J. Infrastructure Plan. and Man., JSCE, No. 744/IV-61, 61-67, 2003 October

[Special Topic]

DEVELOPMENT OF A STOCHASTIC MODEL OF PAVEMENT DISTRESS INITIATION

Hee Cheol SHIN[1] and Samer MADANAT[2]

[1] Post-doctoral Researcher, Dept. of Civil and Environmental Engineering, University of California
Berkeley, CA 94720, U.S.A.
E-mail: hcshin@ce.berkeley.edu

[2] Professor, Dept. of Civil and Environmental Engineering, University of California,
Berkeley, CA 94720, U.S.A.
E-mail: madanat@ce.berkeley.edu

Pavement deterioration models provide useful predictions of distress initiation, for purposes of pavement design and management. A common problem in modeling the initiation is the inappropriate treatment of data censoring. If the censoring is not accounted for properly, the model may suffer from statistical biases. In this paper, an analysis of pavement crack initiation data based on the duration modeling techniques is presented. Duration models enable the stochastic nature of pavement crack initiation to be represented as well as censored data to be incorporated in the statistical estimation of the model parameters. The results show that the model predictions are more accurate than those obtained with the original AASHO model.

Key Words: *crack initiation, censoring, duration model*

1. INTRODUCTION

The planning of maintenance and rehabilitation (M&R) for highway agencies involves the allocation of limited resources to a number of pavements within the highway network. This planning process uses as inputs both measured and predicted highway pavement performance indicators. Therefore, accurate predictions of future pavement performance indicators are necessary for efficient planning of maintenance and rehabilitation activities.

The focus of this paper is the development of a stochastic duration model for crack initiation. Section 2 provides a review of previous research conducted on cracking of asphalt concrete, including the causes of cracking and the modeling approaches that have been used. Section 3 formulates the model. Section 4 presents the specification and results of the estimation of the model parameters using a data set from the American Association of State Highway Officials (AASHO) Road Test [1]. Section 5 concludes this paper.

2. LITERATURE REVIEW

Pavement deterioration models relate indicators of pavement condition to explanatory variables such as traffic loads, age, and environmental factors. The most common indicators of pavement condition are surface distresses such as cracking, rutting, potholes, etc. These surface distresses are caused by load, moisture, temperature, construction defects or a combination of the above.

Cracking performance depends on many factors, including

- The thickness of various pavement layers;
- The quality of the construction materials and practices;
- Environmental considerations, such as temperature and moisture; and
- The axle loads and axle configurations to which the pavement is subjected.

The focus of this research is on the statistical estimation of empirical models that relate field data on crack initiation to explanatory variables representing the pavement structure, traffic loading, and climate. The AASHO Road Test was an accelerated loading test experiment. A crack initiation model was developed as part of the test. The crack initiation model uses traffic repetitions as the dependent variable and pavement thickness and load type as explanatory variables. Though the AASHO cracking model's functional form was relatively arbitrary, the model has been widely accepted. It forms the basis for most current pavement design procedures in the world today. The AASHO Road Test Report 5 proposed the following crack initiation equation.

$$Wc = \frac{A_0(a_1D_1 + a_2D_2 + a_3D_3 + a_4)^{A_1} L_2^{A_3}}{(L_1 + L_2)^{A_2}} \tag{1}$$

where

Wc = number of weighted axle applications sustained by the pavement before appearance of Class 2 Cracking;

D_1, D_2, D_3 = thickness of surfacing, base and subbase respectively, in inches;

L_1 = nominal axle load, in kips;

L_2 = 1 for single axle configuration and 2 for tandem axle configuration;

a_1, a_2, a_3, a_4 = coefficients that were assigned earlier; and

A_1, A_2, A_3, A_4 = regression coefficients.

