar{y})^2}$	Close to 1	0.986	0.969
6	$m = \frac{R^2 - R_0^2}{R^2}$	<0.1	-0.121	-0.18
7	$n = \frac{R^2 - R'^2_0}{R^2}$	<0.1	-0.137	-0.218
8	$R_m^2 = R^2(1 - \lvert\sqrt{R^2 - R_0^2}\rvert)$	>0.5	0.592	0.505

Note: n is the sample size, i is the number of the sample, y_i is the i-th measured CCR, $\tilde{y}_i$ is the i-th predicted CCR, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, $\bar{\tilde{y}} = \frac{1}{n}\sum_{i=1}^{n} \tilde{y}_i$, $y_i^{r_0} = k\tilde{y}_i$ and $\tilde{y}_i^{r_0} = k'y_i$.

5.2 Appearance Frequency and Sensitivity Analysis

Appearance frequency refers to the frequency that an input variable has appeared in the best thirty GEP programs [19]. It describes the importance and contribution of each input variable to the final GEP model. As appearance frequencies in Figure 10 show, among the 12 variables, softening point, binder aggregate ratio and mean temperature had the most significant contribution to the GEP model, followed by overloading rate and proportion of rainy days.

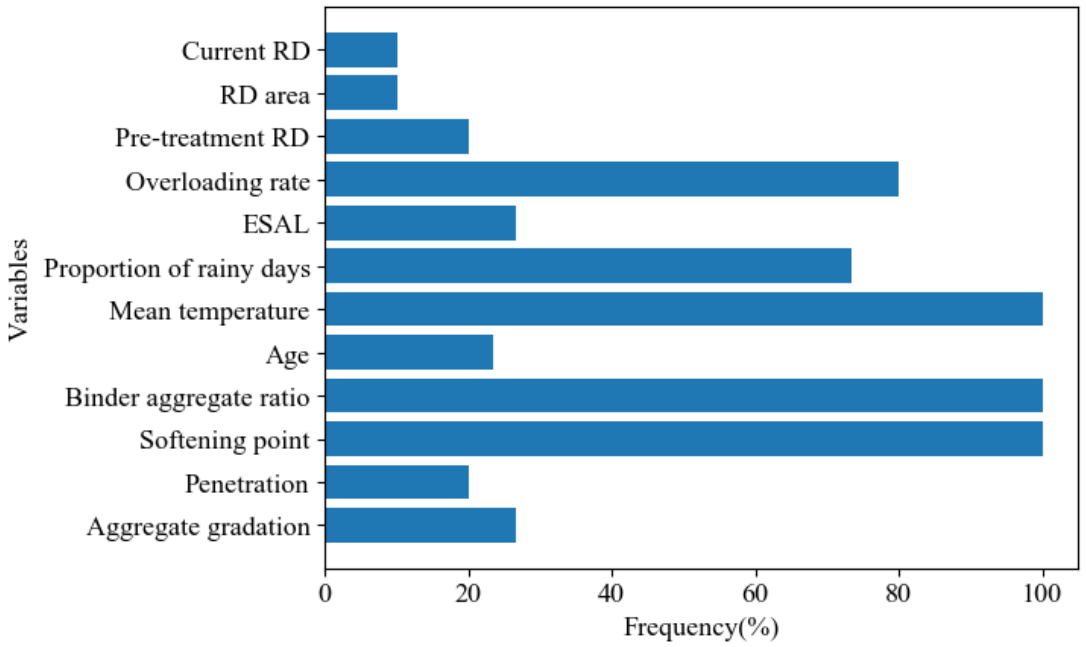


Figure 10 Appearance frequencies of the input variables in the GEP analysis.

Besides, one-dimensional sensitivity analysis (1D-SA) was performed to identify how different values of each input variable affect the model output. The following equation defines the computational method of sensitivity [38]:

$$S_i = \frac{f_{\max}(x_i) - f_{\min}(x_i)}{\sum_{j=1}^{n} f_{\max}(x_j) - f_{\min}(x_j)} \times 100 \quad (7)$$

where $f_{\max}(x_i)$ and $f_{\min}(x_i)$ are the maximum and minimum output of the i-th input variable when other input variables are equal to their mean values, n is the number of input variables and S_i is the 1D-sensitivity of the i-th input variable.

As Figure 11 shows, CCR is most sensitive to the softening point, followed by mean temperature and overloading rate. Binder aggregate ratio also has a significant influence on the predicted CCR. It can be concluded then the effect of the inherent properties of asphalt binder on the rutting resistance of asphalt mixtures is dominant, implying the necessity of using modified asphalt in the middle layer. In addition, CCR is quite sensitive to mean temperature and overloading rate, which is in close agreement with previous studies [39][40] and the practical experience. Other variables that show an evident impact on the predicted CCR include current RD, ESAL, and proportion of rainy days.

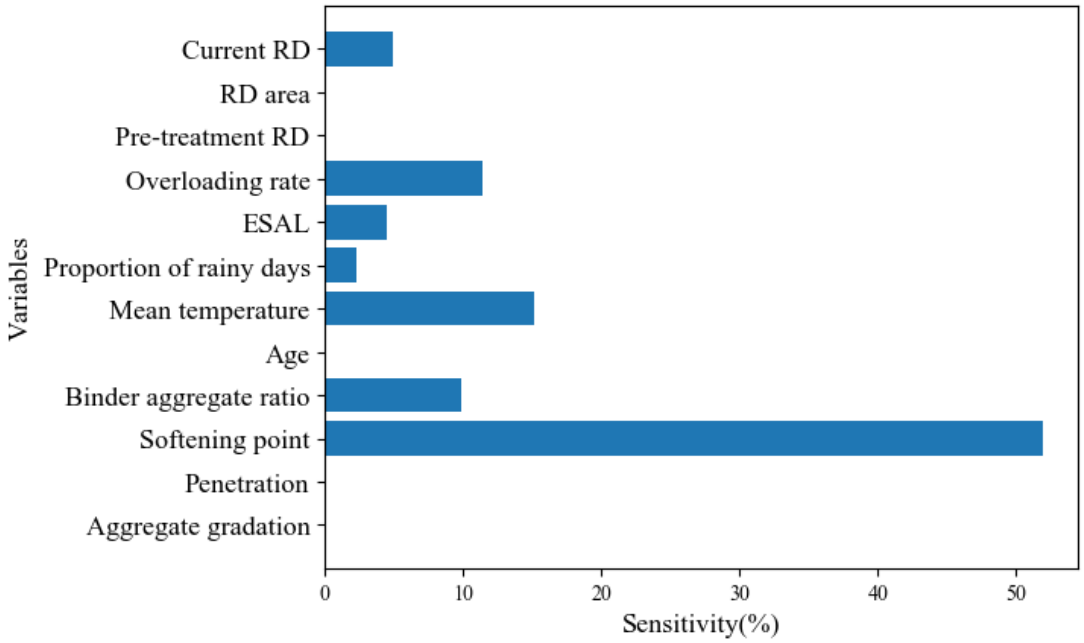


Figure 11 Results of 1D-SA

Furthermore, two-dimensional sensitivity analysis (2D-SA), which is measured as the maximum range in the output response produced by each pair of input variables[41], was also conducted to investigate the interaction effect of variables. Let $\left\{f\left(x_{ik}, x_{jm}\right): k \in \{1, \dots, L\}, m \in \{1, \dots, L\}\right\}$ denotes the sensitivity responses associated with the $L \times L$ changes of input pair $\left(x_i, x_j\right)$. The 2D-SA can be given by:

