

PREDICTION MODELLING APPROACH FOR CRACK PROGRESSION OF HEAVY DUTY FLEXIBLE PAVEMENTS

NAHLA H. AL ASWADKO

Dept. of Civil Engineering, University of Duhok, Kurdistan Region-Iraq

ABSTRACT

Pavement management at a network level requires reliable accurate performance prediction models to help road agencies make useful complex decisions about highways maintenance and rehabilitating activities. The purpose of this paper is to report the approach adopted for model development and validation for heavy duty flexible pavements representing by seven rural freeways segments. Hierarchical generalized linear modelling approach has been applied to predict multilevel model to capture the effect of variations among time series data, among road sections and among highways with same duty pavements. The estimation of pavement cracking progression has been based on longitudinal dataset contain cracking data (reported as a percent of the affected area) as dependent variable and cumulative traffic loading, pavement strength and environmental conditions as independent variables.

The study illustrates how panel data can be nested to predict the probability of crack progression to capture the effect of significant unobserved heterogeneity. The significance of relevant contributing factors in predicting crack progression were presented and elucidated. The validation results indicate that the model replicates the pavement behavior well, and that the inclusion of additional factors in addition to time is improving the model prediction.

KEYWORDS: Cracking model, Heavy duty pavement, Prediction model, Crack progression, Flexible pavement

1. INTRODUCTION

Cracking is one of many measurable distress modes that can be used to assess pavement condition. Predicting surface performance of spray sealed pavement is important for both pavement design and management due to the fact that cracking is one of the primary distress modes in pavement surface performance and its prediction is a major concern for pavement engineers. In addition, the extended cracking in pavement surface layers frequently speeds up pavement deterioration because it allows water ingress and weakens the pavement and subgrade layers (Paterson, 1987) by increasing the moisture content. Although road roughness has been the most suitable measure for evaluating long-term functional performance of road pavements, it may not always be an effective measure for evaluating pavement maintenance and rehabilitation requirements. There are other more appropriate pavement distress measures such as cracking and rutting for triggering intervention (Toole et al., 2009). Further, monitoring of cracking data is needed to indicate pavement deterioration where rutting data is not sufficiently

extensive at network level (Moffatt and Hassan, 2006).

Cracks often occur for two main reasons, namely traffic loading and environmental factors. The first cause is due to repeated loading or overstressing by traffic, while the second cause is due to moisture changes, expansion potential of subgrade soils, oxidation or chemical shrinkage of the pavement and/or surfacing materials (Moffatt and Hassan, 2006).

Modelling cracking is a relatively complex process, making a probabilistic modelling approach more suitable for such complex phenomena (Yang, 2004). By using a probabilistic approach, the model can quantify the probabilities of pavement failure, while considering the full life of pavement surface with an associated probability for cracking in every year and identifying all factors that significantly affect crack behavior (Henning, 2008). In this study, hierarchical generalized linear model was developed to predict the probability of severity levels for surface cracking. The model has been developed to take into account the effect of variations among observations, among sections and among highways. Readily available historical time series data from seven highway segments have been

collected and prepared for modelling. These time series include surface cracking as a performance parameter and traffic loading, expansion potential of subgrade soil, climate condition, condition of drainage system and pavement strength as predictor parameters. Cracking data includes all types of cracking: transverse, longitudinal and crocodile cracking and is reported as a percent of the affected area. The study predicts the probability of a pavement maintaining its current level of cracking.

2. NETWORK SELECTION

A representative network was selected from rural highway network of the State of Victoria/Australia. The sample network includes heavy duty pavement sections from seven highways of class M roads with a total length around 170 km (170 of 100m-sections). This class refers to roads that have a high standard of driving conditions, including four traffic lanes, sealed shoulders, divided carriageways and visible line marking. Road of this class connects Melbourne (the capital of Victoria) with other capital cities and major provincial centers (VicRoads, 2016). In Victoria, the rural network is essentially spray sealed surface over natural gravels. Road agencies practice for almost all rural roads is to prime or prime seal, then a single or double coat seal of bitumen with one sized aggregate (size 10 or 14 mm).

The selected sample has a reasonable coverage of network characteristics and conditions. It covers wide ranges of all major parameters that contribute to pavement cracking progression including traffic loading, pavement strength, climate and drainage condition. The pavements of all these highway sections have granular bases and sub-bases with single or double coat spray/chip seal.

3. DESCRIPTION OF STUDY VARIABLES

Based on the availability of information for the sample network, the variables considered in this study include surface cracking as the dependent variable (DV) and the following parameters as the

independent or predictor variables: traffic loading, pavement strength, soil type, climate, and drainage condition. Provided in the following sections are brief descriptions of the study variables:

1. Cracking

Cracking data were collected in Victoria by using a manual crack recording method and relying on the following requirements (Moffatt and Hassan, 2006):

- Using a visual system for rating surface cracking of a lane from an interpretation of digital video images of a moving vehicle in dry and daylight conditions.
- Surveying a minimum of one lane for each carriageway in a preferred direction (generally the outer lane).
- Measuring the full lane width of the section, between the centers of lane lines.
- Using 100m intervals in one lane for longitudinal sampling frequency.
- Including all types of cracking (visible and repaired): transverse, longitudinal and crocodile cracking types and reporting as a percentage of cracking.
- Surveying at a frequency equal to once every two years.