The AASHO model suffered from severe problems. These are discussed below. First, the analysis did not account for censoring. The data are considered censored when cracking is not actually observed. In the case in which the section had cracked before the first inspection, the observation is left censored, or if it had yet to crack at the last inspection, it is considered right censored. In the AASHO Road Test, there were several sections that had not cracked by the time the experiment ended, and these constitute right-censored data. If censoring is not accounted for correctly in the statistical estimation of model parameters, the estimates can be expected to be biased [2]. Second, the model form was arbitrary. One specific problem is the variable $(L_1 + L_2)$, which consists of the sum of two quantities with different units. Finally, the coefficients, a_1 to a_4, were determined a-priori instead of being estimated simultaneously with the other parameters. The predetermined parameters were used to compute the Structural Number (SN) of the pavement.

The Structural Number is related to the thickness of flexible pavement layers through the use of layer coefficients that represent the resistance of the material being used in each layer of the pavement structure. The following general equation for structural number reflects the relative impact of the layer coefficients (a_i) and thickness (D_i):

$$SN = \sum_{i=1}^{3} a_i D_i \tag{2}$$

The estimated values of the coefficients, a_1, a_2, and a_3, were: 0.33, 0.10, and 0.08 respectively.

The Queiroz-GEIPOT models [3), 4)] have separate regression equations that predict crack initiation and the rate of crack progression. The crack initiation model used the number of equivalent single axles to initiation as the dependent variable and the structural number as the explanatory variable. The equation for the crack initiation model is as follows:

$$\log_{10}N_c = \alpha + \beta \log_{10}SN \tag{3}$$

where

N_c = the number of Equivalent Single Axle Loads (ESAL) needed to initiate cracking;

SN = structural number; and

α, β = regression coefficients.

The World Bank's Highway Design and Maintenance (HDM) models[5)] predict the initiation and progression of various pavement distresses such as cracking, rutting, raveling and roughness. Each distress model includes a number of explanatory variables such as age, traffic, design parameters, environmental factors and other distresses. A probabilistic parametric duration model represented crack initiation, where the dependent variable is the probability distribution of the time to cracking. The basic concepts of probabilistic duration models will be described in section 3 of this paper. The HDM-III crack initiation model used a hazard function, $h(t)$ of the following form:

$$h(t) = \gamma \exp(-\gamma\mu)\, t^{\gamma-1} \tag{4}$$

When $\gamma < 1$, the hazard function is decreasing through time; when $\gamma = 1$, it is a constant; and when $\gamma > 1$, the hazard function is increasing. In the analysis of crack initiation, the parameter μ is replaced by a linear vector function of explanatory variables $\underline{x}$, $\mu = \underline{x}'\underline{\beta}$.

The resulting model for prediction of expected cumulative traffic loading to crack initiation is:

$$TE_{CR2} = \beta_1\, SN^{\beta_2}\, e^{\beta_3\, SY} \tag{5}$$

where

TE_{CR2} = mean cumulative traffic loading at initiation of narrow cracking (in millions of ESAL);

SN = structural number;

$SY = SN^4/(1{,}000\, YE_4)$, where YE_4 is the annual traffic loading (in millions of ESAL/lane/year); and

β_1, β_2, β_3 = regression coefficients.

Several deterioration models reviewed above include separate equations for distress initiation and progression. Most crack initiation models were developed without accounting for censoring, which may introduce bias in the parameter estimates.

Madanat et al [6] applied a structured econometric method for developing deterioration models of pavement crack initiation and progression. A model system consisting of a discrete model for distress initiation and a regression model for distress progression was developed. The estimation sample for the progression model is self-selected, as it contains a disproportionately large fraction of weaker pavements, because they are more likely to have already started cracking (they have lower initiation times). This selectivity bias was corrected by using Heckman's sequential procedure. Madanat and Shin [7] extended this research to account for unobserved heterogeneity in the panel data set, using random-effects specifications in both the discrete and continuous models.