$$f_{\max}\left(x_i, x_j\right) = \max\left(f\left(x_{ik}, x_{jm}\right): k \in \{1, \dots, L\}, m \in \{1, \dots, L\}\right) \quad (8)$$

$$f_{\min}\left(x_i, x_j\right) = \min\left(f\left(x_{ik}, x_{jm}\right): k \in \{1, \dots, L\}, m \in \{1, \dots, L\}\right) \quad (9)$$

$$S_{i,j} = \frac{f_{\max}(x_i, x_j) - f_{\min}(x_i, x_j)}{\sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} f_{\max}(x_i, x_j) - f_{\min}(x_i, x_j)} \times 100 \quad (10)$$

where $f_{\max}\left(x_i, x_j\right)$, $f_{\min}\left(x_i, x_j\right)$ are the maximum and minimum output by varying the input pair $\left(x_i, x_j\right)$ simultaneously while holding the other variables at their median values, and $S_{i,j}$ is the 2D-sensitivity of the input pair $\left(x_i, x_j\right)$.

Figure 12 uses a color matrix to illustrate the 2D sensitivity of each input pair with a deeper color denoting the higher interaction. It can be observed that the interaction between softening point and mean temperature is the strongest.

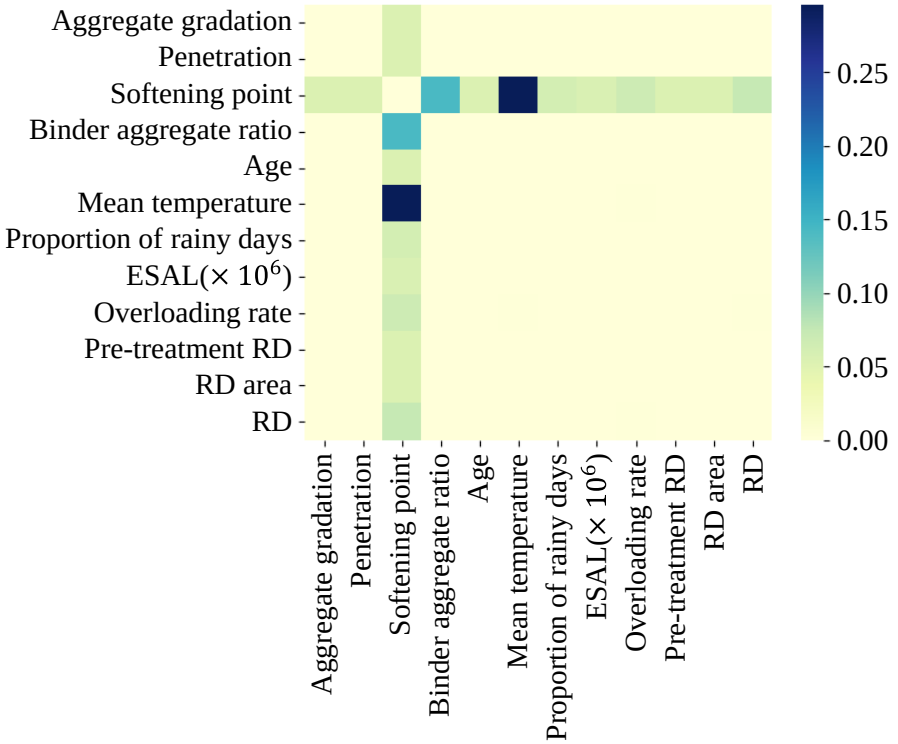
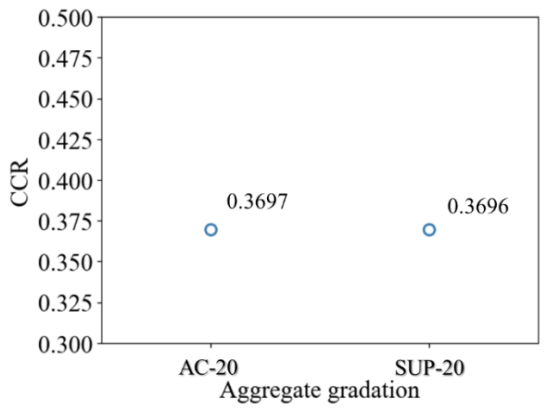


Figure 12 2D sensitivity matrix

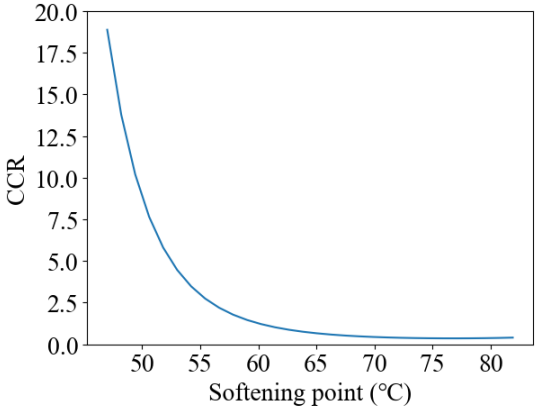
5.3 Parametric Analysis

To further verify how well the predicted target values agree with the underlying physical behavior of the investigated problem, a parameter analysis was performed by moving the objective variable within its domain while keeping other variables equal to their median values. Figure 13 shows the tendency of the predicted CCR to the variation of each input variable. AC-20 has slightly better rutting resistance than SUP-20. CCR decreases drastically with the increase of softening point, and this decline slows down significantly after the softening point is over approximately 55°C. Positive correlations were found as expected between the predicted CCR and binder aggregate ratio, mean temperature and proportion of rainy days. CCR almost keeps steady up to ESAL equal to 3×10^7 and after that, it continuously increases with increasing ESAL. It indicates that the rutting resistance of the middle asphalt layer is susceptible to extremely heavy loads. Further evidence of this can be found in the correlation between the overloading rate and the predicted CCR, as shown in Figure 13(g). As expected, the CCR of the middle layer materials still shows an increase as the current rutting becomes deeper but not that sensitive, which may be attributed to the disturbance of M&R activities implemented on the wearing course or the lagging between the external rutting performance and inner material degradation. The rational