2. Traffic Loading

Traffic volume data in terms of number of Heavy Vehicles (HV) for different road classes was extracted from the relevant database for 2002, 2009, 2010 and 2011. Estimates of traffic data for missing years were obtained for each highway by using the average growth factor for all its segments. HV numbers at the time of construction (HV_{con}) for different sections along each highway were estimated using current HV (HV_{cur}), section age and relevant growth rate. This data was then used to determine cumulative traffic loading in terms of million equivalent standard axles (MESA) (Jameson, 2012) in conjunction with relevant parameters from VicRoads' code of practice (VicRoads, 2013). The number of heavy vehicles at time of construction and the cumulative growth factor (in each year for which condition data was available) was calculated then cumulative traffic loading data (MESA) was determined using the following equation (Jameson, 2012):

$$MESA = (HV_{con} * CGF * DF * LDF * NHVAG * (ESA/HVAG) * 365) / 10^6 \dots\dots\dots (1)$$

Where:

MESA = million equivalent standard axles, cumulative ESA (equivalent standard axle loads) from construction time to any year condition data.

HV_{con} = number of heavy vehicles at time of construction.

$$= HV_{cur} / [(1 + GF)^{Age \text{ at current HV year}}]$$

HV_{cur} = number of heavy vehicles in any year of actual traffic is available.

GF = average annual growth rate of heavy vehicles.

Age = pavement age.

CGF = Cumulative Growth Factor.

$$= (((1 + 0.01 * GF)^{Age} - 1) / (0.01 * GF))$$

DF = direction factor = 1; assuming all HVs (100%) travel in the direction of design lane (VicRoads, 2013).

LDF = lane distribution factor = 1; assuming 100% of HVs travel along the design lane (outer lane). According to VicRoads' code of practice document (VicRoads, 2013); the LDF value is considered as 1 when the number of road lanes is less than 3 in one direction (Class M roads have two lanes in each direction).

$$SNC_0 = 0.55 * \log_{10} (MESA_{DL} / 120 * 10^6) + 0.6$$

$$SNC_i = SNC_0 * (2 - \exp(0.33 * Age / DL))$$

$$\dots \dots \dots (2)$$

$$\dots \dots \dots (3)$$

Where:

$MESA_{DL}$ = cumulative traffic loading that was expected to be experienced over the nominal pavement design life and calculated based on cumulative growth factor over the design life (DL), where DL = 30 years for class M roads (VicRoads, 2013).

All other terms are as previously defined.

4. Climate Condition

Thornthwaite Moisture Index (TMI) deals with engineering applications that lie on or beneath the ground surface, such as road pavements (Byrne and Aguiar, 2010). It is defined as the combination of annual effects of precipitation, moisture deficit, evapotranspiration, soil water storage and runoff (Thornthwaite, 1948). Historical climate time series data in terms of TMI was extracted from the climate extraction tool developed by Byrne and Aguiar (2010). It is provided as an Excel database which uses latitude and longitude values to access relevant data over time for each 100m road section. TMI values were

NHVAG = average number of axle groups per heavy vehicle = 3.1 (VicRoads, 2013).

ESA/HVAG = average ESA per heavy vehicle axle group = 0.82 (VicRoads, 2013).

3. Pavement Strength

The pavement consists of different layers of materials that typically have different properties and behavior under load. Insufficient pavement strength or deformation and displacement in the upper layers of pavement can contribute to pavement deterioration (Paterson, 1987). As pavement deflection data were not available for the whole network, the following two equations were used to estimate the structural number (SN). The SN is an index providing an indication of the strength of the pavement layers and of the total pavement structure. The initial value of the structural number (SNC_0) at the time of pavement construction (Age = 0) (Chen and Martin, 2012) and modified structural number (SNC_i) at any time (*i*) during the life of pavement (Martin, 2008), have been estimated as follows:

extracted along all highway sections for all relevant years. Generally, a positive sign of TMI refers to a wet area while a negative sign of TMI refers to a dry area.

5. Subgrade Soil Type

Subgrade soil provides support to the upper layers of road pavement and withstands the stresses applied to it under load. Roads are constructed on different types of soils and seasonal moisture variations affect their strength and/or volume differently. Though, all these roads are expected to deteriorate over time. Yet, if roads are constructed on expansive subgrade soils, they can deteriorate at a faster rate than those with stable subgrade (Mann, 2003). Expansive soils (i.e. reactive soils which are sensitive to moisture changes during seasonal variation cycles) cover a large area of the State of Victoria which coincides with a significant portion of the State's rural highway network. More than half of Victorian road pavements are built on expansive subgrade soils with varying levels of expansion potential.

The integrated color coded map of expansive soil regions in Victoria (Mann, 2003) was used in conjunction with AutoCAD software to establish the type of subgrade soils for all selected sites using their start and end chainages as a reference. The different colors in the map represent different soil types with different reactivity levels. For this study, two levels of expansion potential of subgrade soils were identified; namely: moderate to highly expansive soils and the non-expansive soils.

