Development of a Distress Index and Rehabilitation Criteria for Continuously Reinforced Concrete Pavements Using Discriminant Analysis

CHIA-PEI CHOU AND B. FRANK MCCULLOUGH

Discriminant analysis is applied to developing the distress index and rehabilitation criteria of the network-level pavement management system for continuously reinforced concrete pavements in Texas. The results are intended to provide the Texas State Department of Highways and Public Transportation with guidelines for evaluating the present pavement conditions and for scheduling rehabilitation. For the discriminant analysis, historical condition survey data were evaluated and separated into two groups, overlaid and nonoverlaid pavements, for which detailed descriptions were given. Each set of data comprised several distress manifestations. A discriminant equation was derived from the analysis that linearly combined the distress manifestations and forced the calculated scores Z″ of the two pavement groups to be as statistically different as possible. The relative magnitudes of Z″ could then be used as a distress index. Several modifications were made in the derivation of the discriminant equation in order to minimize the overlapping area of the distribution of the Z″ scores of the two groups. For the final equation developed, the Z″ scores ranged from −1.8 to 1.0, with zero as the criterion for rehabilitation. Thus, a pavement would be classified as a candidate to be overlaid if its distress index was less than zero. With these results, there was a 92.6 percent confidence that any pavement section would be assigned to the group it actually belonged to. Although some unrealistic assumptions were made in this study, the prediction results obtained were encouraging. It is believed that this approach is a further step in the evaluation of pavement distress condition.

The total expenditures required to maintain and rehabilitate pavements in the United States have been increasing rapidly in the past years primarily because most highway sections that were built in the 1960s and 1970s have gradually reached their terminal serviceability condition. Because of the large amount of money involved, improvements in management and technology for the maintenance and rehabilitation of pavements result in significant savings.

In this paper, the use of distress concepts in the pavement management system (PMS) for rigid pavements is discussed. Special emphasis is focused on the application of discriminant analysis techniques *(1)* to the evaluation of the distress condition of continuously reinforced concrete pavement (CRCP) for the purpose of defining the level of pavement performance and determining criteria for major rehabilitation. This scheme is intended to help the Texas State Department of Highways and Public Transportation (SDHPT) in the management of its highway network.

Department of Civil Engineering, The University of Texas at Austin, Austin, Tex. 78712.

BACKGROUND

Review of existing schemes for maintenance and rehabilitation management revealed that the present serviceability index (PSI) was used nationwide for deciding whether a major rehabilitation or an overlay was necessary. The PSI measure developed by Carey and Irick was used at the AASHO road test *(2)*. They showed that the serviceability of a pavement is largely a function of its roughness.

However, a study has been made of a sample of the different degrees of complexity in the existing network-level maintenance and rehabilitation prioritization methods *(3)*. It was concluded that a scheme that uses only the serviceability index is not applicable to CRCP in Texas. The serviceability history of a pavement with heavy maintenance does not appear to change with time or traffic, whereas the distress condition does (Figures 1 and 2). Each point in the figure represents a surveyed section of CRCP in Texas *(4, 5)*. The number of failures (punchouts and patches) per mile was obtained from the records of the CRCP condition surveys performed in Texas in 1974 and 1978 and described in the following sections. The most likely reason for the consistency over time of the serviceability index is the continuous repair of the highway sections by SDHPT staff. This routine maintenance provided tre-

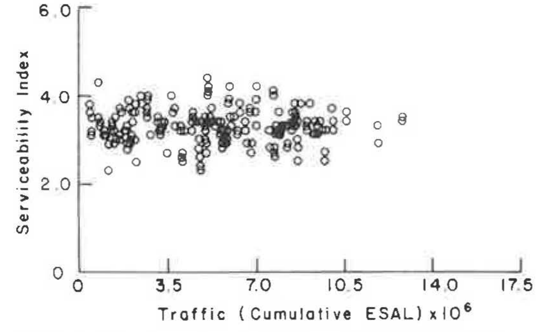

FIGURE 1 Serviceability index versus traffic applications in both directions for Texas CRCP sections surveyed in 1974 and 1978.