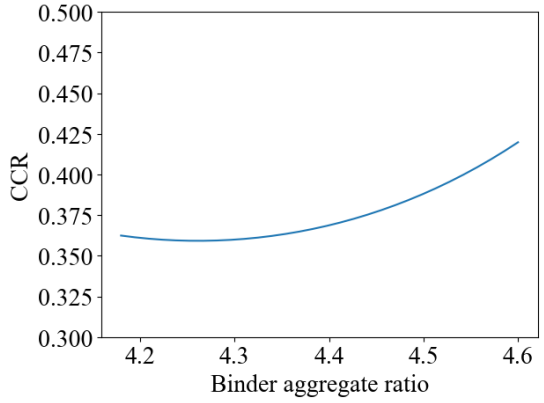
relationship between the input variables and the predicted CCR in the GEP model confirms that the developed GEP model is superior to the MLR model and thus can be used for prediction purposes.



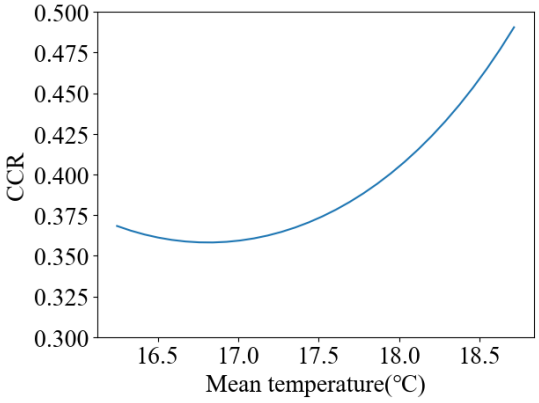
(a)



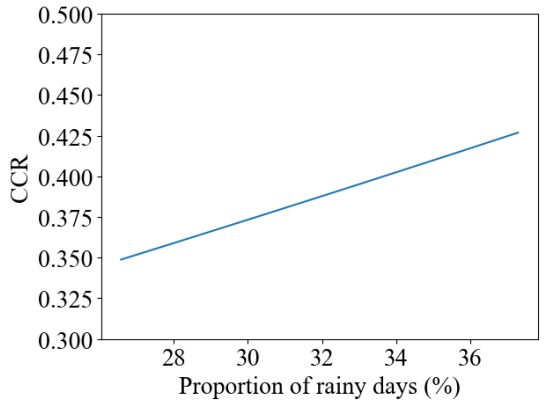
(b)



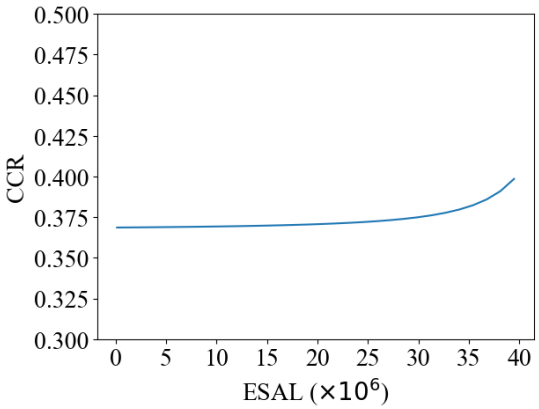
(c)



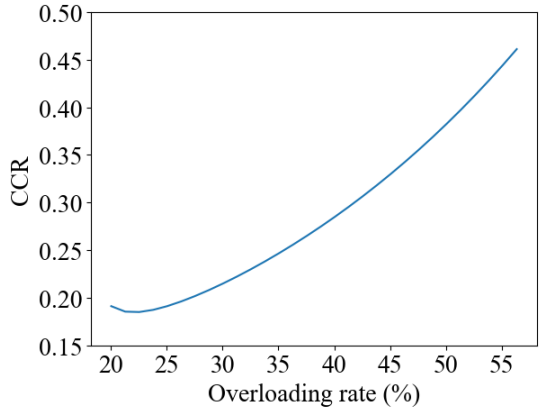
(d)



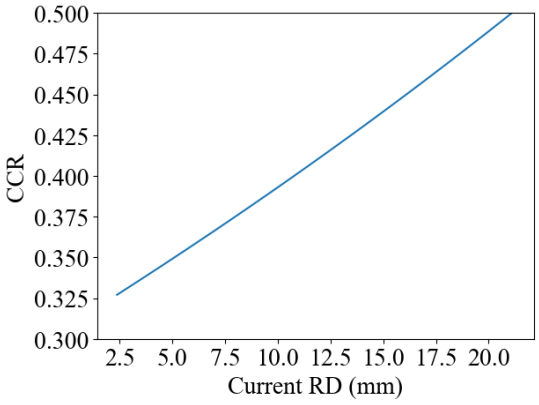
(e)



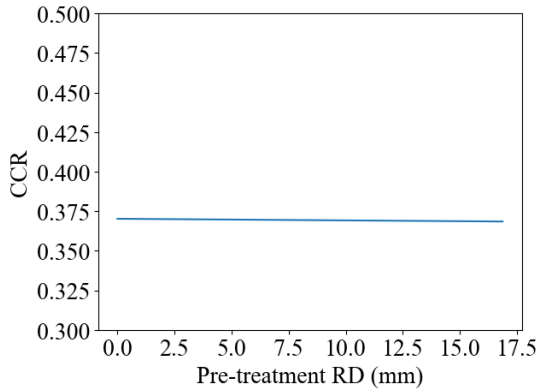
(f)



(g)



(h)



(i)

Figure 13 Results of parametric analysis. (a) Aggregate gradation, (b) Softening point, (c) Binder aggregate ratio, (d) Mean temperature, (e) Proportion of rainy days, (f) ESAL, (g) Overloading rate, (h) Current RD and (i) Pre-treatment RD.

Moreover, to futher explore the impact of the interaction between softening point and mean temperature on the model output, a variable effect characteristic (VEC) surface[42] was drawn. As shown in Figure 14, the combination of low softening point and high mean temperature accounts for a high CCR value. The GEP model becomes more sensitive to the softening point as the mean temperature goes up. Similarly, the GEP model becomes more sensitive to the mean temperature as the softening point goes down. It indicates that binders with a high softening point must be used in relatively hot areas, while this is less important in areas where the mean temperature is lower than 17.5 °C.