6. Drainage

Drainage has been identified as an important factor for both the functional and structural performance of road pavement (Pearson, 2012). The condition of drainage system for the selected sections was extracted from relevant database between 2004 and 2011; rated as good or poor.

4. STUDY APPROACH

Although a deterministic approach is still widely used in developing cracking models and can provide good prediction results (Martin et al., 2011), this approach cannot effectively take into account the stochastic nature of pavement performance (Toole et al., 2009). In order to capture the uncertainty and variability associated with the pavement deterioration process, there is recently an increased interest in developing probabilistic rather than deterministic models (Kadar et al., 2015). Pavement cracking is characterized as a random process due to stochastic variation in its mechanism and the non-linear properties of pavement surface layers.

When using historical time series data for many pavement sections (panel data) in estimating future condition, examining the cause of heterogeneity across sections data is an essential element (Greene, 2004). This heterogeneity may be due to differences in construction quality, subgrade soil type, climate condition and maintenance activities applied in different ways at different times. Therefore, the data structure is hierarchical with three levels of variation within any road class. Time series observations (level-1) are nested within sections (level-2) which are nested within highways (level-3). Therefore, pavement performance variability should be included in model parameters and it is likely that the variability (i.e. heterogeneity) may be due to observed variables or unobserved variables

(factors beyond those included in the proposed model) in the network (Hong, 2007).

However, as recommended by statistical studies (Raundebush and Bryk, 2002), when the available data have a multi-level hierarchical structure, it is necessary to capture the effect of variance at higher levels to build up a more robust model. The overall aim of the present study is to apply a multilevel modelling approach to predict the probability of pavement crack progression and to capture the effects of variances at high levels through logistic models. Also, the paper aims to study the effects of several factors on pavement crack progression.

5. STUDY METHODOLOGY

One of the most useful nested data analysis techniques is hierarchical linear modelling (HLM). It is a statistical modelling approach that captures the effects of variation at multiple levels (Raudenbush and Bryk, 2002, Field, 2009). HLM explicitly models the dependency between observation data, producing more stable intercept and slope estimates with unbiased standard errors. These models are able to accommodate unbalanced data (Field, 2009).

Cracking data are mostly reported as either predominant cracking type, cracking severity or extent of cracking (percentage of affected area). However, the cracking data used for this study was reported only in terms of extent and was considered as continuous data. However, Wang (2013) recommended that converting cracking data from a continuous variable into a discrete categorical variable would help smooth out abnormality in the dataset.

Accordingly, this continuous cracking extent data was divided into four discrete categories where used to predict crack progression in terms of the probability of pavement falling into each category. The four categories of crack extent used were insignificant, limited, considerable and significant affected area, with the ranges shown in Table (1) (Moffatt and Hassan, 2006).

Ordinal logistic regression models are employed when there are more than two categories of the dependent variable and there is a natural order between these categories. In the context of pavement cracking extent, the progress of the affected area has an order structure starting from an insignificant affected area, progressing to limited affected area and then considerable and

finally a significant affected area. This means that the four categories of cracking data have an ordinal structure relating to cracking progress over time. Hence, the ordinal logistic model was used to predict the cumulative probability of achieving

each category. In terms of multilevel modelling, the logistic regression is extended to include multiple levels of nesting and is known as hierarchical generalized linear modelling (HGLM) (Raudenbush et al., 2011).

Table (1): Classification of the ranges of affected area for crack categories

Category description	Range of affected area	Cracking category
Insignificant affected area	0% to 1%	Insignificant
Limited affected area	1% to < 5%	Limited
Considerable affected area	5% to < 15%	Considerable
Significant affected area	≥ 15%	Significant

6. HIERARCHICAL GENERALIZED LINEAR MODEL SPECIFICATION

In traditional logistic regression, the logit of the odds serves as the dependent variable (logit link function). This logit in binary regression

model is the natural logarithm of the odds that an event occurs (Raudenbush et al., 2011). In this study, the event is the crack occurred or not. The general form of the logit (η) is calculated using the following formula:

$$\begin{aligned} \eta &= \text{LN} [\text{odds (cracked)}] \\ \eta &= \text{LN} [\text{Probability (cracked)} / \text{Probability (uncracked)}] \\ \eta &= \text{LN} [P_1 / (1 - P_1)] \end{aligned} \dots\dots\dots(4)$$

Where:

LN: is the natural logarithm

P_1 : is the probability of cracked

$(1 - P_1)$: is the probability of uncracked

The predicted log-odds can be used to derive the probability of crack pavement (P_1) by computing:

$$P_1 = 1 / [1 + \text{Exp} (-\eta)] \dots\dots\dots(5)$$

The log-odds (η) and P_1 can be estimated via a linear combination of predictor(s) X_1, X_2, \dots, X_n :

$$\eta = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots\dots\dots + \beta_n * X_n \dots\dots\dots(6)$$

$$P_1 = 1 / [1 + \text{Exp} - (\beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots\dots\dots + \beta_n * X_n)] \dots\dots\dots(7)$$

Where:

X_1, X_2, \dots, X_n : are the independent variables.