3. MODEL FORMULATION

(1) Stochastic duration models

Let T denote the time to cracking of a pavement in a test experiment. T is a random variable that takes values in $(0, \infty)$. Its continuous distribution is specified by a cumulative function $F(t)$ with a density function $f(t)$. The cumulative distribution function is

$$F(t) = \int_0^t f(s)\,ds = \text{Prob}(T \le t) \qquad (6)$$

The probability that the pavement cracks after t is given by the survival function,

$$S(t) = 1 - F(t) = \text{Prob}(T \ge t) \qquad (7)$$

Since we collect data at specific times, and we know the condition of pavements at these times, the hazard rate is a more useful function than the cumulative density function or the survival function. The probability that a pavement cracks in the next small interval, Δt, given it lasts at least until time t, is given by

$$g(t) = \text{Prob}(t \le T < t + \Delta t \mid T \ge t) \qquad (8)$$

The hazard function, $h(t)$, which is the instantaneous rate of change of $g(t)$, is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}(t \le T < t + \Delta t \mid T \ge t)}{\Delta t} \qquad (9)$$

The hazard function quantifies the instantaneous risk that the pavement sections crack at time t. The cumulative (or integrated) hazard function is expressed as

$$H(t) = \int_0^t h(u)\,du \qquad (10)$$

The density function, the survival function, and the hazard function are all related;

$$h(t) = \frac{f(t)}{S(t)} \qquad (11)$$

The following relationships between these functions also hold:

$$H(t) = -\log S(t) \qquad (12)$$

and

$$S(t) = e^{-H(t)} \qquad (13)$$

Right censoring of the data occurs when the test ends at time C, before all pavement sections have failed. For each section i, we either know T_i, if $T_i \le C$, or that $T_i > C$, in which case the time to cracking for pavement i is censored. The time recorded is min (T_i, C_i) together with the censoring indicator variable δ_i, which is a dummy variable that takes the value 0 if the observation (pavement section) is censored and 1 otherwise.

The full likelihood function is obtained by multiplying the respective contributions of values of density function f for uncensored observations and values of survival function S for censored observations. In the presence of right censoring the likelihood for all observations in a sample of size n is [8] :

$$l = \prod_{\delta_i=1} f(t_i) \prod_{\delta_i=0} S(t_i) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i}\ [S(t_i)]^{1-\delta_i} \qquad (14)$$

To estimate the parameters of the distributions, we maximize the log likelihood function,

$$L = \log l = \sum_{i=1}^{n} \{\delta_i \log[f(t_i)] + (1-\delta_i)\log[S(t_i)]\} \quad (15)$$

Upon choosing a particular distribution, we can substitute the appropriate expressions for the density function and the survival function.

(2) The Weibull hazard model

Though the hazard function *h(t)* could be constant over time, there are many situations in which it is more realistic to suppose that *h(t)* either increases or decreases over time. A flexible form for such a hazard function is given by

$$h(t) = \alpha\gamma t^{\gamma-1} \qquad t > 0 \quad (16)$$

where α and γ are positive constants. The hazard function given by Equation (16) is called the Weibull hazard function with parameters α and γ. The parametric model that follows the Weibull hazard function is called the Weibull model. Note that *h(t)* increases when $\gamma > 1$; decreases when $\gamma < 1$; and is constant when $\gamma = 1$.

The Weibull distribution function obtained from Equation (16) is:

$$F(t) = 1 - \exp\left(-\int_0^t h(s)\, ds\right) \qquad t > 0$$
$$= 1 - \exp(-\alpha t^{\gamma}) \quad (17)$$

Its density function is:

$$f(t) = \alpha\gamma t^{\gamma-1}\exp(-\alpha t^{\gamma}) \qquad t > 0 \quad (18)$$