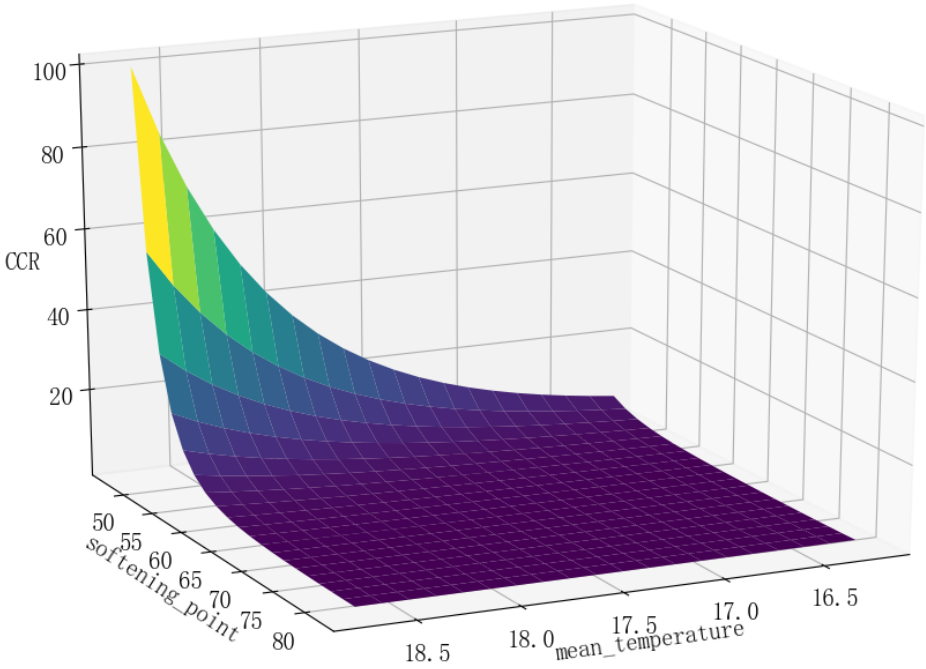


Figure 14 The VEC surface of the input pair: softening point and mean temperature.

6 Uncertainty Analysis

The CCR model shown in Eq. (2) was deterministic, which means all input variables have deterministic values despite the varying levels of uncertainty embedded in them. However, some input variables utilized by the GEP model have inherent variability due to the uncertainty in the construction, testing, and deterioration process of pavement. Furthermore, although the GEP model will faithfully emulate the exact value of CCR, model error is inevitable.

6.1 Monte Carlo Simulation

To investigate the effect of the combined uncertainties on the model output, Monte Carlo simulation (MCs) was performed in this study to propagate different sources of uncertainty into the CCR prediction. Figure 15 shows the schematic diagram of MCs. After the uncertainties have been characterized by quantifying their variances and distributions, a probability density function (PDF) was constructed for each random variable. MCs was then executed to randomly sample data from the predefined distributions into the GEP model and to compute the corresponding CCR values. By repeating this process, the probability distributions of the predicted CCR could be estimated. More details on the parameters and assumptions of the distribution of each

random variable are introduced in the following section.

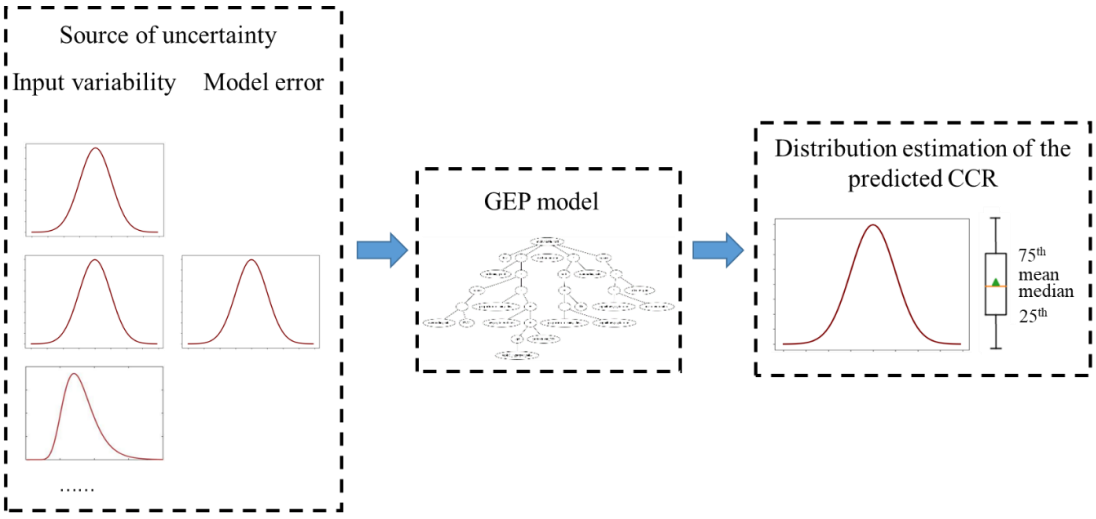


Figure 15 Schematic diagram of Monte Carlo simulation.

6.2 Parameters and Assumptions

The aggregate gradation of middle layer represented by its name is constant and fixed in the same road section. However, the rutting depth, softening point, penetration, and binder aggregate ratio are all random variables and vary from core to core. Besides, the real-time data of overloading rate, mean temperature and proportion of rainy days are difficult to obtain. Some areas even have no such data at all. As a result, data collected in earlier years or other areas have to be used to replace or estimate the true value and hence cause variability in the input variables.

Numerous historical data stored in PMS are capable of providing statistical distribution properties regarding some input variables, e.g., rutting depth. In this research, most of the field core locations were roughly documented in a one-kilometer interval, whereas the original RD data were detected every 10 meters. Hence, there were about 100 data samples in each coring section that can be used to investigate the distribution and coefficient of variation (COV) of the RD data. Kolmogorov-Smirnov test with 95% confidence was employed to determine whether the data follow a specific distribution. The result shows it is statistically acceptable to consider RD data to follow a lognormal distribution and normal distribution in 96.9% and 87.5% of the sections, respectively, as seen in Table 5. So, it is preferable to sample RD from a lognormal distribution, which is in agreement with the results of Sivilevičius and Vansauskas [43]. Thus, both current RD and pre-treatment RD were considered to

follow a lognormal distribution, with the COV ranging from 10% to 40%.

Table 5 Result of Kolmogorov-Smirnov test

Distribution	Null hypothesis	
	Accept	Reject
Normal	87.5%	12.5%
Lognormal	96.9%	3.1%

The variability of the overloading rate could be estimated from the axle load spectra of the same section in different years. However, considering the axle load spectra of the whole service period would overestimate its variability. This may be ascribed to the significant decline of the overloading rate in recent years due to the effective control measures adopted by the transport department. Thus, only the traffic data in the last five years were included here. Haan et al. [44] indicated that reasonable estimates of the mean and variance of the input variable are more important than the form of distribution. Since five data samples were not enough to construct a specific distribution, the overloading rate is therefore assumed to be normally distributed with the COV varying from 20% to 75% according to the statistical results. Similarly, it was also decided to assume a normal distribution for ESAL, mean temperature and proportion of rainy days. As ESAL increase over time, annual equivalent single-axle loads(AESAL) data instead of ESAL data in the past five years, were collected to determine the COV range of ESAL.