$\beta_0, \beta_1, \beta_2, \dots, \beta_n$: are fixed and unknown coefficients, where β_0 is the intercept and β_1, β_2 and β_n are the slopes.

As mentioned in section 5 in this paper, it is expected that there are three levels of random variation (heterogeneity) within the existing panel dataset, including Level-1 (e) variation among time series observations within same sections, Level-2 (r_0) variation among pavement sections within same highways, and Level-3 (u_{00}) variation among highways within same road classes.

Theoretically, the effect of heterogeneity can be captured by implementing randomness over the model parameter(s) (Raudenbush and Bryk, 2002, Field, 2009). Hence, the random intercept approach is used by allowing the intercepts to vary at level-2 (β_0) and level-3 (β_{00}). The multilevel model for the binary logistic regression depending on one predictor can be presented as follow:

Level-1: $\eta = \beta_0 + \beta_1 * X_1 + e$

Level-2: $\beta_0 = \beta_{00} + r_0$

Level-3: $\beta_{00} = \beta_{000} + u_{00}$

The final mixed model is:

$$\eta = \beta_{000} + \beta_{10} * X_1 + e + r_0 + u_{00} \dots\dots\dots(8)$$

Where:

η , X_1 , β_0 , and β_1 : are as defined previously.
 e , r_0 , and u_{00} : are the random variables for level-1, level-2 and level-3, respectively.

β_{00} and β_{000} : are the fixed intercept coefficients for level-2 and level-3, respectively.

By incorporating other network variables considered in this study, the above multilevel

models can be extended by including variables that vary over time within section (Time, MESA, SNC_i and TMI) at level-1, and variables that vary from one section to another (SSR and DRA) at level-2. The extended model for crack progression model would include the following variables:

Level-1: $\eta = \beta_0 + \beta_1 * \text{Time} + \beta_2 * \text{MESA} + \beta_3 * \text{TMI} + \beta_4 * \text{SNC}_i + e$

Level-2: $\beta_0 = \beta_{00} + \beta_{01} * \text{SSR} + \beta_{02} * \text{DRA} + r_0$

Level-3: $\beta_{00} = \beta_{000} + u_{00}$

The mixed model that incorporates all the above levels is referred to as the conditional model below:

$$\eta = \beta_{000} + \beta_1 * \text{Time} + \beta_2 * \text{MESA} + \beta_3 * \text{TMI} + \beta_4 * \text{SNC}_i + \beta_{01} * \text{SSR} + \beta_{02} * \text{DRA} + e + r_0 + u_{00} \quad \dots\dots\dots(9)$$

Where:

η : is the predicted logit odds.

Time: is the time variable in years.

MESA: is the traffic loading variable in terms of Million Equivalent Standard Axles load /lane.

TMI: is the climate condition variable in terms of Thornthwaite Moisture Index.

SNC_i is the pavement strength at time (i) in terms of modified structural number.

SSR: is the subgrade soil reactivity variable (coded as, non-expansive = 0 and expansive = 1).

DRA: is the drainage condition variable (coded as, good = 0 and poor = 1).

All other variables are as defined previously.

In ordinal logistic regression, multiple logit functions are utilized to yield the predicted cumulative probability (CP) of each cracking category. As clarified earlier in this study, four discrete categories can be used to predict crack progression in terms of the probability (P) of pavement falling into each category, as presented below:

$$\left. \begin{aligned} \text{Cumulative probability of significant cracking (CP}_{\text{sig}}) &= P_{\text{sig}} \\ \text{Cumulative probability of considerable cracking (CP}_{\text{con}}) &= P_{\text{sig}} + P_{\text{con}} \\ \text{Cumulative probability of limited cracking (CP}_{\text{lim}}) &= P_{\text{sig}} + P_{\text{con}} + P_{\text{lim}} \\ \text{Cumulative probability of insignificant cracking (CP}_{\text{ins}}) &= P_{\text{sig}} + P_{\text{con}} + P_{\text{lim}} + P_{\text{ins}} = 1 \end{aligned} \right\} \dots\dots\dots(10)$$

From the above formulas, the probability of each cracking category can be obtained as follows:

$$\left. \begin{aligned} \text{Probability of significant cracking (P}_{\text{sig}}) &= (\text{CP}_{\text{sig}}) \\ \text{Probability of considerable cracking (P}_{\text{con}}) &= (\text{CP}_{\text{con}}) - (\text{CP}_{\text{sig}}) \\ \text{Probability of limited cracking (P}_{\text{lim}}) &= (\text{CP}_{\text{lim}}) - (\text{CP}_{\text{con}}) \\ \text{Probability of insignificant cracking (P}_{\text{ins}}) &= 1 - (\text{CP}_{\text{lim}}) \end{aligned} \right\} \dots\dots\dots(11)$$

