

Hybrid Nonlinear and Machine Learning Methods for Analyzing Factors Influencing the Performance of Large-Scale Transport Infrastructure

Yongze Song^{ID}, Member, IEEE, Peng Wu^{ID}, Qindong Li, Yuchen Liu, and Lalinda Karunaratne

Abstract—Strategic maintenance is essential for sustainable road infrastructure development. Accurate estimation of road maintenance effects can support the assessment of maintenance strategies and reasonable allocation of budgets and resources. Road deterioration is affected by sophisticated factors, but accurate investigation of the integrated deterioration factors is limited. This study developed a dynamic trade-off model (DTOM), a hybrid nonlinear and machine learning method, for quantifying temporally varied impacts of factors and examining maintenance effects at the network level. Pavement deterioration factors are classified into three categories: (i) historical observations of roughness, (ii) pavement age, and (iii) traffic, climate and environment factors. Their respective impacts on pavements are estimated using a non-linear least square regression, a joinpoint regression and a random forest model, respectively. Vehicle-based laser scanner monitored high-resolution deterioration data was collected for a large spatial scale road network in Western Australia from 2007 to 2018. Results show that the resurfacing and rehabilitation are essential for strategic reduction of deterioration. Twelve-year maintenance activities reduced the distress of roughness by 7.5% and increased road performance (the percentage of roads with roughness lower than 2.085 IRI) by 14.5% for the whole road network. The DTOM has great potentials in accurately assessing infrastructure maintenance effects and predicting deterioration scenarios.

Index Terms—Road infrastructure, smart transportation infrastructure, intelligent transportation systems, spatial analysis, spatial big data, sustainable infrastructure development.

I. INTRODUCTION

ROAD infrastructure provides a basic foundation for socio-economic development of the whole society by supporting passenger and freight transportations [1]–[4]. Effective road maintenance strategies are required to ensure an acceptable level of service provided by roads to satisfy the

needs of different road users [5]–[7]. Strategic road maintenance efforts aim to develop comprehensive solutions, based on predictions that address current and future deterioration and defects for targeted roads, to ensure the sustainability of road infrastructure [8], [9]. Strategic maintenance efforts have been applied by national and state/territory road agencies to effectively manage road assets over many past decades [10].

The effects of road maintenance are generally assessed by three criteria. First, from the perspective of road asset management, maintenance cost is used as a direct indicator to estimate the activities and quality requirements of road maintenance. For instance, the cost of road maintenance is used to evaluate the ability of roads to adapt to the challenge of climate change in the US [11], [12], and is used to compare different models of road maintenance services in Sweden [13]. In addition, sensor monitoring indicators of road surface conditions are also commonly used to quantify road quality and maintenance effects. The wide applications of smart sensor observations of road performance provide opportunities for quantitative and accurate understanding of road deterioration. For instance, road roughness (rideability) is one of the most common indicators for capturing riding quality and it is measured by vehicles travelling at highway speed using laser profilometer technology [14], [15]. Images from videos and unmanned aerial vehicle (UAV) mounted cameras are also important data sources of dynamic road surface conditions and the evidence of maintenance effects [16], [17]. Finally, effects of road maintenance are estimated by the satisfaction in meeting road users' requirements, such as riding comfort, population accessibility to public facilities [5], [18] and road safety [19], [20].

Road deterioration is affected by a combination of sophisticated factors [21]. The factors can be classified into four categories according to their paths affecting pavement and determining current deterioration. First, current and future deterioration is closely associated with historical conditions. When the time range between the historical conditions and the current and future deterioration is enlarged, the associations between them will probably be decreased [17], [22]. Next, road deterioration usually can be predicted as an increased non-linear function of the pavement age, which is defined as the time range from the paved or rehabilitation year to the observation year [23]–[25]. Third, road deterioration is also directly affected by loads from passengers and freight transport, and climate and environmental conditions. The

Manuscript received 27 September 2020; revised 21 February 2021 and 3 July 2021; accepted 27 August 2021. Date of publication 27 September 2021; date of current version 9 August 2022. This work was supported in part by the Australian Government through the Australian Research Council's Discovery Project under Project DP180104026 and in part by the Main Roads Western Australia. The Associate Editor for this article was A. Nunez. (Corresponding author: Peng Wu.)

Yongze Song and Peng Wu are with the School of Design and the Built Environment, Curtin University, Perth, WA 6102, Australia (e-mail: yongze.song@curtin.edu.au; peng.wu@curtin.edu.au).

Qindong Li and Lalinda Karunaratne are with the Main Roads Western Australia, Perth, WA 6004, Australia (e-mail: qindong.li@mainroads.wa.gov.au; lalinda.karunaratne@mainroads.wa.gov.au).

Yuchen Liu is with the School of Earth and Planetary Sciences, Curtin University, Perth, WA 6102, Australia (e-mail: yuchen.liu@curtin.edu.au).

Digital Object Identifier 10.1109/TITS.2021.3112458

impact of loads on road deterioration can be quantified with proxy variables of traffic volumes, heavy vehicle transportation, and estimations of total vehicle masses [26]–[29]. The potential climate and environmental factors include temperature, rainfall, soil moisture, surface evaporation and vegetation coverage [30]–[34]. Finally, from the perspective of pavement engineering and management, road deterioration is also associated with pavement related variables, such as pavement type, material, depth, and techniques [35], [36]. Experimental and mathematical studies also reveal the critical impacts of pavement structural parameters, such as surface and base layers, and rigid, semi-rigid, and flexible pavements, on roughness performance of pavements and maintenance work [37]–[41]. For instance, pavements with surface of sprayed seal, asphalt concrete (flexible and semi-rigid pavements), and Portland cement concrete (rigid pavements) have distinct deterioration modes [42]–[44]. Materials and thickness of base and sub-base layers of pavements, such as semi-rigid pavements, also have essential impacts on road surface roughness [45]–[47].

The impacts of each category of deterioration factors have been investigated in previous studies, but a comprehensive assessment of the integrated impacts of the factors on deterioration and estimation of maintenance effects is still required. A critical issue to modelling the different categories of effects is that the effects have distinct characteristics and relationships with road deterioration. Historical records of road deterioration indicate the start points of future variations of deterioration, pavement age contains information of the relationship between designed pavement age and actual age, and factors demonstrate the integrated traffic, climate and environment conditions during road usage. Their respective impacts on road deterioration are also varied over time. Therefore, it is essential to model the effects respectively and compare them to reveal the temporally varied influence of the effects.

This study develops a dynamic trade-off model (DTOM), a hybrid nonlinear and machine learning method, for quantifying the temporally varied impacts of deterioration factors and examining the effects of maintenance at the network-level. Pavement deterioration factors are classified into three categories, including historical observations of roughness, pavement age, and traffic, climate and environment factors. Their respective impacts on the pavement are defined as the baseline effect, pavement age effect and factors effect, which are estimated using a non-linear least square (NLS) regression, a joinpoint regression and a random forest model (RFM), respectively. Then, temporally varied contributions of the three types of effects are computed to construct a dynamic prediction model. Finally, effects of different types of maintenance activities are assessed, and future deterioration scenarios are predicted.

In this study, the DTOM is applied in investigating the effects of strategic maintenance activities for the large-scale road network in Western Australia (WA), which represents one of the largest regional road networks globally. Vehicle-mounted laser profilometers monitored high-resolution pavement condition data and corresponding traffic and environment factors were collected during five maintenance periods from 2007 to 2018. The effects of resurfacing and

rehabilitation are estimated, and future deterioration scenarios from 2007 to 2030 are predicted.

II. STUDY AREA AND DATA

A. Study Area

The arterial road network in WA, Australia, with the total length of 18,500 km, links cities and residential areas, major ports, and primary industry regions, including mining sites, oil and gas plants, and agriculture regions [48]. In the road network, the majority of pavement surface type is sprayed seal (91.3%), followed by asphalt (7.3%) and concrete (0.2%). The percentages of pavement surface types in WA are generally consistent with that in the whole Australian road network, where sprayed seals account for about 90% of the surfaced roads [49]. In Australia, commonly used materials for sprayed seals include bituminous materials and aggregates, and the primary advantages of sprayed seals are low cost, speed in construction, and safe for all rural roads [49]. In the past decades, Main Roads WA (MRWA), the state road agency, has developed various approaches to optimize maintenance strategies and improve road performance to satisfy road user needs [50]–[52]. Up until now, both innovative knowledge and practical experience regarding road performance, deterioration factors, and maintenance strategies have been accumulated in managing the WA arterial road network.

B. Road Segment-Based Deterioration

In this study, long-term and large spatial scale road roughness data has been collected for investigating the road deterioration and maintenance effects in WA. The road roughness has been collected using heavy vehicle-mounted laser profilometers that collect point clouds of road surface and report the roughness every ten metres (m) along roads using IRI (International Roughness Index), a calibrated and standardized roughness measurement [45], [46]. Then the roughness data with its 10 m resolution is upscaled to the road segment-based data with resolutions of 100–500 m using a spatial heterogeneity-based segmentation (SHS) model for effective management of road assets from both project and network levels [53]. The SHS model is an accurate homogeneous segmentation method developed based on spatial heterogeneity for defining road segments using spatially continuous observations of road attributes [53]. In this study, roads are segmented using roughness data.