The survival function is thus:

$$S(t) = 1 - F(t) = \exp(-\alpha t^{\gamma}) \quad (19)$$

If a vector of explanatory variables x is observed with the duration data, the Weibull hazard function is written as:

$$h(t, \underline{x}, \underline{\beta}) = e^{-\gamma\underline{\mu}}\gamma t^{\gamma-1}$$
$$= e^{-\gamma\underline{x}\underline{\beta}}\gamma t^{\gamma-1} \quad (20)$$

where $\underline{\mu} = \underline{x}\ \underline{\beta}$. Then, the distribution function, density function, and survival function are as follows:

$$F(t, \underline{x}, \underline{\beta}) = 1 - \exp\left(-e^{-\gamma\underline{x}\underline{\beta}}t^{\gamma}\right) \quad (21)$$

$$f(t, \underline{x}, \underline{\beta}) = e^{-\gamma\underline{x}\underline{\beta}}\gamma t^{\gamma-1}\exp\left(-e^{-\gamma\underline{x}\underline{\beta}}t^{\gamma}\right) \quad (22)$$

$$S(t) = \exp\left(-e^{-\gamma\underline{x}\underline{\beta}}t^{\gamma}\right) \quad (23)$$

The parameters γ and β of the model can be estimated by maximum likelihood. In the Weibull model, the expected time to crack initiation is given by *(9)*:

$$E[t|\underline{x}] = e^{\underline{\mu}}\Gamma\left(1 + \frac{1}{\gamma}\right)$$
$$= e^{\underline{x}\underline{\beta}}\Gamma\left(1 + \frac{1}{\gamma}\right) \quad (24)$$

where the gamma function, $\Gamma(z)$, is defined as

$$\Gamma(z) = \int_0^{\infty} w^{z-1}e^{-w}dw \quad (25)$$

for $z > 0$.

Because the model is not a linear regression model, there is no obvious equivalent to the conventionally reported standard error.

4. MODEL SPECIFICATION AND ESTIMATION RESULTS

(1) Model specification

The AASHO Road Test Data were used for development of the pavement cracking initiation model. We have a total of 252 observations (test sections). The number of sections that had cracked by the end of the test was 185. The remaining 67 observations were censored (i.e., cracking had not occurred yet by the time of the test). The model shown in this section used right wheel path cracking as the dependent variable. The variables included are:

- *avt*: number of accumulated load repetitions in the traffic lane before crack initiation (the dependent variable);
- D_1: surface thickness in inches (1 to 6 inches);
- D_2: base thickness in inches (0, 3, 6, and 9 inches);
- D_3: sub-base thickness in inches (0, 4, 8, 12, and 16 inches);
- *LOAD*: nominal axle load (in kips); and
- *TYPE*: single dummy variable, 1 for single axle and 0 for tandem axle.

We estimated a Weibull model using two different model specifications. Our first specification (Model 1) was a modified version of the original AASHO model, which was presented earlier (equation 1). The function μ (of the hazard model in equation 20), using this first specification, is shown below:

$$\mu = \beta_1 + \beta_2 D_1 + \beta_3 D_2 + \beta_4 D_3 + \beta_5 L_2 + \beta_6(L_1 + L_2) \quad (26)$$

where L_1 and L_2 are defined after equation (1). Essentially, this is similar to the AASHO specification, with the predetermined coefficients of the structural number in equation (1) replaced by parameters to be estimated.

Our second model specification improved on the original AASHO specification, which is problematic because it assumes that the effects of L_1 and L_2 are separable and additive. This is physically unrealistic for two reasons. First, it assumes a constant rate of substitution between axle load and axle type, which is inconsistent with pavement engineering knowledge. Pavement engineers typically use nonlinear equations to convert between single and tandem axle loads. Second, the specification involves an addition of two variables with different units (L_1 is in units of kips while L_2 is in units of axles).

Our improved specification accounts for the difference in the effects of a single axle and a tandem axle of equal load through the use of interaction terms. The function μ in the hazard model with the improved specification (Model 2) is:

$$\mu = \beta_1+\beta_2 D_1+\beta_3 D_2+\beta_4 D_3+\beta_5 TYPE*LOAD+\beta_6(1-TYPE)*LOAD \quad (27)$$

Note that we use two interactive terms for *LOAD* and *TYPE*: the first interactive term is $\beta_5 TYPE*LOAD$ and the second interactive term is $\beta_6(1\text{-}TYPE)*LOAD$. In this specification, β_5 is the effect of 1 unit of a single axle load, while β_6 is the effect of 1 unit of a tandem axle load.