The set of cumulative predicted probability (equation 10) has one redundant probability (CP_{ins}) due to the constraint that the sum of probabilities equal one, so only three equations are needed in this case. Hence, the log-odds can be

predicted via the linear combination of predictors for the first three categories which are separated by the threshold value (δ) as the following formulas:

$$\left. \begin{aligned} \eta (\text{CP}_{\text{sig}}) &= \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots\dots\dots + \beta_n * X_n \\ \eta (\text{CP}_{\text{con}}) &= \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots\dots\dots + \beta_n * X_n + \delta_1 \\ \eta (\text{CP}_{\text{lim}}) &= \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots\dots\dots + \beta_n * X_n + \delta_2 \end{aligned} \right\} \dots\dots\dots(12)$$

Where:

- β_0 : is the first threshold value between significant and considerable categories.
 - $(\beta_0 + \delta_1)$: is the second threshold value between considerable and limited categories.
 - $(\beta_0 + \delta_2)$: is the third threshold value between limited and insignificant categories.
- All other variables are as defined previously.

7. DATA ANALYSIS USING MULTILEVEL MODELLING APPROACH

Multilevel hierarchical generalized linear model (HGLM) was used to develop pavement crack progression (CR) probability models with the Logit Link Function. The study analysis was performed using Hierarchical Linear and Nonlinear Modelling software (HLM7, 2016) and Statistical Package for Social Sciences software (SPSS, 2016). This type of analysis was selected to ensure that the hierarchical structure in the dataset was not ignored; otherwise, very different conclusions would be reached (Field, 2009).

Data splitting is a simple technique for validating a developed model, ensuring prediction accuracy. Good and Hardin (2003) recommended that one-fourth to one-third of the data should be set aside for validation purposes. Before developing the models, a random data split was used to divide the dataset into two parts; approximately two-third of the data was used for model development and the remaining one-third of the data was used for model validation. The statistics of continuous independent variables that were used for model development are presented in Table (2).

Table (2): Statistics of continuous independent variables used for progression models development

Statistics	CR	TMI	SNC _i	MESA
Mean	3.79	12	3.28	6.50
Standard Deviation	7.33	28.06	0.18	4.14
Minimum	0	-13	2.73	0.93
Maximum	91	94	3.96	31.28

8. PREDICTING PROBABILITY MODEL OF CRACK PROGRESSION (CR)

A three-level model was used to predict the probability of crack progression in terms of the probability of a pavement falling into each of four discrete categories. The four categories are presented in Table (1) for different ranges of percent cracking area and described as insignificant, limited, considerable and significant area affected.

The null model indicates that the variance components for r_0 and u_{00} are highly significant ($p < 0.001$). The proportion of variance over time is

$$CP_{sig} = 1 / (1 + \text{Exp}(-\eta)) = 1 / (1 + \text{Exp}(-(-3.18))) = 0.0399 \quad \dots\dots\dots(13)$$

$$CP_{con} = 1 / (1 + \text{Exp}(-(\eta + \delta_1))) = 1 / (1 + \text{Exp}(-(-3.18 + 2.83))) = 0.4134 \quad \dots\dots\dots(14)$$

$$CP_{lim} = 1 / (1 + \text{Exp}(-(\eta + \delta_2))) = 1 / (1 + \text{Exp}(-(-3.18 + 3.93))) = 0.6792 \quad \dots\dots\dots(15)$$

Transforming the above predicted cumulative probabilities to the probabilities of cracking for each category using the formulas in equation (11); it is found that on average only 4% of the observations within sections in the selected network are expected to exhibit significant

high (36%) within segments and very high (53%) between highways. Around 11% of the variance was found between sections. The significant variance between observations, sections and highways confirms that there is statistical justification for using a multilevel logistic analysis approach to predict the probability of crack progression in order to capture the variance between levels efficiently. The estimated null models for the cumulative probabilities (CP) of cracking are:

cracking during the study period. Around 37% (41% - 4%) and 27% (68% - 41%) of the observations within sections are expected to exhibit considerable and limited affected cracking areas, respectively. However, about 32% (100% - 68%) of the observations within sections in the

selected network are expected to reveal insignificant affected areas (uncracked) during the study period. Available independent variables are added to the null model as predictors to estimate

the conditional predicted cumulative probability (CP) for each cracking category. The final mixed model for the CP of significant, considerable and limited categories are presented below:

$$CP_{sig} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) \dots\dots(16)$$

$$CP_{con} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 3.807)) \dots\dots(17)$$

$$CP_{lim} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 5.380)) \dots\dots(18)$$

The analysis results show that the variances for the random errors (r_0 and u_{00}) and the fixed parameters (Time, MESA, SNCi, SSR and TMI) are statistically significant with p -values less than 0.001. However, drainage condition (DRA) was not significant hence excluded from the model. This means that drainage condition has no significant contribution to pavement cracking in the selected sample. The reason behind that is heavy duty pavement (class M road sections) has high standards of design and construction, well maintained, and generally exhibit high levels of smoothness. Also, road cross sections' crowns are generally high, with deep table drain inverts and sub-soil drains may also be present, and therefore there is a little opportunity for water to gain access to the pavement (Toole et al., 2004)." Time, MESA, TMI and SSR were positively related to the probability of CR, whereas, SNCi was negatively related to the probability of CR. The three models estimated in equations (16), (17) and (18) have the same estimated slopes for all predictors with estimated thresholds of (-0.072), (-0.072+3.807) and (-0.072+5.380), respectively.