For other variables, however, these parameters are unknown. Therefore, a normal distribution is simply assumed for them, and the uncertainty of these variables in terms of COV were decided according to available literature and expert opinions. Research has shown that the traditional asphalt binder property indexes, such as penetration, softening point and binder aggregate ratio (asphalt binder content), have substantially lower variability than the asphalt mixture property indexes, with the value of COV less than 10% [45][46]. Accordingly, COV ranges for softening point and binder aggregate ratio were set to be 0%-10%. The overall results are outlined in Table 6.

Table 6 Probability distributions of random variables

Random variable	Probability distribution	COV range
Softening point	normal	0%~10%
Binder aggregate ratio	normal	0%~10%

Mean temperature	normal	5%~15%
Proportion of rainy days	normal	0%~25%
ESAL	normal	5%~65%
Overloading rate	normal	20%~75%
Current RD	lognormal	10%~40%
Pre-treatment RD	lognormal	10%~40%

The second significant source of uncertainty derives from the GEP model error. This uncertainty was quantified by the residuals calculated as the difference in prediction values and measured values. Figure 16 shows the distribution of the residuals, on which the Kolmogorov-Smirnov test was used for testing normality. It can be observed from Figure 16 that the p-value (0.7431) is much higher than 0.05, indicating that the normal distribution was not rejected at the 95% significance level. Thus, the uncertainty incurred by the model error is considered to be normally distributed with a mean and standard deviation calculated based on model residuals.

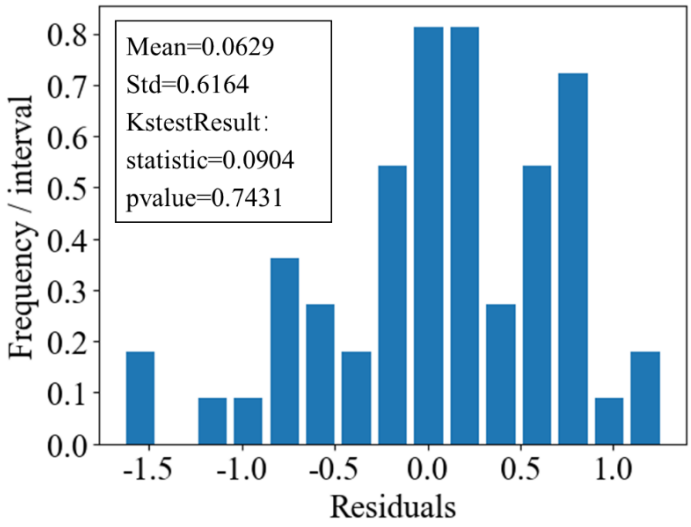


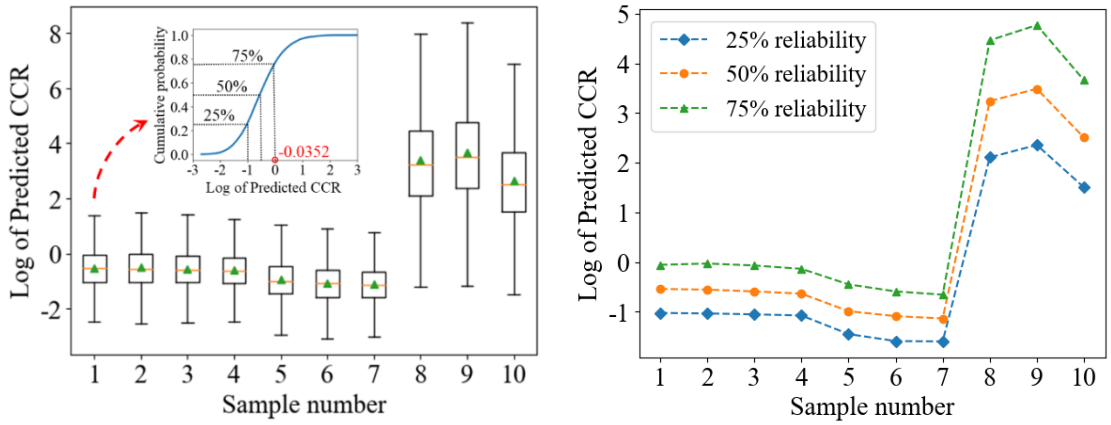
Figure 16 Distribution of the residuals.

6.3 Results of Uncertainty Analysis

A COV of 5% was first assigned to all the random variables to display how the uncertainties were propagated to the model outputs by performing MCs. A distribution rather than a single value of the predicted CCR of the testing samples were obtained, as presented in Figure 17a, where the subplot shows the cumulative frequency distribution (CDF) for sample one. The CDF can be used to interpret the predicted

CCR values for a target reliability level [47]. For example, the predicted CCR of sample one at 75% reliability is about 0.9654 (equal to $e^{-0.0352}$). That means 75 of 100 predictions for this sample will exhibit CCR less than 0.9654, i.e., there is a 75% chance that the predicted CCR will be less than 0.9654.

Figure 17b shows the predicted CCR values of the testing samples at a level of 75%, 50%, and 25% reliability, respectively. As seen, the predicting results vary significantly at different reliability levels. For the middle layer materials to be at least 75% reliable in terms of its high-temperature performance, the green dashed curve should not exceed the threshold of CCR. The reliability level must be selected based on the annual budget, as a high-reliability value may sharply increase the cost for middle layer rehabilitation. It can thus be concluded that disregard of the uncertainties will result in a prediction or evaluation that does not perform to the required level of reliability.