The roughness of roads and maintenance records in the 2007–2018 period were collected in 2007, 2009, 2012, 2014, 2016 and 2018, and therefore, five periods were established as shown in Table I. The maintenance records consist of no maintenance actions, resurfacing and rehabilitation. Primary types of resurfacing and rehabilitation works are summarized in Table II. In general, resurfacing includes sprayed seals, asphalt surfacing, slurry surfacing and microsurfacing [54], and rehabilitation includes structural overlays, mechanical and chemical stabilization, heavy patching and pavement reconstruction [55]. Figure 1 shows maps of road segment-based mean roughness and roughness changes from 2007 to 2018,

TABLE I
A SUMMARY OF OBSERVATIONS OF ROAD SURFACE ROUGHNESS AND MAINTENANCE RECORDS

Year	No. of observations	Percentage of missing data (%)	Period	Percentage of maintained roads		
				No actions	Resurfacing	Rehabilitation
2007	61385	13.15%	/	/	/	/
2009	65460	7.39%	2007-2009	89.41%	7.65%	2.92%
2012	64779	8.35%	2009-2012	84.70%	12.92%	2.36%
2014	67048	5.14%	2012-2014	90.67%	8.22%	1.10%
2016	64877	8.21%	2014-2016	85.81%	11.74%	2.43%
2018	67416	4.62%	2016-2018	93.38%	5.71%	0.89%
/	/	/	2007-2018	44.06%	46.24%	9.70%

TABLE II
TYPES OF ROAD MAINTENANCE ACTIONS

Action	Type	Details
Resurfacing [54]	Sprayed seals	Initial and subsequent seal treatments and special purpose treatments
	Asphalt surfacing	Dense graded asphalt, open graded asphalt and stone mastic asphalt
	Slurry surfacing	Slurry seals and microsurfacing
Rehabilitation [55]	Structural overlays	Granular, asphalt and concrete overlays and inlays
	Stabilization	Mechanical and chemical stabilization of pavement layers and subgrade
	Heavy patching	Carrying out deep (>75 mm) and severely distressed pavement
	Pavement reconstruction	Partial or total reconstruction

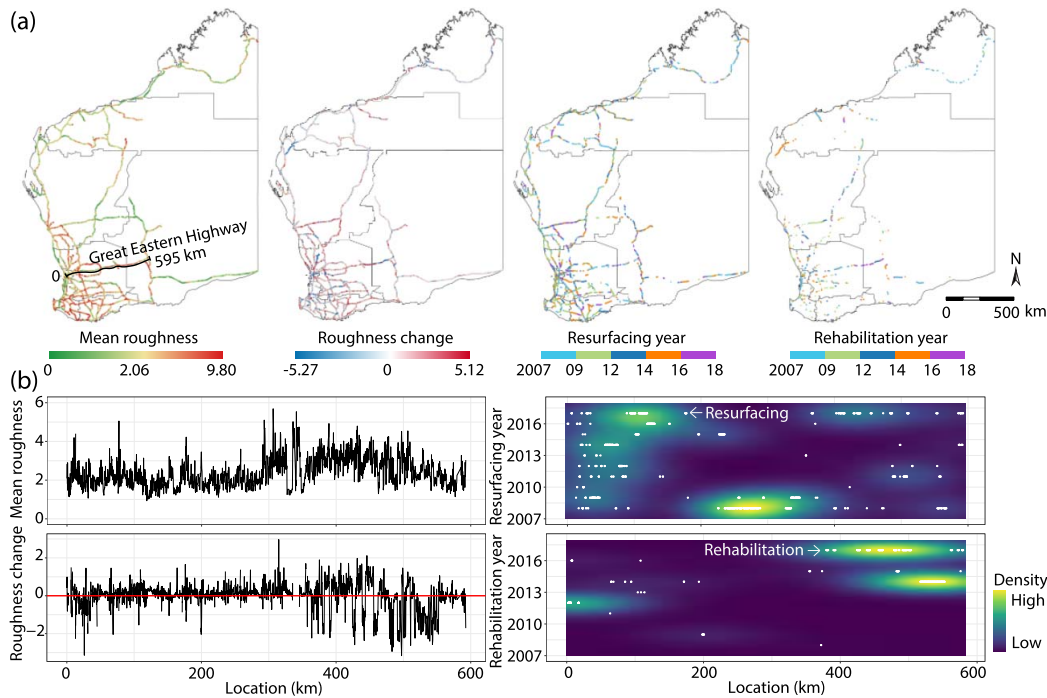


Fig. 1. Distribution of road deterioration and maintenance activities in 2007-2018 in Western Australia (a) and a case data of Great Eastern Highway (b).

and the resurfacing and rehabilitation records during the five periods. The mean roughness in the six years is 2.182 IRI and the mean roughness change during the five periods is -0.0013 IRI, which indicates that the roughness was stable at the network level during the period. Table I lists the statistical summary of roughness observations and maintenance records. On average, the annual number of road segment-based observations is 65,161, and the percentage of missing data is about 7.81%. During the five periods, 46.2% of the roads were resurfaced, 9.7% of the roads were rehabilitated, and no resurfacing or rehabilitation were recorded for 44.1% of the

roads. On average during each period, 9.25% of the roads were resurfaced and 1.94% of the roads were rehabilitated.

C. Explanatory Variables

To predict road deterioration, three categories of explanatory variable data were collected corresponding to the road segment-based roughness data. The general associations between roughness and the explanatory variables are visualized in Figure 2. First, roughness data in the start year of each period were collected and regarded as the baseline to predict the roughness at the end year of the periods. Due to

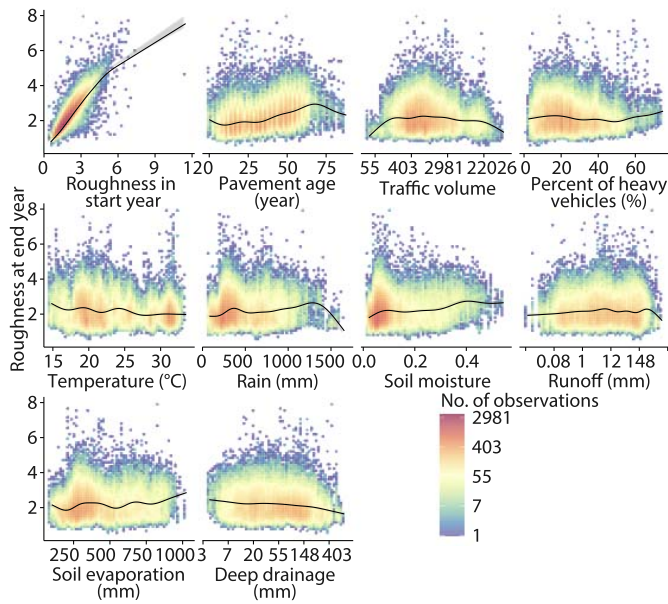


Fig. 2. Explanatory variable data for estimating road maintenance effects and their relationships with the end-year roughness in the five periods.

TABLE III

STATISTICAL SUMMARY OF PAVEMENT AGE OF ALL ROADS IN WESTERN AUSTRALIA FROM 2007 TO 2018

Statistics of pavement age	Year					
	2007	2009	2012	2014	2016	2018
Mean	29.92	31.36	33.53	35.03	35.92	37.58
Standard deviation	15.34	15.79	16.48	16.89	17.76	18.06
Standard error	0.064	0.065	0.067	0.069	0.071	0.072

the inconsistent missing data in the six observation years, only the roughness data with valid observations both in the start year and at the end year were considered for each period. The scatter plot shows that the roughness in the start year and at the end year are closely correlated.

Then, pavement age data were collected to evaluate its association with deterioration. Table III shows a statistical summary of pavement age in the six observation years. The pavement age ranged from 0 to 88 years.

Finally, the traffic, climate and environment deterioration factors were collected. The traffic variables include the traffic volume measured by the annual average daily traffic (AADT) and the percent of heavy vehicles, which were provided by MRWA. According to vehicle categories guide in WA, heavy vehicles are vehicles used for freight transportation with the maximum mass about 42.5–147.5 t and classified into ten categories according to the axle combination, length and mass [56]. The climate variables include annual average temperature and annual rainfall. The temperature data with a 1-km resolution was sourced from the land surface temperature product (MOD11A2) from the Moderate Resolution Imaging Spectroradiometer (MODIS) [57]. The annual rainfall with a 1-km resolution was collected from the Australian Soil Resources Information System (ASRIS) dataset and the Australian Bureau of Meteorology [58]. The local environmental variables include annual mean soil moisture, runoff,

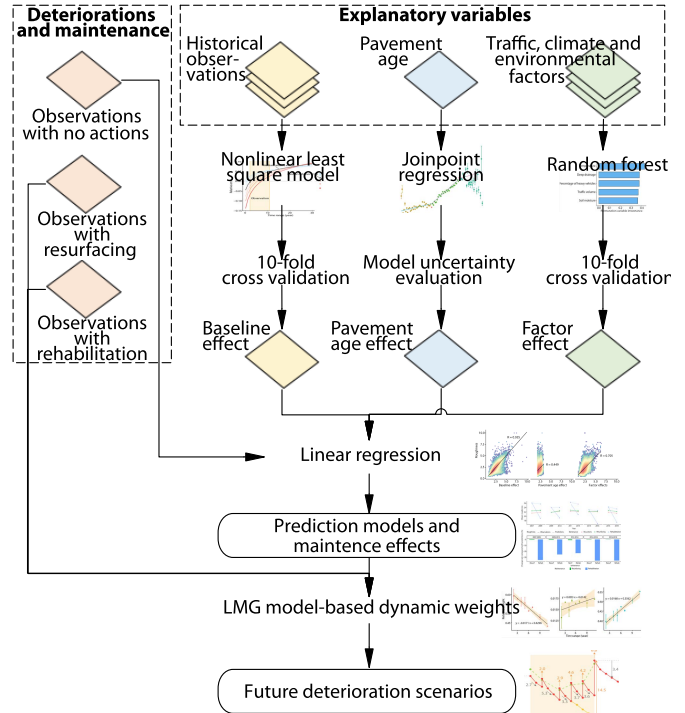


Fig. 3. Schematic overview of the dynamic trade-off model (DTOM) for deterioration prediction and maintenance effects estimation.

soil evaporation and soil deep drainage, which are of 1-km resolution and also sourced from the ASRIS dataset [58]. All the factor data were processed to the road segment-based data with identical spatial units with the roughness of roads.

III. DYNAMIC TRADE-OFF MODEL (DTOM)

This study developed a DTOM, a hybrid nonlinear and machine learning method, for quantifying deterioration factors, examining effects of road maintenance and predicting future deterioration scenarios. The DTOM includes five stages as shown in Figure 3. The deterioration factors were divided into three categories: historical observations during a period, pavement age, and traffic, climate and environmental factors, which are corresponding to the baseline effect, pavement age effect and factors effect of deterioration, respectively. The pavement related variables listed in the introduction section are not included, since structured data of these variables are difficult to be collected.