The absolute values of t -ratios indicate that the effect of Time is stronger than the effects of other variables on crack progression, followed by TMI, SNCi and SSR. Yet, MESA has a smaller effect on CR. The effect of odds ratios for the predictors of the CR conditional model can be explained as follows:

1. Time factor: for every additional year in time the odds of greater cracking (significant cracking category) as opposed to less cracking (insignificant cracking category) increase 1.62 times on average when the other factors are statistically controlled.

2. Traffic loading factor: for every additional MESA in traffic loading the odds of greater cracking as opposed to less cracking increase 1.02 times on average when the other factors are statistically controlled.

3. Pavement strength: if all the other factors remain the same, the decrease of pavement strength by one unit of SNCi results in increasing the odds of greater cracking as opposed to less cracking by 90% ($[1-0.099]*100\%$).

4. Soil type factor: after controlling all other factors, the odds of greater cracking as opposed to less cracking is 45% higher (or 1.45 times) for the pavements built on expansive subgrade soil than for pavements built on non-expansive soil.

5. Climate factor: for every additional unit in TMI the odds of greater cracking as opposed to less cracking increase 1.05 times on average when the other factors are statistically controlled.

By using the cumulative probability equations (equations 16 to 18), the probability of each cracking category can be found using formulas in equation (11). These probabilities are presented below:

$$P_{sig} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) \dots\dots(19)$$

$$P_{con} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 3.807 - 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) \dots\dots(20)$$

$$P_{lim} = CP_{lim} = 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 5.380)) - 1 / (1 + \text{Exp} (-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNCi} + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 3.807)) \dots\dots(21)$$

$$P_{ins} = 1 - CP_{lim} = 1 / (1 + \text{Exp}(-(-0.072 + 0.482 * \text{Time} + 0.019 * \text{MESA} - 2.317 * \text{SNC}_i + 0.369 * \text{SSR} + 0.047 * \text{TMI}))) + 5.380) \dots\dots\dots(22)$$

On the basis of above equations, Figure (1) shows the simulation for the change of probability of crack progression for the four categories with time, when MESA and SNC_i are at their mean values with wet climate and expansive soil. The inference from this figure is that a section say at year 6 has a 3% probability of having insignificant affected area, a 10% probability of having limited affected area, a 74% probability of having a

considerable affected area and a 13% probability of having significant affected area. Therefore, due to the highest probability of considerable cracking category, it is the most likely cracking state for that section. However, these observations could be more or less in different conditions when considering various values of MESA and SNC_i with wet or dry climate conditions and expansive or non-expansive soils.

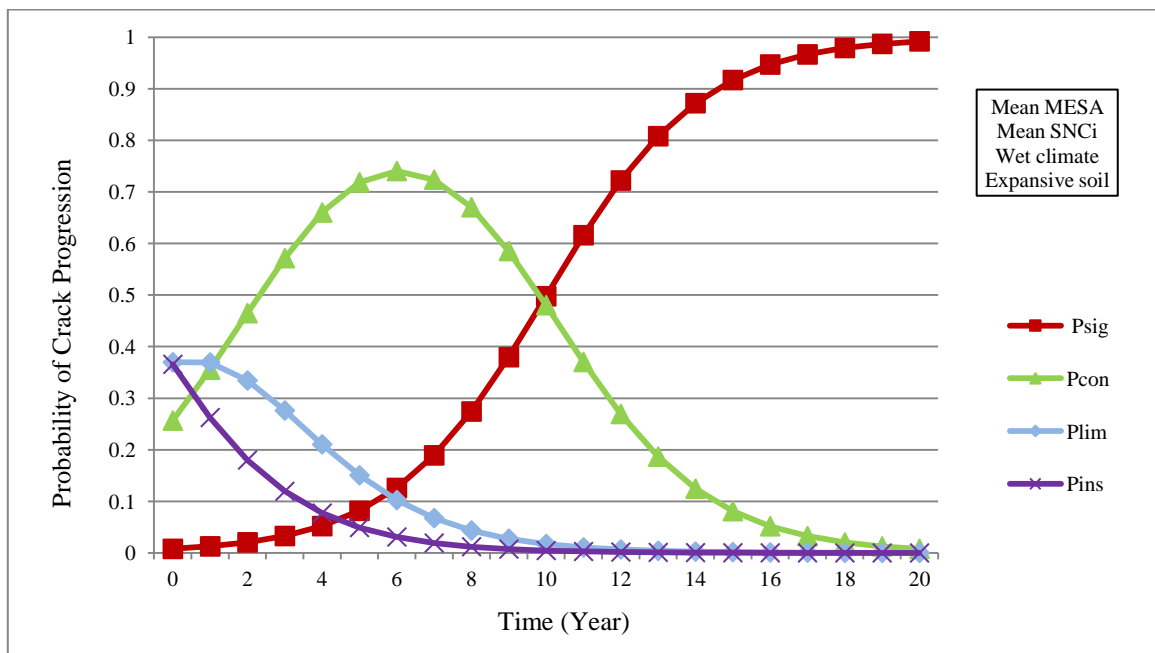


Fig. (1): Simulation for the probabilities of crack progression model over time

9. MODEL EVALUATION

Cross-tabulation analysis is used to test the ability of the developed models to correctly predict crack progression. The developed model is evaluated by testing the success of each of the developed probability model. The analysis result shows the frequency distribution of the predicted and observed cracking data. The numbers of observations that are being correctly predicted in the dataset are used to determine the success rate of the developed model. The results indicate that out of the 142 observations as