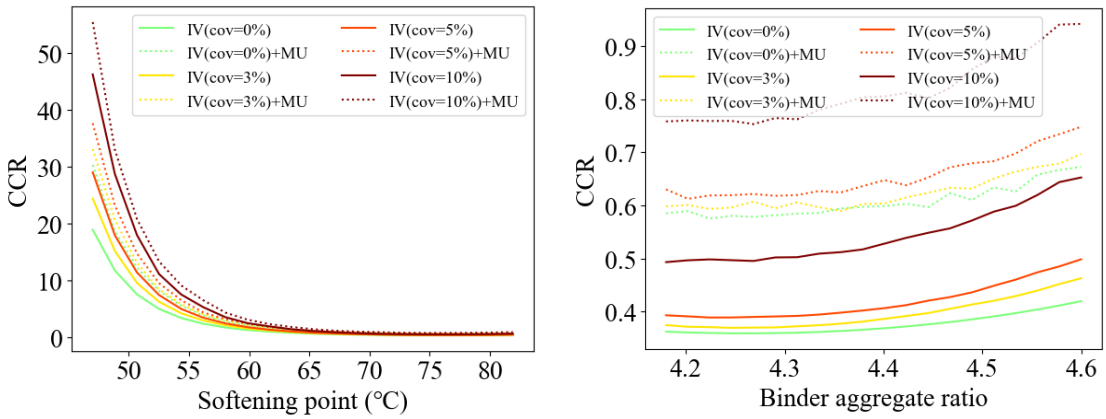
(a) (b)

Figure 17 Results of uncertainty analysis for testing samples(COV=5% for all random variables). (a) Distribution estimation of the log of predicted CCR and (b) Log of predicted CCR at different reliability levels.

Moreover, different sources of uncertainty may affect the predicting results to varying degrees. Hence, it is of great importance to gain some insight into how every single uncertain factor works and the relative contributions of these uncertainties. This will allow engineers to understand whether the uncertainty is reducible through stricter quality control or refining the prediction model. To this end, four levels of COV derived from Table 6 were considered for each random variable, among which COV=0% represents the deterministic results. Only one input variable was set to be uncertain with a specified COV value and a mean value that is taken within the range in Table 2, whereas other variables are equal to their median values. The contributions

of input variability and model uncertainty to the overall variance were also compared.

Figure 18 shows the result, where the x-axis represents the mean value and the y-axis is the prediction at 75% reliability level. It is observed from Figure 18a that the CCR values at different COV levels is significantly different for softening point less than 55 °C. Comparing the CCR values corresponding to the green and yellow solid lines shows that even a low COV would cause great variance in CCR prediction. The almost overlapping of the solid red line and the dotted green line in Figure 18a implies that the variations resulted from input variability and model uncertainty are comparable when the COV is lower than 5%, and afterward, input variability contributes more. From Figure 18b, d, e, g and h (binder aggregate ratio, proportion of rainy days, ESAL, current RD and pre-treatment RD) where all the dotted lines are above the solid lines, it can be concluded that the influence of model uncertainty is more significant than the input variability. However, the much smaller y-axis values in these figures indicate that the degree of impact is limited. Moreover, the partial overloping of the solid lines and dotted lines in Figure 18e and g suggest that when ESAL is less than about 1.8×10^7 or current RD is lower than 7.5 mm or so, the uncertainty of these variables is almost negligible. From Figure 18c, the relatively higher position of the green dotted line than the red solid line indicates that model uncertainty has more contributions when the COV of mean temperature is less than 10%. Concerning the overloading rate, the relative position between the green dotted line and the yellow, red and dark red solid lines signifies that the contributions of model uncertainty are greater, comparable and smaller than that of input variability, corresponding to the 20%, 50%, and 75% COV, respectively. Finally, it is clear that the effect of variability in mean temperature, ESAL, overloading rate and current RD is more significant at their respective higher values as the distance between solid lines in different colors is increasing in Figure 18c, e, f and g. In summary, the proposed model is robust provided that the variability of softening point is strictly controlled.



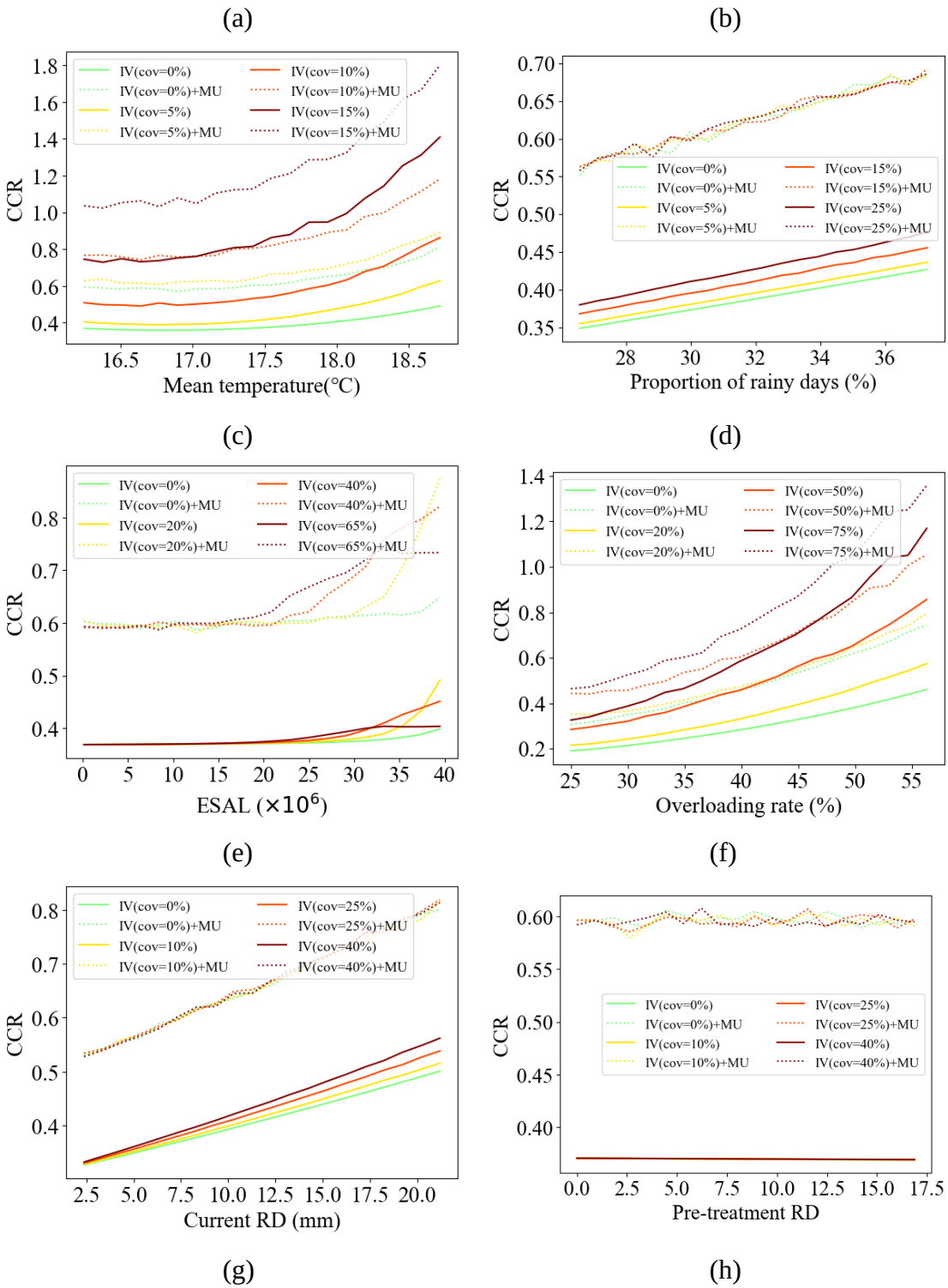


Figure 18 Effect of model uncertainty and input variability on the predicted results at 75% reliability level (IV=input variability; MU=model uncertainty). (a) Softening point, (b) Binder aggregate ratio, (c) Mean temperature, (d) Proportion of rainy days, (e) ESAL, (f) Overloading rate, (g) Current RD and (h)Pre-treatment RD.