The first to the third stages aimed to investigate the three types of deterioration effects. The first stage constructed an NLS model to evaluate the relationship between observations at the end year and the start year of a period. The second stage developed a joinpoint regression to characterize the “S”-shape relationship between deterioration and pavement age. In the third stage, an RFM, a well-performed machine learning algorithm, was developed to explore the associations between deterioration and the traffic, climate and environmental factors. In the above three stages, a 10-fold cross validation was performed to evaluate the accuracy of the NLS model and the RFM, and model uncertainty was evaluated for the joinpoint regression. In the fourth stage, the deterioration prediction model was developed and effects of different types

TABLE IV
COMBINATIONS OF PERIODS AND TIME RANGES OF DATA
FOR NLS-BASED BASELINE EFFECT ESTIMATION

No. of period	Start year	End year	Time range
1	2007	2009	2
2	2007	2012	5
3	2007	2014	7
4	2007	2016	9
5	2007	2018	11
6	2009	2012	3
7	2009	2014	5
8	2009	2016	7
9	2009	2018	9
10	2012	2014	2
11	2012	2016	4
12	2012	2018	6
13	2014	2016	2
14	2014	2018	4
15	2016	2018	2

of maintenance activities were assessed. In the fifth stage, the contributions of the three types of deterioration effects were investigated and future deterioration scenarios were predicted.

The methods and processes of the five stages are presented in the following subsections.

A. NLS-Based Baseline Effect Estimation

The relationship between predictions and historical observations during a time period was constructed using an NLS model. In the model, the time range was added to the model to characterize the varied correlations between the roughness in the start year and that at the end year. In the study, data of 15 time period combinations, including eight time ranges, derived from the six-year observations during 2007 – 2018 were used for the NLS-based baseline effect estimation (Table IV). Only the data of roads without maintenance activities during the time periods were used for modelling. As such, the spatiotemporal road deterioration dataset used for the NLS modelling includes 858,630 groups of observations. The NLS model form is:

$$\log(Y_{t_2}) = c_0 + (c_1 \log(Y_{t_1}) + c_2)(\Delta T + 1)^{c_3} + \epsilon \quad (1)$$

where Y_{t_1} and Y_{t_2} are roughness observations at the years t_1 and t_2 , $\Delta T = t_2 - t_1$ is the time range during a period, c_0 , c_1 , c_2 and c_3 are coefficients, and ϵ is the random error.

B. Joinpoint Regression-Based Pavement Age Effect Estimation

To identify changes in the pavement age effect of deterioration, a joinpoint regression was estimated for the five-period deterioration data listed in Table I. The joinpoint regression was constructed as:

$$\log(Y(a)) = \mu(a) + \epsilon(a) \quad (2)$$

where $Y(a)$, $a = 1, \dots, A$ is the proportion of roughness at A different pavement ages; when there are K join points, $\mu(a)$ is a succession of $(K+1)$ linear segments over the entire range of pavement age: $[0, \tau_1], \dots, [\tau_k, \tau_{k+1}], \dots, [\tau_K, 88]$; and $\epsilon(a)$

is the random error for the a th pavement ages. The parameter τ_k is the join point for a statistically significant change in the slopes β_k and β_{k+1} of two neighbor successions of linear segments.

The trends of the pavement age effect of deterioration over the range $[\tau_k, \tau_{k+1}]$ can be computed with the annual percentage change (APC) in roughness between join points. The APC from the pavement age τ_k to τ_{k+1} is computed as:

$$APC_{k+1} = 100(e^{\beta_{k+1}} - 1) \quad (3)$$

The trend over the entire range of pavement age $[0, 88]$ was assessed with the average annual percentage change (AAPC), which is the sum of time range weighted APCs [59]. The AAPC is calculated as:

$$AAPC = 100(e^{\sum_{k=1}^{K+1} w_k \beta_k / \sum_{k=1}^{K+1} w_k} - 1) \quad (4)$$

where w_k is the length of the k th range of pavement age $[\tau_{k-1}, \tau_k]$. The joinpoint regression was performed using the Joinpoint Regression Program, Version 4.8.0.1 [60].

C. Random Forest-Based Factors Effect Estimation

The deterioration factors effect for the five-period spatiotemporal data in Table I was estimated using an RFM, which can be expressed as:

$$Y(s, t) = f(X(s, t)) + \epsilon(s, t) \quad (5)$$

where Y is the roughness at the road segment s and the observation year t , X includes traffic, climate and environmental variables, f is the RFM, and ϵ is the random error. The RFM is an ensemble learning model containing a set of various regression trees that are trained via a bagging approach, that can decrease complexity of models that overfit the training data, and random variable selection [61], [62]. A bootstrap sample of the data is used in the bagging approach to build regression trees. The nodes of trees are split with optimal abundance thresholds from the subsets of randomly selected variables to deal with the potential overfitting issues. A permutation is used to determine the optimal split during the growth of trees and to quantify the importance of variables [63]. Different from deterioration prediction models developed in terms of fixed parameter and equations, the random forest and other machine learning algorithms act a black box in modelling [42]. Due to the amount of observations in the study, the ‘‘ranger’’ R package is used to perform the RFM for its fast computation [64].

In the study, the optimal maximum depth of nodes and number of trees were determined with the comparison of out-of-bag (OOB) errors. First, to select the optimal maximum depth of nodes (MDN), the number of trees was set as 1, optional MDNs were values ranging from 1 to 50, and the OOB mean of squared residuals (MSE_{OOB}) was computed to determine the optimal MDN:

$$MSE_{OOB} = \frac{1}{N} \sum_{i=1}^N (o_i - \bar{o}_i^{OOB})^2 \quad (6)$$

where \bar{o}_i^{OOB} is the mean value of the OOB predictions for the i th observation [62]. The trend of MSE_{OOB} with the

increased MDNs was modelled with a locally estimated scatterplot smoothing (LOESS) function to calculate the change rates of the MSE_{OOB} trend. The point where the change rate is zero is regarded as the optimal MDN. Since the MSE_{OOB} decreased with the increased MDN, the model underfits when the MDN was lower than the point with the optimal value, and overfits when the MDN was higher than the point with the optimal value. Then, the optimal number of trees was calculated using the similar process with the identified optimal MDN and optional numbers of trees ranging from 1 to 50.

D. Model Validations

Models for estimating above three types of effects on road deteriorations have been validated using respective approaches according to the characteristics of models. First, the accuracy of the NLS-based baseline effect estimation model is evaluated using a ten-fold cross validation, where the model accuracy was compared with the linear regression model (LM):

$$\log(Y_{t_2}) = c_0 + c_1 \log(Y_{t_1}) + \epsilon \quad (7)$$

In the ten-fold cross validation, the model accuracy was evaluated with the mean absolute error (MAE), root mean squared error (RMSE) and the coefficient of determination (R^2). The three accuracy evaluation indicators are computed as:

$$MAE = \frac{1}{N} \sum_{i=1}^N |o_i - p_i| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (o_i - p_i)^2} \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (o_i - p_i)^2}{\sum_{i=1}^N (o_i - \bar{o})^2} \quad (10)$$

where N is the number of observations, o is the observation, p is the prediction and \bar{o} is the mean value of the observations.

Next, the joinpoint regression-based pavement age effect estimation model is evaluated with an uncertainty analysis. The confidence intervals of estimations were computed for both APCs over all ranges of pavement age and AAPC over the entire range to test if the APCs and AAPC were significantly different from zero [65].

Finally, the random forest-based factors effect estimation model is evaluated using a ten-fold cross validation due to the consideration of model accuracy. The cross-validation indicators are MAE, RMSE and R^2 presented in equations (8)–(10). In the cross validation, the accuracy of the RFM was compared with the stepwise linear regression (SLR):

$$Y(s, t) = g(X(s, t)) + \epsilon(s, t) \quad (11)$$

and generalized additive model (GAM):

$$Y(s, t) = \sum_j h_j(X_j(s, t)) + \epsilon(s, t) \quad (12)$$

where Y is the roughness at the road segment s and the observation year t , X includes traffic, climate and environmental variables, g is the SLR, h_j is the spline function for

variable X_j in GAM, and ϵ is the random error. The GAM is a non-parametric regression model for assessing non-linear relationships between response and explanatory variables with a set of spline smooth functions [66], [67]. The GAM was produced using the R package “mgcv” [68].

E. Prediction Model and Maintenance Effects

In this stage, linear regression was used to construct the relationship between roughness and the three types of deterioration effects, including the NLS-based baseline effect, joinpoint regression-based pavement age effect, and RFM-based factors effect. The roughness prediction model was constructed with spatiotemporal observations of roads without maintenance activities.

To evaluate the maintenance effects, the roughness prediction model was applied to the maintained roads to compute the expected roughness of these roads if no maintenance action was taken. Thus, the difference between predicted and observed roughness of these roads revealed the practical effects of different types of maintenance, including resurfacing and rehabilitation.

F. Future Deterioration Scenarios

The future deterioration is a sum of weighted deterioration effects:

$$Z(s, t) = \sum_j \gamma_j (E_j(s, t)) \quad (13)$$

where $Z(s, t)$ is the predicted roughness at road segment s and year t , $E_j(s, t)$ is the j th ($j = 1, 2, 3$) type of the deterioration effects and γ_j is weight. The weights were computed using the averaging over orderings method proposed by Lindeman, Merenda and Gold (LMG) that determines the relative importance of variables of linear regressions with sequential sums of squares of the regressions [69]–[71]. The weights applied to the three deterioration effects were computed for the roughness without maintenance activities from data of the five periods, and data of the 15 time period combinations derived from the six-year observations during 2007 – 2018 to quantify the variation trends of weights of the three components. The LMG model was performed using the R package “reaiampo” for its fast computation and the availability of the bootstrap confidence intervals [72].