significant affected area category, only 105 are correctly assigned to that category using the 50% predicted probability. Out of the 763 observations as considerable affected area category, 417 are correctly assigned. For the observations of limited affected area category, 430 out of 540 are correctly assigned. For the observations as insignificant affected area category, 1155 out of 1499 are correctly assigned. These results indicate that the overall success rate of the crack progression model is 72%, which is calculated as below:

$$\% \text{ Success rate of CR progression} = \frac{105 + 417 + 430 + 1155}{2944} = 72\% \dots\dots\dots(23)$$

The above success rate increases to 86% for developed crack progression model, when the estimated probability is increased to 60%. However, the accuracy of cracking data is subjective in a way and certainty subject to human error and this certainly affects the accuracy of model and results in addition to the fact that some model parameters are estimated and not measured (such as pavement strength).

10. MODEL VALIDATION

Internal validation method is used to ensure that the developed models have the ability to predict the future conditions accurately. As mentioned before in this paper, around one-third of the data is set aside to use for model validation. This dataset is used to develop a validation model with the same variables that have been used for the developed model. Multiple statistical testing using a Bonferroni correction (Field, 2009) is applied when checking whether the coefficients for the validation models fell within the 95% confidence intervals for the coefficients of the developed models. The confidence interval (CI) estimate provides a range of likely values for each of the models parameters. Based on the general form of a confidence interval, the lower and upper bounds of the 95% confidence intervals are calculated using the following formula (Field, 2009):

$$95\% \text{ confidence interval} = \text{estimated parameter} \pm 1.96 * \text{standard error} \dots\dots\dots(24)$$

The internal validation result for the developed crack progression model is presented in Table (3). The result indicates that all parameters of the model based on the validation dataset fell within the upper and lower bound intervals for the parameters of the developed model. This means that the probability model exhibit internal validity.

11. SUMMARY AND CONCLUSIONS

The study presents multilevel hierarchical generalized linear model that can account for the correlation among time series data of the same pavement section and capture the effect of unobserved factors on pavement deterioration.

The study demonstrates the application of multilevel analysis using HLM7 software for modelling the probability of pavement progression. The procedure can be applied to any pavement condition variable that has ordinal classification with data that has a hierarchical structure, such as pavement rutting and skid resistance.

From the analysis approach performed for the sets of road sections of heavy duty pavement, the results indicate that unobserved heterogeneity is a critical aspect that should be considered among sections and among highways for modelling heavy duty cracking pavement. The developed progression model is statistically significant and the parameter estimates are significant and have correct signs. The developed models indicate that time, traffic loading, climate condition and subgrade soil type have positive contributions to crack progression. However, pavement strength has negative contribution to crack progression. Drainage condition has no significant contribution to pavement cracking in the selected sample. The effect of time is stronger than the other variables on crack progression. The effect of climate condition at any time is stronger than the effect of pavement strength and traffic loading in crack progression.

The model can generate the probability of a pavement staying at a certain distress level and the odds ratio, which enables highway agencies not only predict probability of cracking but also assess the confidence of making such predictions. The probabilistic model format for cracking data provides such flexibility in the application of the model when triggers are set according to risk considerations.

Further, the results of model validation indicate that the developed probability model is well estimating the crack condition and has the ability to predict future condition accurately. The model can be used only within the range of input data (independent variables) used in their development. Therefore, it is recommended that the developed models should be used only within the data limits presented in this study and only for spray sealed pavements.

Table (3): Internal validation result for developed crack progression model

Cracking model variables	CDM ¹	p-value (Developed model)	Standard Error	95% CI ³ LB ⁴	95% CI ³ UB ⁵	CVM ²	p-value (Validated model)
Intercept	- 0.071691	0.978	2.523377	-5.018	4.874	-1.690793	0.633
Time	0.481903	<0.001	0.039007	0.405	0.558	0.462373	<0.001
MESA	0.019106	0.04	0.026491	-0.033	0.071	0.055041	0.01
SNC _i	- 2.316598	<0.001	0.714724	-3.717	-0.916	-1.764982	0.05
SSR	0.368997	0.01	0.146766	0.081	0.657	0.267209	0.01
TMI	0.046572	<0.001	0.007362	0.032	0.061	0.046266	<0.001
Threshold (δ_1)	3.806797	<0.001	0.197419	3.420	4.194	3.441873	<0.001
Threshold (δ_2)	5.379809	<0.001	0.298725	4.794	5.965	4.895593	<0.001