7 Conclusions

In this study, a machine learning method named GEP was utilized to predict the rutting resistance of the middle asphalt layer after long-term field service, and compared with the traditional MLR method. Different sources of uncertainty were propagated into the GEP model through Monte Carlo simulation to investigate their effects on the predicting results and enhance the reliability of the prediction. Based on the outcome of this study, the following conclusions have been obtained:

- The developed GEP model was capable of predicting the CCR of middle layer asphalt mixtures with high accuracy. It can be used as an alternative to the field core test and can generate a continuous profile of the high temperature performance of the in-service asphalt mixture. Compared with the MLR model, the GEP model shows better prediction performance. The GEP model satisfies all the criteria for external validation, while the MLR model failed one criterion.
- The input variables selected in the optimal GEP model included aggregate gradation, softening point, binder aggregate ratio, mean temperature, proportion of rainy days, ESAL, overloading rate, current RD and pre-treatment RD. Among them, softening point, binder aggregate ratio and mean temperature were found to have the most significant contribution to the GEP model, followed by overloading rate and proportion of rainy days. CCR was most sensitive to softening point. The material properties of the asphalt binder dominate the rutting resistance of the in-service asphalt mixtures. The interaction between softening point and mean temperature was the strongest. The combination of low softening point and high mean temperature led to a high CCR value.
- The parametric analysis demonstrated that the predicted target values were in close agreement with the underlying physical behavior of the investigated problem. CCR decreased drastically with the increase of softening point. Positive correlations were found between the predicted CCR and binder aggregate ratio, mean temperature, proportion of rainy days and current RD. Extreme heavy loads or high overloading rate would greatly increase the rutting potential of the middle asphalt layer.
- The implementation of uncertainty analysis enables the prediction of CCR for a target reliability level. Transport agencies are then allowed to make their M&R plans according to the specified reliability target. The influence of

model uncertainty is more significant compared with that of input variability in binder aggregate ratio, proportion of rainy days, ESAL, current RD and pre-treatment RD. However, for other variables, the contributions of input variability and model uncertainty to the overall variance could be either different or comparable, depending on the COV levels. Among all the uncertainties, the variability in the softening point should be strictly controlled, especially when the softening point is less than 55 ºC.

Despite the contributions this study has made, there are still opportunities to further improve this research. First, the estblished GEP model was found to have some difficulties in distinguishing smaller CCR values. This can probabily be solved by improving the data quality of smaller CCR value. Second, the sample size of CCR data is still limited which implies the needs to accumulate more data samples. Meawhile, it should be mentioned that the accumulation of test data for in-service asphalt mixtures will never stop. By expanding the experimental database and allowing the model to continue learning, the model can be further improved. This study just prove the feasibility and construct the framework of modelling in-service material properties of asphalt pavement. Further research could also include other material property indicators and asphalt layers.

Acknowledgments

This study was conducted under the support of the Research Institute for Sustainable Urban Development (RISUD) at the Hong Kong Polytechnic University. In addition, the data used in this research were collected from the Pavement Management System in Jiangsu province, China. The engineers and professors who established the system and collected the data are also acknowledged for their contribution.

References

[1] O.E. Gungor, I.L. Al-Qadi, All for one: Centralized optimization of truck platoons to improve roadway infrastructure sustainability, Transportation Research Part C: Emerging Technologies 114 (2020) 84-98.

[2] C.Y. Chan, B. Huang, X. Yan, S. Richards, Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS), Journal of advanced transportation 44(3) (2010) 150-161.

[3] X. Chen, Q. Dong, H. Zhu, B. Huang, Development of distress condition index of asphalt pavements using LTPP data through structural equation modeling, Transportation Research Part C: Emerging Technologies 68 (2016) 58-69.
[4] G. Zhou, L. Wang, Co-location decision tree for enhancing decision-making of pavement maintenance and rehabilitation, Transportation Research Part C: Emerging Technologies 21(1) (2012) 287-305.
[5] A.G. Karlaftis, A. Badr, Predicting asphalt pavement crack initiation following rehabilitation treatments, Transportation Research Part C: Emerging Technologies 55 (2015) 510-517.
[6] Q. Li, H. Yang, F. Ni, X. Ma, L. Luo, Cause analysis on permanent deformation for asphalt pavements using field cores, Construction and Building Materials 100 (2015) 40-51.
[7] J. Ling, F. Wei, H. Chen, H. Zhao, Y. Tian, B. Han, Accelerated Pavement Testing for Rutting Evaluation of Hot-Mix Asphalt Overlay under High Tire Pressure, Journal of Transportation Engineering, Part B: Pavements 146(2) (2020) 04020009.
[8] N. Dong, F. Ni, L. Zhou, X. Ma, Comparison of the Hamburg, indirect tensile, and multi-sequenced repeated load tests for evaluation of HMA rutting resistance, Construction and Building Materials 216 (2019) 588-598.
[9] Q. Li, F. Ni, L. Gao, Q. Yuan, Y. Xiao, Evaluating the rutting resistance of asphalt mixtures using an advanced repeated load permanent deformation test under field conditions, Construction and Building Materials 61 (2014) 241-251.
[10] L. Gao, F. Ni, S. Charmot, Q. Li, High-temperature performance of multilayer pavement with cold in-place recycling mixtures, Road materials and pavement design 15(4) (2014) 804-819.
[11] S. Ghaffarpour Jahromi, Introduction of a simple method for prediction of the rutting resistance factor of nanoclay-modified bitumen, International Journal of Pavement Engineering 20(2) (2019) 216-221.
[12] M.R. Mirzahosseini, A. Aghaeifar, A.H. Alavi, A.H. Gandomi, R. Seyednour, Permanent deformation analysis of asphalt mixtures using soft computing techniques, Expert Systems with Applications 38(5) (2011) 6081-6100.
[13] A.H. Gandomi, S. Sajedi, B. Kiani, Q. Huang, Genetic programming for experimental big data mining: A case study on concrete creep formulation, Automation in Construction 70 (2016) 89-97.
[14] M. Elhenawy, H. Chen, H.A. Rakha, Dynamic travel time prediction using data