Further, future deterioration scenarios from 2007 to 2030 were predicted with the prediction model and parameters that the traffic volume was increased at an annual rate of 1.91% and the percentage of heavy vehicles was increased at an annual rate of 4.93%, which were the mean annual increase rates during 2007 – 2018, while other parameters were not changed. As shown in Table I, there were 4.62%–13.15% of missing data in the six observation years. To avoid the impacts of missing data in the roughness observations and pavement age records, only data of roads without missing data during all five periods listed in Table I were used for future deterioration scenarios prediction. The data used in the analysis included 49,025 observations which was 69.36% of all the data. In the study, deterioration scenarios were

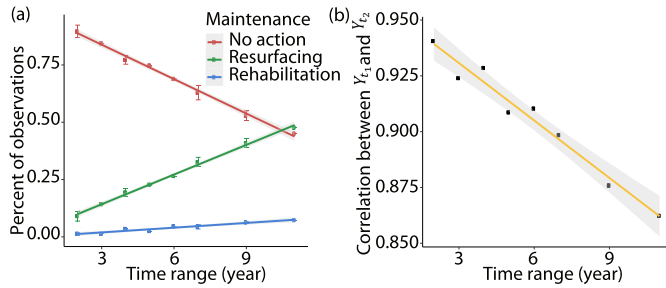


Fig. 4. Characteristics of roughness variations with time ranges. Baseline effects of deterioration explored by the NLS model. (a) Percentages of observations with different maintenance activities during the periods with time ranges ranging from 2 to 11 years. (b) The decreased trend of the relationship between the correlation of the roughness in the start year and that at the end year, and the time range.

predicted based on the roughness observations in each of the six observatory years from 2007 to 2018. The road surface conditions were characterized from two aspects in the future scenarios: distress of roughness and road performance. The distress was defined as the median roughness of roads across the whole network, and road performance was defined as the percentage of roads with roughness lower than the median roughness (2.085 IRI) of observations from 2007 to 2018.

IV. RESULTS AND DISCUSSION

A. Baseline Effect of Deterioration

The NLS model revealed that the time range of predictions and historical observations were critical for more accurate prediction of deterioration. Figure 4(a) shows that the percentages of resurfaced and rehabilitated (maintained) roads increased linearly with the increase of the time range. This means that more roads are treated from 2007 to 2018. Figure 4(b) reveals the decreased correlations between roughness in the start year and that at the end year during the periods with enlarged time ranges. The decreased correlations demonstrate that when the time range was increased, the prediction accuracy will be reduced when historical data are used for prediction. Therefore, the time range is an important parameter that should be considered in the baseline effect estimation.

The baseline effect is presented as the relationship between roughness in the start year and that at the end year during a series of time periods. Equation (14) is the estimated NLS relationships between predictions and historical observations of roughness that were adjusted with the parameter of time range:

$$\log(Y_{t_2}) = -0.8653 + (0.9337\log(Y_{t_1}) + 0.8920)(\Delta T + 1)^{0.0114} \quad (14)$$

Figure 5 shows the comparison of the LM and NLS in modelling the baseline effect, where the time range is ignored in the LM. The values of the time range are integer years ranging from 0 to 30, and the values of the roughness in the start year are ten quantiles of the roughness observations that are used to characterize statistical distributions of the roughness observations. The comparison demonstrates the varied relationships between predictions and observations with

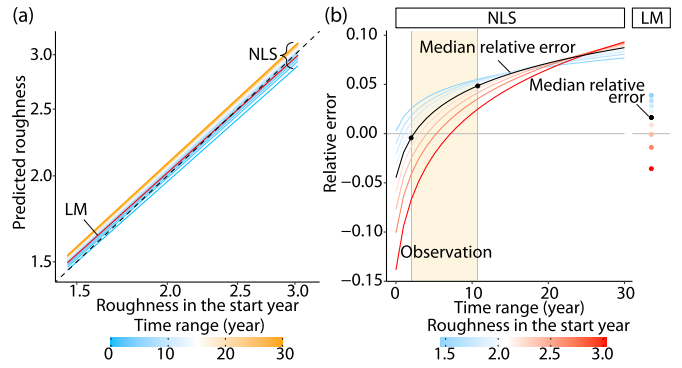


Fig. 5. Comparison of the linear regression model (LM) and NLS in modelling the baseline effect: (a) baseline effects and (b) relative errors.

TABLE V
TEN-FOLD CROSS VALIDATION OF THE NLS MODEL COMPARED WITH LINEAR REGRESSIONS

Cross validation indicator	LM		NLS	
	Training	Testing	Training	Testing
MAE	0.1679	0.1679	0.1670	0.1670
RMSE	0.2728	0.2728	0.2718	0.2718
R^2	0.8354	0.8354	0.8366	0.8366

changed time ranges. Further, the NLS model reflects the reduced prediction accuracy with the increased time range.

The NLS model for the baseline effect estimation was validated using a ten-fold cross validation approach and by the comparison with LM. The ten-fold cross validated MAE, RMSE and R^2 are presented in Table V. Compared with LM, MAE and RMSE for both training and testing data are slightly reduced, and R^2 for both data are slightly increased by the NLS model. In summary, the NLS model can reveal the impact of time range between observations and predictions on the prediction accuracy, and the validation indicates that the NLS model can provide slightly more accurate prediction due to the integration of the parameter of time range.

B. Pavement Age Effect of Deterioration

In the study, the joinpoint regression revealed an “S”-shape relationship between roughness and pavement age, as shown in Figure 6. In each pavement age group, roughness was summarized with mean (points) and standard error (error bars) values, where the standard error was the ratio between standard deviation and the squared number of observations. The result reveals that the pavement significantly becomes rougher with the increased age. For instance, the “S”-shape function indicates that the mean roughness across the road network in WA would increase 25.1% from the age 0 (mean roughness = 1.74) to age 40 (mean roughness = 2.18), and it would increase 60.1% from age 0 to age 60 (mean roughness = 2.79), if there is no maintenance activity.

When the pavement age was lower than 5 or higher than 65, the standard errors were relatively high due to the limited number of observations. The “S”-shape relationship consists of three line segments with two joinpoints (red points) at the pavement ages 33 and 63, respectively. The pavement age 63 is the estimated end of effective life (EEL) of the whole

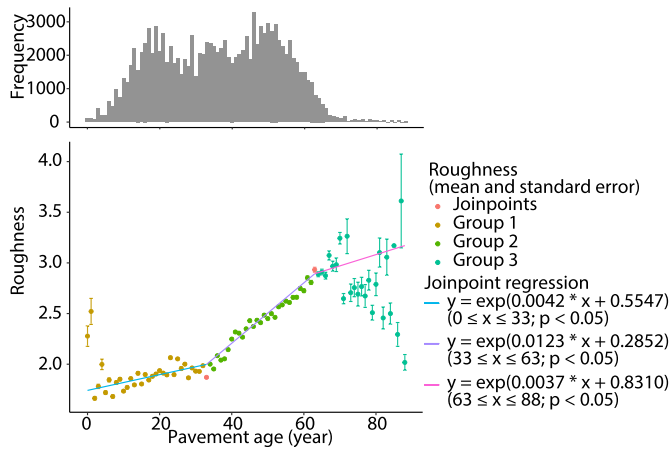


Fig. 6. Pavement age effects on deterioration identified by a joinpoint regression.

TABLE VI
SUMMARY OF JOINPOINT REGRESSION

Joinpoint	Parameters	Estimation	$CI_{0.025}$	$CI_{0.975}$
Joinpoint	α_1	33	28	38
	α_2	63	54	78
APC	$[0, \alpha_1]$	0.42 *	0.2	0.6
	$[\alpha_1, \alpha_2]$	1.23 *	1.1	1.5
	$[\alpha_2, 88]$	0.37 *	-1.2	0.6
AAPC	$[0, 88]$	0.7 *	0.6	0.7

* Parameter is significantly different from zero at 95% confidence interval (CI).

road network, which is close to the EEL of 56 reported by Austroads in 2008 [73]. The EEL is an important indicator for estimating the trigger point (TP) of pavement age for maintaining roads. Figure 6 also presents the equations of the joinpoint regression that is connected piecewise exponential function, together with the significance of slopes using the t-test.

Table VI shows a summary and evaluation of the joinpoint regression between roughness and pavement age. The results evaluation includes the estimations and 95% confidence interval (CI) for three parameters: joinpoints of pavement age, annual percent change (APC) for each of the line segments, and average annual percent change (AAPC) for the relationships over the whole range of pavement age. Compared with the first range of pavement age ($[0, 33]$, $APC = 0.42\%$), the annual deterioration rate in the second range has markedly increased ($[33, 63]$, $APC = 1.23\%$). Finally, the annual deterioration rate has increased with a low slope when the pavement age is higher than 63 ($APC = 0.37\%$). Observations at roads older than 63 are usually not referred in practice, since only a few roads are still used when their ages are older than planned life (e.g. 40 years) and roughness observations are greatly varied as shown in Figure 6. Over the entire range of pavement age, the AAPC is 0.7, which is also significantly higher than zero.

C. Deterioration Factors Effect

Contributions of traffic, climate and environmental factors to road roughness deterioration are investigated using an

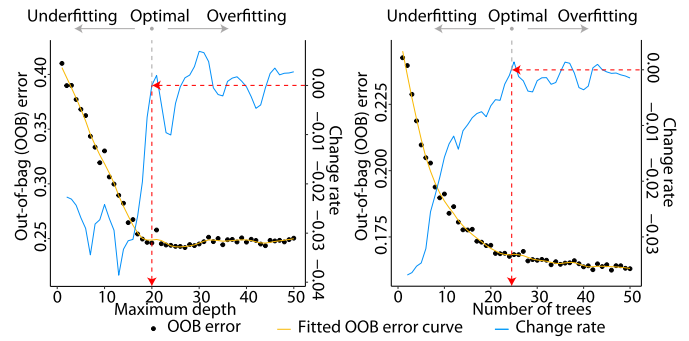


Fig. 7. The process of selecting the optimal maximum depth of nodes (MDN) and number of trees for the random forest model (RFM) in terms of the out-of-bag (OOB) errors.

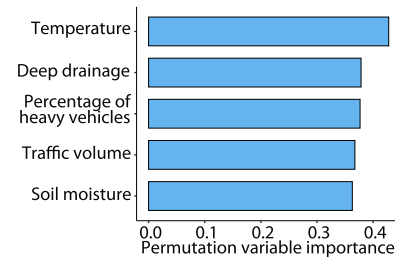


Fig. 8. Importance of variables for prediction in RFM.

RFM. Before modelling, a correlation analysis and collinearity diagnosis were performed to select variables. The criterion for diagnosing collinearities between factor variables is that the values of variance inflation factor (VIF) value less than 5 [74]–[76]. As a result, five factor variables were selected in the model, including traffic volume, percentage of heavy vehicles, temperature, soil moisture and soil deep drainage, where their highest VIF was less than 1.9. In the following analysis, these variables, together with baseline effects and pavement age, will be used to assess their combined impacts on deterioration. Figure 7 shows the process of selecting the optimal MDN and number of trees for the RFM based on the OOB errors. As a result, the optimal MDN is 20 and the optimal number of trees is 24.