(2) Estimation results

It is expected that an increase in pavement layer thickness increases the time to cracking of the pavements, and that an increase in axle load, of either type, will decrease the time to pavement crack initiation. The effect of the surface layer should be greater than the effects of the two unbound layers, and the effect of the sub-base should be the smallest. Moreover, the effect of single loads should be greater than that of tandem loads. The dependent variable in the model is hazard rate, not the time to cracking of the pavements, so the parameters should be interpreted accordingly.

Table 1 presents the results of the Weibull model using the first specification. The t-statistics show that each variable is a significant explanatory variable of crack initiation at one percent significance level. Furthermore, it can be seen that the coefficients have the correct signs, which confirm our a priori hypotheses. The ratio of the estimated resistance of the asphalt concrete to that of the base is less than 3, which is lower than what was obtained in the original AASHO model. On the other hand, the ratio of the estimated resistance of the base to that of the sub-base is about 1.5, which is higher than what was obtained in the original AASHO model. The estimated value of the parameter γ is close to 1.0, which seems to indicate a relatively constant hazard rate in terms of load repetitions.

Table 1 Parameter Estimates for Model 1

Variable	Coefficient	t-statistic
Constant	7.22	15.4
D_1	0.706	8.95
D_2	0.255	9.78
D_3	0.172	7.43
L_2	3.33	9.43
L_1+L_2	-0.180	-12.7
γ	1.08	16.4

Table 2 presents the results of the Weibull model using our second specification. Again, the t-statistics show that each variable is a significant explanatory variable of crack initiation at one percent significance level. Again, it can be seen that the coefficients have the expected signs.

The ratios of the estimated resistances of the three layers are close to those obtained in the original AASHO model. The results indicate that the asphalt concrete layer is about 3.1 times more effective in reducing the rate of crack initiation than the base layer, and the base layer is about 1.3 times more effective than the sub-base layer. The AASHO results indicated that the asphalt concrete layer is about 3.3 times more effective in reducing crack initiation than the base layer, and the base layer is about 1.3 times more effective than the sub-base layer [1].

The coefficients of the two interactive load terms indicate that a tandem axle load of 1.85 kip has the same effect on crack initiation as single axle load of 1 kip. These relative magnitudes are consistent with pavement engineering knowledge. Finally, note that the estimated value of the parameter γ is greater than one, indicating an increasing hazard rate with load repetitions.

The Survival function and the cumulative hazard

Table 2 Parameter Estimates for Model 2

Variable	Coefficient	t-statistic
Constant	10.9	60.4
D_1	0.783	11.5
D_2	0.253	12.6
D_3	0.191	10.7
TYPE*LOAD	-0.230	-21.5
(1-TYPE)*LOAD	-0.124	-18.8
γ	1.39	17.5

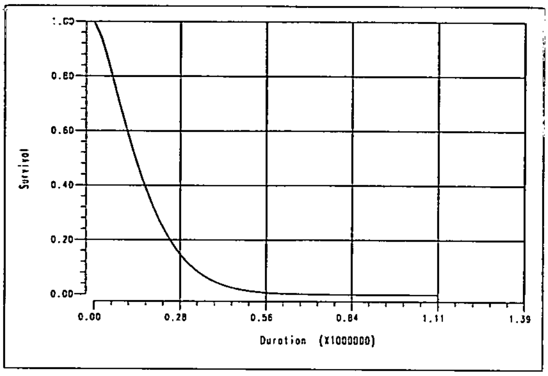

Figure 1 Survival Function

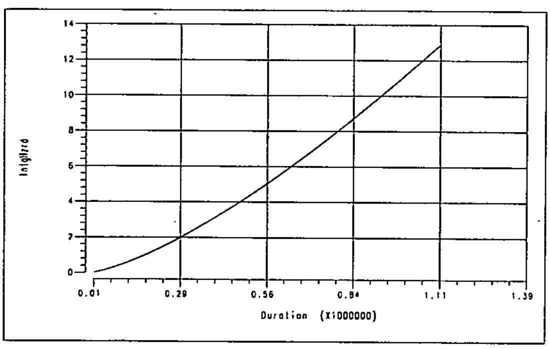

Figure 2 Cumulative Hazard Function

function computed at the means of the explanatory variables are shown in **Figure 1** and **Figure 2**.. The unit of duration at the x-axis is number of accumulated load repetitions.