1: Coefficient of developed model 2: Coefficient of validated model 3: Confidence interval 4: Lower bound 5: Upper bound

REFERENCES

- BYRNE, M. & AGUIAR, G. (2010). Impact of Climate Change on Road Performance: Updating Climate Information for Australia. *Austrroads Publication No. AP-R358-10*, Austrroads, Sydney, NSW, Australia.
- CHEN, H. & MARTIN, T. (2012). dTIMS Asset Management Tool User Documentation, Project No. 001538- Version 1.1.
- FIELD, A. (2009). *Discovering statistics using SPSS statistics*. Third Edition, London: Sage.
- GOOD, P. I. & HARDIN, J. W. 2003. *Common Errors in Statistics (and How to Avoid Them)*. John Wiley & Sons, Ltd, USA.
- GREENE, W. (2004). *Interpreting Estimated Parameters and Measuring Individual Heterogeneity in Random Coefficient Models*. Department of Economics, Stern School of Business, New York University.
- HENNING, T. F. (2008). *The Development of Pavement Deterioration Models on the State Highway Network of New Zealand*. Ph.D. Thesis, Research Space at Auckland University, New Zealand.
- HLM7 (2016). HLM7 for Windows [Computer software] by Raudenbush, S.W., Bryk, A.S, & Congdon, R., Skokie, IL: Scientific Software International, Inc.
- HONG, F. (2007). *Modeling Heterogeneity in Transportation Infrastructure Deterioration: Application to Pavement*. Ph.D. Thesis, University of Texas at Austin.
- JAMESON, G. (2012). Guide to Pavement Technology: Part 2: Pavement Structural Design. *Austrroads Publication No. AGPT02-12*, Austrroads, Sydney, NSW, Australia.
- KADAR, P., MARTIN, T., BARAN, M. AND SEN, R. (2015). Addressing Uncertainties of Performance Modelling with Stochastic Information Packages-Incorporating Uncertainty in Performance and Budget Forecasts. In *9th International Conference on Managing Pavement Assets*. In: *Proceeding of the 9th international conference on managing pavement assets*, Washington, DC.
- MANN, A. (2003). *The Identification of Road Sections in Victoria Displaying Roughness Caused by Expansive Soils*. MSC Thesis, Swinburne University of Technology, Victoria, Australia.
- MARTIN, T. C. (2008). *Predicting Sealed Granular Pavement Deterioration At A Road Network Level*. Ph.D. Thesis, Monash University, Victoria, Australia.
- MARTIN, T., CHOUMMANIVONG, L. & TOOLE, T. (2011). A New Pavement Deterioration Models for Sealed Low Volume Roads in Australia. *Eighth International Conference on Managing Pavement Assets*, November 15-19, 2011, Santiago, Chile.
- MOFFATT, M. & HASSAN, R. (2006). Guide To Asset Management: Part 5e: Cracking. *Austrroads Publication No. AGAM05E-06*, Austrroads, Sydney, NSW, Australia.

- PATERSON, W. D. (1987). *Road Deterioration And Maintenance Effects: Models For Planning And Management*. Highway Design and Maintenance Standard Series, John Hopkins University Press, Baltimore, USA.
- PEARSON, D. (2012). *Deterioration and maintenance of pavements*. London, SW1P 3AA, United Kingdom.
- RAUDENBUSH, S. W., & BRYK, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*, Second Edition. Newbury Park, CA: Sage.
- RAUDENBUSH, S. W., BRYK, A. S., CHEONG, Y. F., CONGDON, R. & TOIT, M. (2011). *HLM7 Hierarchical Linear and Nonlinear Modeling Manual*. SSI Scientific Software International, Inc, USA.
- SPSS (2016). *IBM SPSS Statistics for Windows*, Version 22.0. Armonk, NY: IBM Corp.
- THORNTON, C. W. (1948). *An Approach Toward a Rational Classification of Climate*. *Geographical Review* (currently published by American Geographical Society), Volume 38, No. 1, pp. 55-94.
- TOOLE, T., MARTIN, T., ROBERTS, J., KADAR, P. & BYRNE, M. (2009). *Guide To Asset Management Part 5H: Performance Modelling*. Austroads Publication No. AGAM05H-09, Austroads, Sydney, NSW.
- TOOLE, T., ROPER, R. & MARTIN, T. 2004. *Improve HDM-4 Model Calibration Factors and Application Guidelines for Sealed Roads in Victoria*. Final Report, Contract Report RC2464- Part 1, ARRB Transport Research.
- VICROADS (2013). *Selection and Design of Pavements and Surfacing*, Code of Practice RC 500.22, VicRoads, Kew, Melbourne, Australia.
- WANG, Y. (2013). *Ordinal logistic regression model for predicting AC overlay cracking*. *Journal of Performance of Constructed Facilities*, Volume 27, Issue 3, pp. 346-353.
- YANG, J. (2004). *Road Crack Condition Performance Modeling Using Recurrent Markov Chains and Artificial Neural Networks*. Ph.D. Thesis, University of South Florida.