In the RFM with optimal parameters, the relative importance of variables was determined by the permutation importance evaluation approach. The results show that variables ranked by the importance from the highest to the lowest are temperature, deep drainage, percentage of heavy vehicles, traffic volume and soil moisture (Figure 8). The variables rainfall, runoff and soil evaporation are not included in the RFM since they were removed for the multicollinearity.

The model performance of the RFM was evaluated by the comparison with the SLR and GAM methods in a ten-fold cross validation. Table VII shows the statistical summary of the ten-fold cross validation for the SLR, GAM and RFM models. Compared with linear regression, non-linear models can significantly reduce prediction errors and improve fitness for both training and testing data. Among the models, the RFM has the lowest MAE and RMSE, and the highest R^2 for both training and testing data, which demonstrates the accuracy and effectiveness of the RFM in evaluating the deterioration factors effect.

TABLE VII
TEN-FOLD CROSS VALIDATION FOR SLR, GAM, AND RFM FOR ESTIMATING THE FACTORS EFFECT OF ROAD DETERIORATION

Cross validation indicator	SLR		GAM		RFM	
	Training	Testing	Training	Testing	Training	Testing
MAE	0.479	0.474	0.471	0.467	0.232	0.295
RMSE	0.616	0.611	0.606	0.602	0.314	0.402
R^2	0.096	0.099	0.126	0.126	0.785	0.636

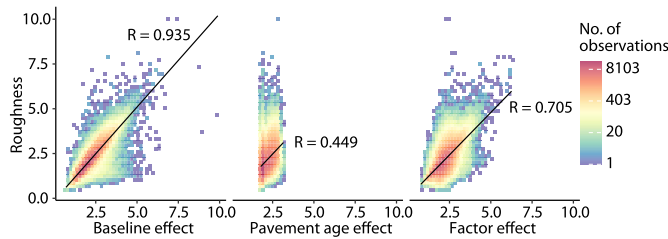


Fig. 9. Estimated baseline, pavement age and factors effects of road deterioration and their relationships with roughness at the end years of five periods.

TABLE VIII
CONTRIBUTIONS OF THE THREE TYPES OF EFFECTS OF DETERIORATION IN THE FIVE-PERIOD MODEL

Period	Baseline effect	Pavement age effect	Factors effect	Total
2007-2009	46.5%	5.9%	35.8%	88.2%
2009-2012	43.9%	7.5%	37.0%	88.4%
2012-2014	46.9%	7.5%	36.8%	91.2%
2012-2016	45.7%	6.7%	37.3%	89.8%
2016-2018	42.3%	6.9%	38.6%	87.7%

D. Deterioration Prediction Model

Figure 9 shows the significantly correlated relationships between roughness observations and the estimated three types of deterioration effects, including the baseline effect, pavement age effect and factors effect, with the correlation coefficients of 0.935, 0.449 and 0.705, respectively. The results indicate that the baseline effect has the highest correlation, and the pavement age effect has the lowest correlation with the roughness. According to the weights computed by the LMG model, contributions of the three types of deterioration effects in each period of the five-period model are listed in Table VIII. The contributions of the baseline effect, pavement age effect and factors effect range from 42.3% to 46.9%, 5.9% to 7.5%, and 35.8% to 38.6%, respectively for the five periods. The total contribution ranges from 87.7% to 91.2%, and the model fitness (R^2) for the entire five-period spatiotemporal dataset is 0.891, which is higher than each individual effect: baseline effect ($R^2 = 0.875$), pavement age effect ($R^2 = 0.201$) and factors effect ($R^2 = 0.497$).

E. Maintenance Effects Analysis

Roads that need to be maintained usually have relatively high risk of increased deterioration. Figure 10 shows the comparison of roughness observations in the start year and at the end year during the five periods for three maintenance groups, including no maintenance actions, resurfacing and rehabilitation, to characterize the roughness of roads that need

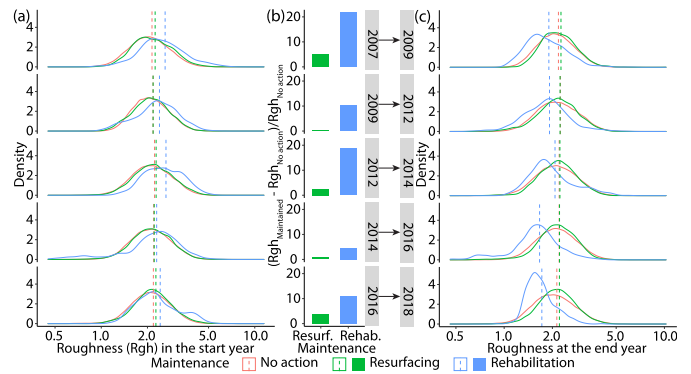


Fig. 10. Comparison of roughness observations in the start year (a) and (b) and at the end year (c) during the five periods for three maintenance groups, no actions, resurfacing and rehabilitation. (b) shows the percentages of roughness in the maintained groups higher than the no action groups in the start year.

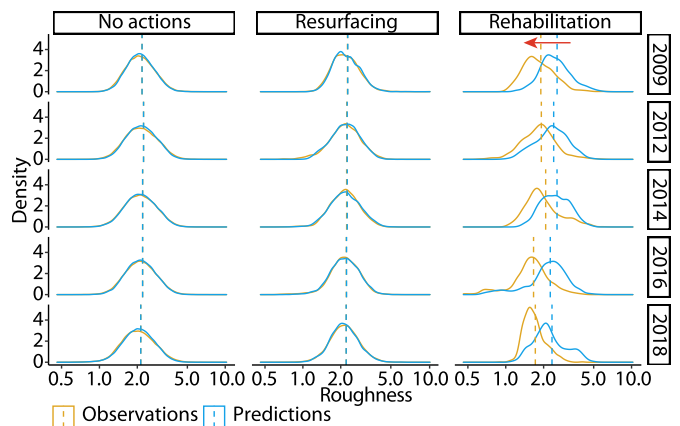


Fig. 11. Distribution comparison of observations and predictions of roughness in three maintenance groups: no actions, resurfacing and rehabilitation.

to be maintained. For instance, for the period 2007 – 2009, roughness of the roads to be rehabilitated was much higher than that of the no action group in 2007 (Figure 10(a)). In 2007, the mean roughness of the roads to be resurfaced was 4.8% higher than the no action group, and the roughness of the roads to be rehabilitated was 21.7% higher than the no action group (Figure 10(b)). In 2009, roughness of the rehabilitated roads was significantly reduced and much lower than the no action group (Figure 10(c)). Compared with the no maintenance action and resurfacing groups, roughness is more effectively reduced in the rehabilitation group. The density plots and statistical summaries show the similar variations patterns of roughness for other four periods. In the five periods, the mean roughness of roads to be resurfaced was 0.1%–4.8% higher than the no action group, and that of roads to be rehabilitated was 4.5%–21.7% higher than the no action group. Therefore, the analysis proves the necessity to maintain the roads with relatively high roughness and describes the changes of roughness during the five periods from 2007 to 2018.

Figure 11 shows the density distributions of the observations and DTOM-based predictions of roughness in the three maintenance groups, including no maintenance action, resurfacing and rehabilitation. The roughness in the no action and resurfacing groups was slightly varied, but that in the rehabilitation

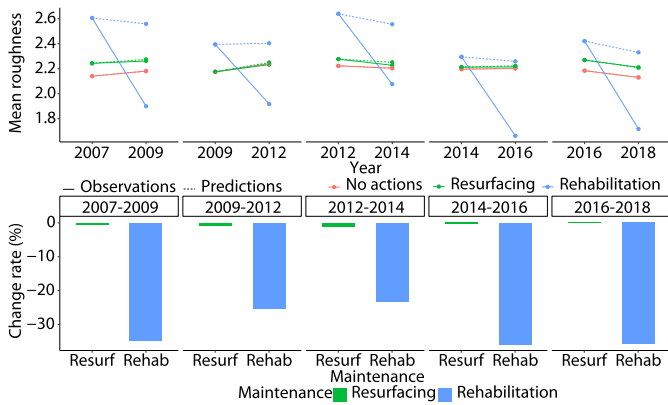


Fig. 12. Comparison of mean observations and predictions of roughness, and maintenance effects of resurfacing and rehabilitation works.

group was significantly reduced for data in all five periods. Figure 12 shows the comparison of mean observations and DTOM-based predictions of roughness to accurately capture the maintenance effects. The top figures present the predicted and practical variation trends of roughness for the three groups of roads, and the bottom figures show the change rate of observations compared with predictions due to resurfacing and rehabilitation activities. Results show that rehabilitation has a high contribution in reducing roughness and maintaining the quality of road surface, while resurfacing has a much lower contribution in reducing roughness. Resurfacing reduced roughness by 0.1%–1.1% and the rehabilitation reduced roughness by 23.1%–35.8% for the five periods from 2007 to 2018. The effects of maintenance tended to have similar and consistent patterns and were not greatly varied during the five periods. The relatively consistent effects of resurfacing and rehabilitation are an outcome of the continuous and effective decision making of the strategic road network maintenance by MRWA.

In addition, the results also reveal potential effects of other types of maintenance activities other than resurfacing and rehabilitation. For the no maintenance action groups, the mean roughness was increased in the periods 2007 – 2009, 2009 – 2012, and 2014 – 2016, and it is slightly decreased in the periods 2012 – 2014 and 2016 – 2018. The decreased roughness of the no maintenance action groups indicates the effects of other forms of road maintenance activities, such as reactive maintenance works, including pothole patching, localised surface corrections, sprayed sealing and dig outs.