clustering and genetic programming, Transportation Research Part C: Emerging Technologies 42 (2014) 82-98.
[15] S. Terzi, Modeling the deflection basin of flexible highway pavements by gene expression programming, Journal of Applied Sciences 5(2) (2005) 309-314.
[16] M. Saltan, S. Terzi, Comparative analysis of using artificial neural networks (ANN) and gene expression programming (GEP) in backcalculation of pavement layer thickness, (2005).
[17] A.H. Gandomi, A.H. Alavi, M.R. Mirzahosseini, F.M. Nejad, Nonlinear genetic-based models for prediction of flow number of asphalt mixtures, Journal of Materials in Civil Engineering 23(3) (2011) 248-263.
[18] J. Liu, K. Yan, L. You, P. Liu, K. Yan, Prediction models of mixtures' dynamic modulus using gene expression programming, International Journal of Pavement Engineering 18(11) (2017) 971-980.
[19] H. Majidifard, B. Jahangiri, W.G. Buttlar, A.H. Alavi, New machine learning-based prediction models for fracture energy of asphalt mixtures, Measurement 135 (2019) 438-451.
[20] A.R. Tenpe, A. Patel, Application of genetic expression programming and artificial neural network for prediction of CBR, Road Materials and Pavement Design 21(5) (2020) 1183-1200.
[21] L.P. Leon, D. Gay, Gene expression programming for evaluation of aggregate angularity effects on permanent deformation of asphalt mixtures, Construction and Building Materials 211 (2019) 470-478.
[22] [1] A. Eleyedath, S.S. Kar, A.K. Swamy, Modelling of expansion ratio and half-life of foamed bitumen using gene expression programming, International Journal of Pavement Engineering 22(3) (2021) 369-381.
[23] C. Ferreira, Gene expression programming: a new adaptive algorithm for solving problems, arXiv preprint cs/0102027 (2001).
[24] J. Jiang, F. Ni, L. Gao, S. Lou, Developing an optional multiple repeated load test to evaluate permanent deformation of asphalt mixtures based on axle load spectrum, Construction and Building Materials 122 (2016) 254-263.
[25] N. Dong, F. Ni, S. Li, J. Jiang, Z. Zhao, Characterization of permanent deformation performance of asphalt mixture by multi-sequenced repeated load test, Construction and Building Materials 180 (2018) 425-436.
[26] Z. Zhao, J. Jiang, F. Ni, Q. Dong, J. Ding, X. Ma, Factors affecting the rutting resistance of asphalt pavement based on the field cores using multi-sequenced

repeated loading test, Construction and Building Materials 253 (2020) 118902.
[27] F. Moghadas Nejad, A. Azarhoosh, G.H. Hamedi, H. Roshani, Rutting performance prediction of warm mix asphalt containing reclaimed asphalt pavements, Road Materials and Pavement Design 15(1) (2014) 207-219.
[28] M.R. Mohd Hasan, J.E. Hiller, Z. You, Effects of mean annual temperature and mean annual precipitation on the performance of flexible pavement using ME design, International Journal of Pavement Engineering 17(7) (2016) 647-658.
[29] J.T. Smith, S.L. Tighe, J. Andrey, B. Mills, Temperature and precipitation sensitivity analysis on pavement performance, Surface Transportation Weather and Snow Removal and Ice Control Technology (2008) 558.
[30] A.L. Simpson, J.F. Daleiden, W.O. Hadley, Rutting analysis from a different perspective, Transportation research record 1473 (1995) 9.
[31] A.R. Archilla, S. Madanat, Development of a pavement rutting model from experimental data, Journal of transportation engineering 126(4) (2000) 291-299.
[32] G. Rusbintardjo, The Influence of Overloading Truck to the Road Condition, Proceedings of the Eastern Asia Society for Transportation Studies, 2013.
[33] H. Zhang, F. Ni, Study of Expressway Axle Load Spectrum Based on Toll Data of Jinghu Expressway, Journal of Testing and Evaluation 40(7) (2012) 1220-1227.
[34] I.E. Frank, R. Todeschini, The data analysis handbook, Elsevier1994.
[35] G.N. Smith, Probability and statistics in civil engineering, Collins professional and technical books 244 (1986).
[36] A. Golbraikh, A. Tropsha, Beware of q2!, Journal of molecular graphics and modelling 20(4) (2002) 269-276.
[37] P.P. Roy, K. Roy, On some aspects of variable selection for partial least squares regression models, QSAR & Combinatorial Science 27(3) (2008) 302-313.
[38] A.H. Gandomi, G.J. Yun, A.H. Alavi, An evolutionary approach for modeling of shear strength of RC deep beams, Materials and Structures 46(12) (2013) 2109-2119.
[39] L. Gao, Z. Wang, C. Deng, J. Shi, Analysis on Effect Factors of Rutting Performance, 2009, pp. 3772-3778.
[40] X. Huang, H. Li, J. Zhang, Simulation of Rutting Behavior of Asphalt Pavement Based on Real Temperature Field (2008) No. 08-0372.
[41] M.J. Embrechts, F.A. Arciniegas, M. Ozdemir, R.H. Kewley, Data Mining for Molecules with 2-D Neural Network Sensitivity Analysis, International journal of smart engineering system design 5(4) (2003) 225-239.

[42] P. Cortez, M.J. Embrechts, Using sensitivity analysis and visualization techniques to open black box data mining models, Information Sciences 225 (2013) 1-17.
[43] H. Sivilevičius, V. Vansauskas, Research and evaluation of ruts in the asphalt pavement on Lithuanian highways, Journal of civil engineering and management 19(5) (2013) 609-621.
[44] D.R. Edwards, C.T. Haan, S. Prabhu, D.E. Storm, T. Al-Issa, G.J. Sabbagh, Effect of Parameter Distributions on Uncertainty Analysis of Hydrologic Models, (1998).
[45] A.E. Alvarez, L.V. Espinosa, S. Caro, E.J. Rueda, J.P. Aguiar, L.G. Loria, Differences in asphalt binder variability quantified through traditional and advanced laboratory testing, Construction & building materials 176 (2018) 500-508.
[46] C. Rodezno, R. Brown, Improving accuracy of asphalt content determination by ignition test, Road materials and pavement design 18(sup4) (2017) 112-127.
[47] E. Coleri, J.T. Harvey, Evaluation of Laboratory, Construction, and Performance Variability by Bootstrapping and Monte Carlo Methods for Rutting Performance Prediction of Heavy Vehicle Simulator Test Sections, Journal of transportation engineering 137(12) (2011) 897-906.