The prediction accuracies of the three models (AASHO, Model 1 and Model 2), for the sample of observations used in estimating the parameters, were compared by computing the Root Mean Squared Error (RMSE) of the mean axle load repetitions until crack initiation, as predicted with each

Table 3. RMSE of the Mean Predicted Axle Load Repetitions to crack initiation

AASHO	Model 1	Model 2
3.09 E+5	2.57 E+5	1.60 E+5

model. The results are shown in **Table 3**. The entries in the table are the RMSE of the predicted mean time to cracking in units of axle load repetitions. It can be seen that Model 1 achieves about 20% improvement in the prediction accuracy over the original AASHO model, while Model 2 is 50% more accurate than the AASHO model. In other words, for this data set, about 20% improvement in prediction accuracy can be attributed to the contribution of rigorous statistical method (stochastic duration techniques) while about 30% improvement can be attributed to the contribution of an improved model specification.

5. CONCLUSION

In this study, an analysis of the pavement cracking initiation data collected during the AASHO Road Test was conducted. This analysis is based on the use of probabilistic duration modeling techniques. Duration techniques enable the stochastic nature of pavement failure time to be evaluated as well as censored data to be incorporated in the statistical estimation of the model parameters. Due to the nature of pavement crack initiation, the presence of censored data is almost unavoidable and not accounting for such data would produce biased model parameters.

The main advantages that distinguish this stochastic duration model from the original AASHO model are as follows. First, the duration model explicitly recognizes the stochastic variations in the pavement cracking initiation process. Second, the stochastic duration model accounts for the fact that some of the data are censored. . Third, all the parameters of the model were estimated simultaneously. Therefore, this model is statistically more efficient than the AASHO model. Fourth, our specification was more realistic than that used in the AASHO model, in that it did not assume that the effects of axle type and load were separate and additive. Finally, the predictions obtained with our hazard rate model, using our improved specification, were about 50% more accurate than those obtained using the original

AASHO model.

ACKNOWLEDGMENT: The authors would like to acknowledge the assistance of Dr. Marshall Thompson in obtaining the original AASHO road test data.

REFERENCE

1) HRB: The AASHO Road Test – Report 5- Pavement Research, Highway Research Board Special Report 61E, Publication No. 954, National Research Council, 1962.
2) Greene, W. H.: *Econometric Analysis*, Third Edition, Prentice-Hall, 1997.
3) Queiroz, C.A.V.: Performance Prediction Models for Pavement Management in Brazil, Ph.D. dissertation, University of Texas, Austin, TX, 1981.
4) GEIPOT: Research on the Interrelationships between Costs of Highway Construction, Maintenance and Utilization (PCIR), Final Report, 12 volumes, Empresa Brasileira de Planejamento de Transportes (GEIPOT), Ministry of Transport, Brasilia, 1982.
5) Paterson, W.D.O.: *Road Deterioration and Maintenance Effects*, The Highway and Maintenance Standard Series, Johns Hopkins University Press, 1987.
6) Madanat, S., Bulusu, S. and Mahmoud, A.: Estimation of infrastructure distress initiation and progression models, *ASCE Journal of Infrastructure Systems*, Vol. 1, No 3, 1995.
7) Madanat, S. and Shin, H.C.: Development of distress progression models using panel data sets of in-service pavements, *Transportation Research Record* 1643, TRB, National Research Council, 1998.
8) Kalbfleisch, J. D. and Prentice, R. L.: *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, 1980.
9) Meeker, W. Q. and Escobar, L.A.: *Statistical Methods for Reliability Data*, John Wiley and Sons, 1998.

(Received November 13, 2002)