F. Contributions of Three Types of Effects in Prediction

Contributions of the three types of effects are changed with the time range between observations and predictions. Figure 13 shows the contribution comparisons of three types of effects between the five-period prediction model where the time range is 2 to 3 years and the 2007 – 2018 prediction model where the time range is 11 years. Note that the collinearity of the three types of deterioration effects is not significant since the VIF values are all less than 2. The contribution of the baseline effect to the roughness deterioration decreased from 45.1% in the five-period model, to 36.2% in the 2007 – 2018 model, but the contribution of the pavement age effect slightly

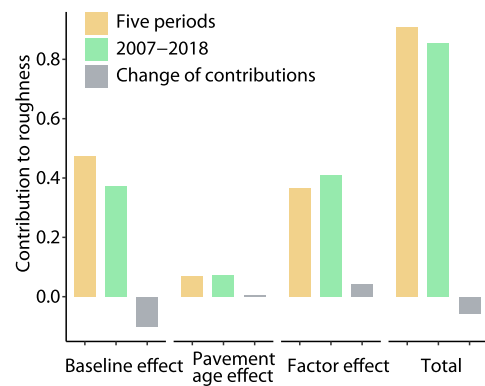


Fig. 13. Contribution comparisons of three types of effects between the five-period and the 2007-2018 prediction models.

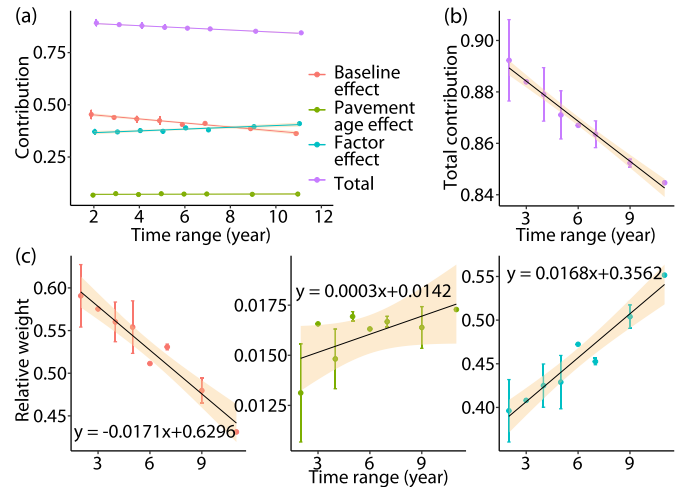


Fig. 14. Contributions of the baseline, pavement age and factors effects on road deterioration for prediction models with the time range from 2 to 11 years: (a) contributions of three types of effects, (b) total contribution, and (c) relative weights of the three types of deterioration effects.

increased from 6.9% in the five-period model to 7.3%, and the contribution of the deterioration factors effect increased from 37.1% to 41.0% in the 2007 – 2018 model. Thus, the changes of contributions of the baseline effect, pavement age effect and factors effect are -8.9%, -0.4% and 3.9%, respectively. In total, 4.6% contribution of the three types of effects decreased from the five-period model to the 2007 – 2018 model. Figure 14 shows the changed weights of the three types of effects, which is computed as their relative contributions, when the time range between historical records and predictions are varied from 2 years to 11 years. In general, when the time range is enlarged, the total contributions of the three types of effects are decreased. Results show the significant linear variation trends of the decreased contribution of the baseline effect and the increased contribution of the factors effect. The linear regressions (Figure 14(c)) between contributions of the three types of effects and time range are used to quantify weights of the three types of effects in predicting future scenarios of deterioration.

G. Future Deterioration Scenarios

The deterioration prediction model and weighting trends of three types of effects are applied in predicting future roughness

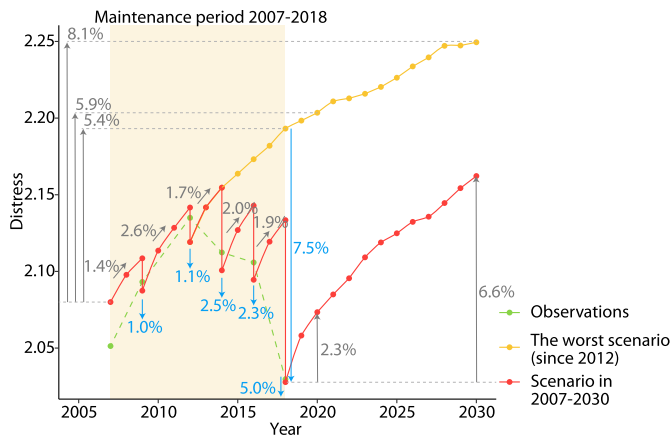


Fig. 15. The DTOM-based future roughness distress (median roughness) scenarios of the road network.

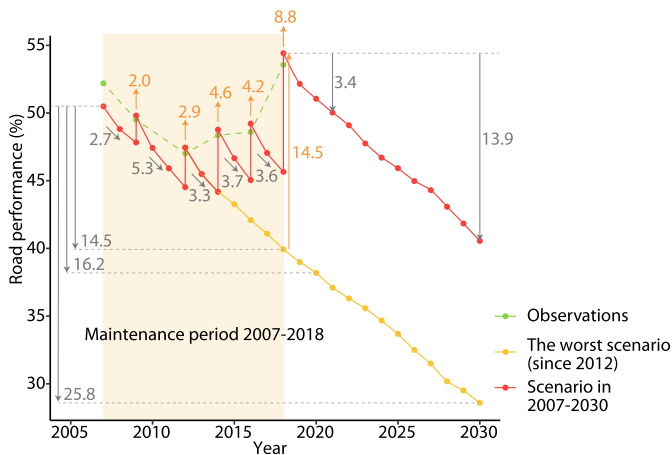


Fig. 16. The DTOM-based future road performance (percentage of roads with roughness lower than 2.085 IRI) scenarios of the road network.

deterioration scenarios. Figure 15 and Figure 16 show the summaries of the DTOM-based future distress scenarios and road performance scenarios of the whole road network in WA, respectively. The red lines show the DTOM-based scenarios from 2007 to 2030. Future scenarios are explained from following four aspects.

First, in general, the network roughness increased from 2007 to 2012 and decreased from 2012 to 2018. The increased roughness in 2007 – 2012 indicated that the wearing impact on the road from road users, aging of roads and local environmental factors was higher than expected. In 2012 – 2018, the accumulation of continuous strategic road infrastructure maintenance activities generated positive outcomes as the roughness was gradually reduced, and the road performance as a consequence increased. The deterioration scenario since 2012 (orange lines) was the worst scenario among those predicted since other years.

Second, in 2007 – 2018, if no action was taken to maintain roads, the roughness would continuously increase, and road performance would decrease. The comparison between the worst scenario since 2012 and observations in 2018 shows that the maintenance activities, including resurfacing and rehabilitation, help reduce the distress of roughness by 7.5%,

and improve the road performance by 14.5%. In the five periods in 2007 – 2018, the annual increase rate of the distress of roughness was 0.7%–1.0%, and the annual reduction rate of road performance was 1.4%–1.9%. The continuous maintenance activities help reduce the distress of roughness by 0.4%–2.5% and improve the road performance by 1.0%–4.4% for the road network.

Third, if no action was taken to maintain the roads after 2018, the average distress of roughness of the road network would increase 2.3% in 2020 and 6.6% in 2030, and the road performance would reduce 3.4% in 2020 and 13.9% in 2030. Predictions also show that if no action was taken after 2007, the average road distress of roughness would increase 5.9% in 2020 and 8.1% in 2030, and the road performance would reduce 16.2% in 2020 and 25.8% in 2030.

Finally, these results also prove that the contribution of maintenance activities in a single year tend to be limited in the whole network, but the accumulation of the maintenance activities over a period can significantly reduce the overall roughness and improve road performance. A long-term pavement performance monitoring and analysis was conducted for over twenty years by the road agencies in Australia and New Zealand [10]. The efforts of long-term data collection are essential for monitoring road deterioration, assessing maintenance effects and improving practices of road agencies [77]. Therefore, the insight drawn from the analysis of multi-source data and emerging technologies can improve decision making that are critically important for the sustainable road infrastructure and sustainable development [78], [79].

H. Challenges Towards Large-Scale Applications of DTOM

Methods for analyzing large-scale transport infrastructure are increasingly required for strategic maintenance and management. However, benchmark studies are still limited due to following reasons. First, most nations lack the large-scale high-resolution observations of road deteriorations. Next, methods are still limited in dealing with different scales of data, from project-level road failures to network-level data aggregation. Finally, it is still a critical challenge to coordinate maintenance operations with limited machines and resources for large-scale road management. This study provided a potential solution and guide to address these challenges. The temporally varied impacts of different types of road deterioration factors have been identified by the DTOM. The quantified temporally varied impacts of factors are critical for more accurate examination of road maintenance effects and prediction of future deterioration scenarios.

There are still limitations in the study and future studies are recommended in following aspects. First, annual routine maintenance has impacts on the roughness deterioration, but it is difficult to be separated from the data. Pavement type, material, depth, and techniques, which are not included in this study, are also closely associated with deteriorations. For instance, pavement surfaces with sprayed seal and asphalt have different deterioration modes. Thus, the analysis of annual routine maintenance and pavement related variables are recommended for future studies. In addition, endogeneity bias may be introduced into the prediction model when data

of roads without maintenance are used to predict the performance of roads with relatively high deterioration, which tend to receive more treatments than others [80]. The selectivity correction methods can be integrated in modelling to reduce the errors [81]. Last, the strength of pavement and subgrade on roughness deterioration is recommended to be measured and added in the model to further improve prediction accuracy.

V. CONCLUSION

This study developed a hybrid nonlinear and machine learning model for investigating the large-scale road deterioration, assessing the effects of strategic maintenance activities and predicting future scenarios. This study provided following findings about road deterioration and maintenance activities.

The baseline effect reveals the non-linear contributions of the time range on the relationship between historical road conditions and current or future deterioration. An “S”-shape trend is identified between the pavement age and deterioration. In addition to historical conditions and pavement age, road deterioration is also impacted by traffic, climate and environmental conditions. With the increase of the time range between historical observations and predictions, the contribution of the baseline effect decreases gradually, and the contribution of the deterioration factors effect increases.

Resurfacing and rehabilitation are essential for strategic reduction of deterioration at different life cycle stages of the road pavement. The temporally continuous, large spatial scale and effective strategic maintenance activities in WA have significantly reduced the potential rapid roughness progression and improved overall road performance. The strategic maintenance in WA is a benchmark case for maintaining roads in a sustainable road infrastructure development at the network level. The experience of using vehicle-based laser scanner sensors monitoring high-resolution deterioration data, multi-source deterioration factors data and comprehensive methods for predicting deterioration and maintenance effects in WA, Australia, can be applied to effectively manage the road infrastructure in other countries and regions.

ACKNOWLEDGMENT

The authors would like to thank Brett Belstead from Main Roads Western Australia and Tim Martin from ARRB Group Ltd., Australia, for providing support, comments and suggestion for this research. Note that the views expressed here do not reflect the views of Main Roads Western Australia.

REFERENCES

- [1] S. G. Perz, “Sustainable development: The promise and perils of roads,” *Nature*, vol. 513, no. 7517, pp. 178–179, 2014.
- [2] M. V. Chester, “Sustainability and infrastructure challenges,” *Nature Sustainability*, vol. 2, no. 4, pp. 265–266, Apr. 2019.
- [3] E. E. Koks *et al.*, “A global multi-hazard risk analysis of road and railway infrastructure assets,” *Nature Commun.*, vol. 10, no. 1, pp. 1–11, Dec. 2019.
- [4] P. Luo, Y. Song, and P. Wu, “Spatial disparities in trade-offs: Economic and environmental impacts of road infrastructure on continental level,” *GIScience Remote Sens.*, vol. 58, no. 5, pp. 756–775, Jul. 2021, doi: [10.1080/15481603.2021.1947624](https://doi.org/10.1080/15481603.2021.1947624).
- [5] T. J. Kwon, L. Fu, and S. J. Melles, “Location optimization of road weather information system (RWIS) network considering the needs of winter road maintenance and the traveling public,” *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 1, pp. 57–71, Jan. 2017.
- [6] H. R. Vignisdottir *et al.*, “A review of environmental impacts of winter road maintenance,” *Cold Regions Sci. Technol.*, vol. 158, pp. 143–153, Feb. 2019.
- [7] F. Zhang, N. Zheng, H. Yang, and N. Geroliminis, “A systematic analysis of multimodal transport systems with road space distribution and responsive bus service,” *Transp. Res. C, Emerg. Technol.*, vol. 96, pp. 208–230, Nov. 2018.
- [8] T. Sullivan and R. Scott, “Strategic road network management: An approach using roughness,” in *Proc. 6th Road Eng. Assoc. Asia Australasia (REAAA)*, Kuala Lumpur, Malaysia, 1990, p. 26.
- [9] S. Moodley, “Prioritising maintenance and rehabilitation projects to ensure sustainable pavement preservation,” in *Proc. Pavement Asset Manage., World Conf. (WCPAM)*, Baveno, Italy. Boca Raton, FL, USA: CRC Press, Jun. 2019, p. 159.
- [10] T. Martin and L. Choumanivong, “The benefits of long-term pavement performance (LTPP) research to funders,” *Transp. Res. Proc.*, vol. 14, pp. 2477–2486, Jan. 2016.
- [11] P. S. Chinowsky, J. C. Price, and J. E. Neumann, “Assessment of climate change adaptation costs for the U.S. road network,” *Global Environ. Change*, vol. 23, no. 4, pp. 764–773, Aug. 2013.
- [12] Y. Qiao, J. Santos, A. M. K. Stoner, and G. Flinisch, “Climate change impacts on asphalt road pavement construction and maintenance: An economic life cycle assessment of adaptation measures in the state of Virginia, United States,” *J. Ind. Ecol.*, vol. 24, no. 2, pp. 342–355, Apr. 2020.
- [13] S. Yarmukhamedov, A. S. J. Smith, and J.-C. Thiebaud, “Competitive tendering, ownership and cost efficiency in road maintenance services in Sweden: A panel data analysis,” *Transp. Res. A, Policy Pract.*, vol. 136, pp. 194–204, Jun. 2020.
- [14] T.-Y. Chuang, N.-H. Perng, and J.-Y. Han, “Pavement performance monitoring and anomaly recognition based on crowdsourcing spatiotemporal data,” *Autom. Construct.*, vol. 106, Oct. 2019, Art. no. 102882.
- [15] I. Wessels and W. J. Steyn, “Continuous, response-based road roughness measurements utilising data harvested from telematics device sensors,” *Int. J. Pavement Eng.*, vol. 21, no. 4, pp. 437–446, Mar. 2020.
- [16] N. H. M. Nasir, W. M. W. Mohamed, and K. N. Tahar, “A review on road distress detection methods,” *Adv. Transp. Logistics Res.*, vol. 1, pp. 230–241, Oct. 2018.
- [17] R. Fan, U. Ozgunalp, B. Hosking, M. Liu, and I. Pitas, “Pothole detection based on disparity transformation and road surface modeling,” *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 897–908, Aug. 2019.
- [18] Y. Song *et al.*, “Spatial and temporal variations of spatial population accessibility to public hospitals: A case study of rural–urban comparison,” *GIScience Remote Sens.*, vol. 55, no. 5, pp. 718–744, Sep. 2018.
- [19] L. Fu and T. Usman, “Safety effects of winter weather and road maintenance operations,” in *Sustainable Winter Road Operations*. Hoboken, NJ, USA: Wiley, 2018, pp. 101–130.
- [20] K. H. Janstrup, M. Møller, and N. Pilegaard, “A clustering approach to integrate traffic safety in road maintenance prioritization,” *Traffic Injury Prevention*, vol. 20, no. 4, pp. 442–448, May 2019.
- [21] Y. Song and P. Wu, “An interactive detector for spatial associations,” *Int. J. Geographical Inf. Sci.*, vol. 20, pp. 1–26, Apr. 2021.
- [22] M. E. Beckley, *Pavement Deterioration Modeling Using Historical Roughness Data*. Phoenix, AZ, USA: Arizona State Univ., 2016.
- [23] G. Chai, R. van Staden, S. H. Chowdhury, and Y. C. Loo, “A study of the effects of pavement ageing on binder deterioration,” *Int. J. Pavement Eng.*, vol. 15, no. 1, pp. 1–8, Jan. 2014.
- [24] S. W. Katicha *et al.*, “Development of enhanced pavement deterioration curves,” Virginia Transportation Research Council, Charlottesville, VA, USA, Tech. Rep. VTRC 17-R7, 2016.
- [25] A. Pantuso, G. W. Flinisch, S. W. Katicha, and G. Loprencipe, “Development of network-level pavement deterioration curves using the linear empirical Bayes approach,” *Int. J. Pavement Eng.*, vol. 22, no. 6, pp. 1–14, 2019.
- [26] K. George, A. Rajagopal, and L. Lim, “Models for predicting pavement deterioration,” *Transp. Res. Rec.*, no. 1215, pp. 1–7, 1989.
- [27] T. C. Martin, “New deterioration models for sealed granular pavements,” in *Proc. Inst. Civil Eng.-Transp.*, 2009, vol. 162, no. 4, pp. 215–226.
- [28] X. Mao, J. Wang, C. Yuan, W. Yu, and J. Gan, “A dynamic traffic assignment model for the sustainability of pavement performance,” *Sustainability*, vol. 11, no. 1, p. 170, 2019.
- [29] Y. Song, X. Wang, G. Wright, D. Thatcher, P. Wu, and P. Felix, “Traffic volume prediction with segment-based regression kriging and its implementation in assessing the impact of heavy vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 232–243, Jan. 2019.

- [30] A. Zareie, M. S. R. Amin, and L. E. Amador-Jiménez, "Thornthwaite moisture index modeling to estimate the implication of climate change on pavement deterioration," *J. Transp. Eng.*, vol. 142, no. 4, Apr. 2016, Art. no. 04016007.
- [31] A. E. A. El-Maaty, "Temperature change implications for flexible pavement performance and life," *Int. J. Transp. Eng. Technol.*, vol. 3, no. 1, pp. 1–11, 2017.
- [32] Y. Song *et al.*, "Segment-based spatial analysis for assessing road infrastructure performance using monitoring observations and remote sensing data," *Remote Sens.*, vol. 10, no. 11, p. 1696, Oct. 2018.
- [33] T. Wang *et al.*, "How can the UK road system be adapted to the impacts posed by climate change? By creating a climate adaptation framework," *Transp. Res. D, Transp. Environ.*, vol. 77, pp. 403–424, Dec. 2019.
- [34] Y. Song, J. Wang, Y. Ge, and C. Xu, "An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data," *GIScience Remote Sens.*, vol. 57, no. 5, pp. 593–610, Jul. 2020.
- [35] A. Susanna, M. Crispino, F. Giustozzi, and E. Toraldo, "Deterioration trends of asphalt pavement friction and roughness from medium-term surveys on major Italian roads," *Int. J. Pavement Res. Technol.*, vol. 10, no. 5, pp. 421–433, Sep. 2017.
- [36] L. Yu *et al.*, "Long-term performance deterioration models for semi-rigid asphalt pavement in cold region," *Int. J. Pavement Res. Technol.*, vol. 14, no. 6, pp. 1–11, 2021.
- [37] T. Watanatada, C. G. Harral, W. D. Paterson, A. M. Dhareshwar, A. Bhandari, and K. Tsunokawa, "The highway design and maintenance standards model. Volume 1, description of the HDM-III model. Volume 2, user's manual for the HDM-III model," *World Bank Highway Des. Maintenance Standards Ser.*, vol. 1, p. 673, 1987.
- [38] A. AASHTO, "Mechanistic-empirical pavement design guide: A manual of practice," 2nd ed., AAoSHaT Officials, Washington, DC, USA, Tech. Rep., 2008.
- [39] G. Morosiuk, M. Riley, and J. Odoki, "Modelling road deterioration and works effects-version 2-highway development and management-HDM-4," *Highway Develop. Manage. Ser.*, vol. 6, May 2004.
- [40] N. Abdelaziz, R. T. Abd El-Hakim, S. M. El-Badawy, and H. A. Afify, "International roughness index prediction model for flexible pavements," *Int. J. Pavement Eng.*, vol. 21, no. 1, pp. 88–99, Jan. 2020.
- [41] H. Pérez-Acebo, A. Linares-Unamunzaga, E. Rojí, and H. Gonzalo-Orden, "IRI performance models for flexible pavements in two-lane roads until first maintenance and/or rehabilitation work," *Coatings*, vol. 10, no. 2, p. 97, Jan. 2020.
- [42] G. Sollazzo, T. F. Fwa, and G. Bosurgi, "An ANN model to correlate roughness and structural performance in asphalt pavements," *Construct. Building Mater.*, vol. 134, pp. 684–693, Mar. 2017.
- [43] K. Hwayyis, R. Hassan, and M. T. Fahey, "Factors affecting performance of sprayed seals in rural Victoria," *Int. J. Pavement Eng.*, vol. 21, pp. 1–15, Nov. 2020.
- [44] S. S. Islam, S. K. Singh, G. D. Ransinchung, and S. S. Ravindranath, "Effect of property deterioration in SBS modified binders during storage on the performance of asphalt mix," *Construct. Building Mater.*, vol. 272, Feb. 2021, Art. no. 121644.
- [45] M. W. Sayers, "The international road roughness experiment: Establishing correlation and a calibration standard for measurements," World Bank, Washington, DC, USA, Tech. Rep. 1084, 1986.
- [46] M. W. Sayers, "Guidelines for conducting and calibrating road roughness measurements," World Bank, Washington, DC, USA, Tech. Rep. 46, 1986.
- [47] H. Pérez-Acebo, H. Gonzalo-Orden, D. J. Findley, and E. Rojí, "Modeling the international roughness index performance on semi-rigid pavements in single carriageway roads," *Construct. Building Mater.*, vol. 272, Feb. 2021, Art. no. 121665.
- [48] Main Roads Western Australia. (2019). *Road Network in Western Australia (MRWA-514)*. [Online]. Available: <https://catalogue.data.wa.gov.au/dataset/mrwa-road-network>
- [49] W. Holtrop, "Sprayed sealing practice in Australia," *Road Transp. Res. A, J. Austral. New Zealand Res. Pract.*, vol. 17, no. 4, pp. 41–60, 2008.
- [50] D. Kennedy and B. Peters, "Pavement and surfacing condition in Western Australia with ten-year maintenance contracts," in *Proc. 23rd ARRB Conf.*, Adelaide, SA, Australia, 2008, p. 16.
- [51] A. Nega, H. Nikraz, S. Herath, and B. Ghadimi, "Distress identification, cost analysis and pavement temperature prediction for the long-term pavement performance for western Australia," *Int. J. Eng. Technol.*, vol. 7, no. 4, pp. 267–275, Aug. 2015.
- [52] Z. Rice, K. Sharp, and G. Cocks, "Assessment of thin asphalt surfacing fatigue design in Western Australia," in *Proc. 28th ARRB Int. Conf.*, Brisbane, QLD, Australia, 2018, p. 14.
- [53] Y. Song, P. Wu, D. Gilmore, and Q. Li, "A spatial heterogeneity-based segmentation model for analyzing road deterioration network data in multi-scale infrastructure systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 23, 2020, doi: [10.1109/TITS.2020.3001193](https://doi.org/10.1109/TITS.2020.3001193).
- [54] Main Roads Western Australia, "Technology and environment directorate: Guidelines for surfacing type selection," Main Roads Western Australia, East Perth, WA, Australia, Tech. Rep. 6706/04/154, 2004. [Online]. Available: <https://www.mainroads.wa.gov.au/technical-commercial/technical-library/>
- [55] G. Jameson, "Guide to pavement technology: Part 5: Pavement evaluation and treatment design," ARRB, Melbourne, VIC, USA, Tech. Rep. AGPT05/19, 2019.
- [56] Main Roads Western Australia, "Prime mover, trailer combinations—Vehicle categories," Main Roads Western Australia, East Perth, WA, Australia, Tech. Rep. D12#361585, 2016.
- [57] Z. Wan, S. Hook, and G. Hulley, *MOD11A2 Modis/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1 km Sin Grid V006; NASA Eosdis Land Processes DAAC, Usgs Earth Resources Observation and Science (Eros) Center, NASA EOSDIS Land Processes DAAC*, Sioux Falls, SD, USA, 2015.
- [58] L. Ashton and N. McKenzie, "Conversion of the atlas of Australian soils to the Australian Soil Classification," in *Proc. Austral. Soil Resource Inf. Syst. (ASRIS)*, 2001.
- [59] L. X. Clegg, B. F. Hankey, R. Tiwari, E. J. Feuer, and B. K. Edwards, "Estimating average annual per cent change in trend analysis," *Statist. Med.*, vol. 28, no. 29, pp. 3670–3682, Dec. 2009.
- [60] *Surveillance, Epidemiology and End Results Program*, NC Institute, SEER Softw., Bethesda, MD, USA, 2020.
- [61] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [62] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [63] S. Nembrini, I. R. König, and M. N. Wright, "The revival of the Gini importance?" *Bioinformatics*, vol. 34, no. 21, pp. 3711–3718, Nov. 2018.
- [64] M. N. Wright and A. Ziegler, "Ranger: A fast implementation of random forests for high dimensional data in C++ and R," 2015, *arXiv:1508.04409*. [Online]. Available: <http://arxiv.org/abs/1508.04409>
- [65] H.-J. Kim, M. P. Fay, E. J. Feuer, and D. N. Midthune, "Permutation tests for jointpoint regression with applications to cancer rates," *Statist. Med.*, vol. 19, no. 3, pp. 335–351, Feb. 2000.
- [66] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*. Boca Raton, FL, USA: CRC Press, 1990, vol. 43.
- [67] S. N. Wood, *Generalized Additive Models: An Introduction With R*. Boca Raton, FL, USA: CRC Press, 2017.
- [68] S. N. Wood, "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 73, no. 1, pp. 3–36, 2011.
- [69] R. H. Lindeman, *Introduction to Bivariate and Multivariate Analysis*. Glenview, IL, USA: Scott Foresman, 1980.
- [70] W. Kruskal, "Relative importance by averaging over orderings," *Amer. Statistician*, vol. 41, no. 1, pp. 6–10, Feb. 1987.
- [71] N. Redell, "Shapley decomposition of R-Squared in machine learning models," 2019, *arXiv:1908.09718*. [Online]. Available: <http://arxiv.org/abs/1908.09718>
- [72] U. Grömping, "Relative importance for linear regression in R: The packagerelaimpo," *J. Stat. Softw.*, vol. 17, no. 1, pp. 1–27, 2006.
- [73] T. Martin and M. Moffatt, *Assessment of Remaining Service Life of Pavements*. Melbourne, VIC, Australia: ARRB, 2008.
- [74] M.-J. Kim and D.-K. Kang, "Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9308–9314, Aug. 2012.
- [75] J. Komlos, "Multicollinearity in the presence of errors-in-variables can increase the probability of type-I error," CESifo, Munich, Germany, CESifo Work. Paper, 2019.
- [76] T. Zhang, X. P. Zhou, and X. F. Liu, "Reliability analysis of slopes using the improved stochastic response surface methods with multicollinearity," *Eng. Geol.*, vol. 271, Jun. 2020, Art. no. 105617.
- [77] T. Martin, *Data Review and Calibration of HDM-4 Road Deterioration Models*. Melbourne, VIC, Australia: ARRB, 2004.

- [78] Y. Song, D. Thatcher, Q. Li, T. McHugh, and P. Wu, "Developing sustainable road infrastructure performance indicators using a model-driven fuzzy spatial multi-criteria decision making method," *Renew. Sustain. Energy Rev.*, vol. 138, Mar. 2021, Art. no. 110538. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032120308224>
- [79] Y. Song and P. Wu, "Earth observation for sustainable infrastructure: A review," *Remote Sens.*, vol. 13, no. 8, p. 1528, Apr. 2021.
- [80] S. Madanat and R. Mishalani, "Selectivity bias in modeling highway pavement maintenance effectiveness," *J. Infrastruct. Syst.*, vol. 4, no. 3, pp. 134–137, Sep. 1998.
- [81] V. V. Gayah and S. Madanat, "Accounting for endogeneity in maintenance decisions and overlay thickness in a pavement-roughness deterioration model," *J. Infrastruct. Syst.*, vol. 23, no. 4, Dec. 2017, Art. no. 04017023.

Yongze Song (Member, IEEE) received the B.S. and M.S. degrees in surveying and mapping from China University of Geosciences, Beijing, China, in 2012 and 2015, respectively, and the Ph.D. degree from Curtin University, Perth, WA, Australia, in 2019.

He is currently a Lecturer with Curtin University. His research interests include spatial statistics, geospatial methods, urban remote sensing, sustainable infrastructure, and sustainable development.

Dr. Song is a Fellow of the Royal Geographical Society (with IBG), U.K., and an Editorial Board Member of journal *GIScience & Remote Sensing*.

Peng Wu received the B.S. degree in project management from Tsinghua University, China, in 2006, the M.S. degree in construction management from Loughborough University, U.K., in 2007, and the Ph.D. degree in project management from the National University of Singapore, Singapore, in 2012.

He is currently a Professor with the Department of Construction Management and an Associate Director with the Australasian Joint Research Centre for Building Information Modeling, Curtin University. His research interests include sustainable construction, lean production and construction, production and operations management, and life cycle assessment.

Dr. Wu received the Discovery Early Career Research Award from the Australian Research Council in 2016, which is a prestigious award to support excellent basic and applied research by early career researchers.

Qindong Li received the B.S. degree in civil engineering from Sichuan University, China, in 1992, and the M.E. degree in civil engineering from RMIT University, Australia, in 2003.

He has over 18 years' experience in design, construction, and research of road transport infrastructures. He is currently the Asset Management Modeling and Analytics Manager at the Main Roads Western Australia, Perth, WA, Australia. His research interests include pavement deterioration, whole of life costing, and the use of business intelligence in asset management.

Yuchen Liu received the M.S. degree in geospatial science from Curtin University, Perth, Australia, in 2016, and the B.S. degree in geo-information science and technology from Sun Yat-sen University, Guangzhou, China, in 2020. He is currently pursuing the Ph.D. degree with Curtin University and Main Roads Western Australia.

His research interests include travel time reliability modeling, spatial data analysis, and geographical information science.

Lalinda Karunaratne received the B.S. degree in civil engineering from the University of Moratuwa, Sri Lanka, in 2005.

He is currently a Road Asset Modeling Specialist at Main Roads Western Australia. He is also a Chartered Asset Management Engineer with experience in entire strategic asset management cycle, including asset component level inventory/condition data collection and analysis, deterioration modeling, works program development, valuation, asset management plan preparation, and updating, gained through working within service delivery, consulting, and government (asset owning) organizations. He received the Graduate Certificates in infrastructure asset management and in pavement engineering from the University of Tasmania, Australia, in 2018.