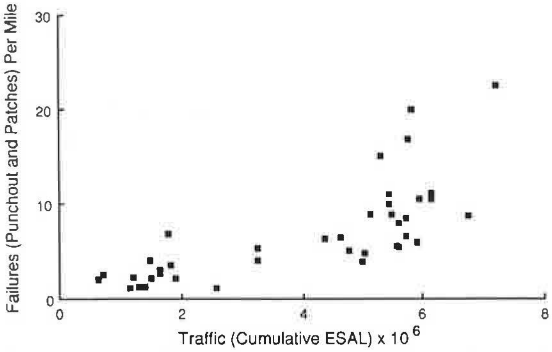

FIGURE 2 Number of failures (punchouts and patches) per mile versus traffic applications in both directions for Texas CRCP sections surveyed in 1974 and 1978.

mendous improvements in pavement roughness, which plays a relatively important role in the serviceability index. Therefore, it is not uncommon for a pavement section to be approaching the end of its life from the structural or economical point of view while the riding quality remains unchanged. Thus, the use of distress measures may be a more realistic way to evaluate a pavement's terminal condition. The development of a distress index to indicate the present pavement condition, therefore, becomes an important task.

In order to obtain a complete data base on the changes in highway distress, a comprehensive condition survey of the CRCP in Texas has been conducted biyearly since 1974. Figure 3 shows the SDHPT districts that were included in the distress condition survey. Several distress manifestations were collected, namely, minor spalling, severe spalling, minor punchout, severe punchout, and repaired patches; the latter were further grouped into asphalt cement, portland cement concrete, and failure patches. Patches result from pavement maintenance. Because the number of patches affects the decision to overlay any particular pavement section, a patch is considered as a distress manifestation in this study.

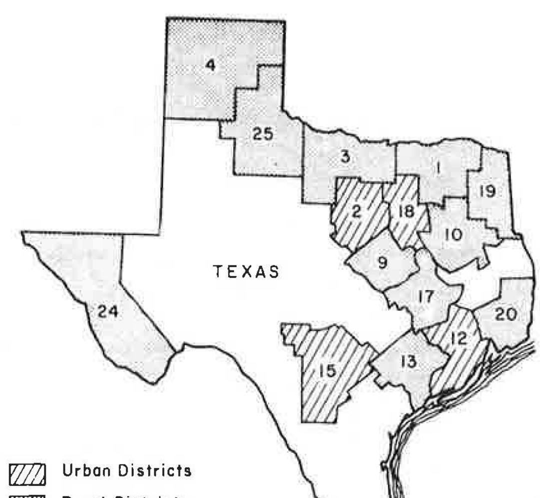

FIGURE 3 Location of rural and urban districts surveyed to collect CRCP information.

The classification of minor or severe is dependent on the size and severity of the distress. For example, a minor punchout is defined as a condition where, although closely spaced transverse cracks are linked by longitudinal cracks to form a block, no sign of movement under traffic is apparent. A severe punchout is recorded when the block moves under traffic. The surrounding cracks will be fairly wide and signs of pumping around the edge of the block may be apparent.

Data were recorded as the cumulative amount of various distress manifestations for every 0.2 mi from the starting point of each surveyed section. Section lengths vary from a fraction of a mile to more than 15 mi. After the condition survey data had been collected and stored, data were reduced for the preparation of the statistical analysis, as follows.

First, the average value of each distress manifestation of every pavement section for each survey year was calculated. This reduced the original survey data of a 10-year period to 1,365 data points. Each datum, representing a pavement section of a certain survey year, was composed of five numbers; each number represented the mean value of a distress manifestation.

It was found that some pavement sections were overlaid in the interval between two successive survey years. No data could be collected after the overlay because the distress could not be seen. Therefore, surveyed sections were grouped into two categories, overlaid and nonoverlaid pavements, and data collected before overlay were used to determine reasons leading to the decision to overlay. Several data points were removed from the database because of the zero values after overlay. Also, survey data of 1984 were removed because no following data were available to check which category the pavement sections of 1984 should be grouped into. Therefore, the final database consisted of 882 data points, 56 overlaid and 826 nonoverlaid. A sample of the summary of reduced data is shown in Figure 4.

Several methods used to calculate the distress index revealed in the literature were evaluated. The proposed methods are

1. Subjective parameters *(6)*,
2. Regression analysis *(7)*,
3. Factor analysis *(8)*, and
4. Discriminant analysis *(9, 10)*.

After the various methods were reviewed, discriminant analysis was selected for the development of the distress index used in this study, because it appeared to be the most appropriate technique for the data available and because of its encouraging results.

DISCRIMINANT ANALYSIS

Discriminant analysis is a statistical technique used to classify data into groups by maximizing the differences between group means. To distinguish between the groups, a collection of discriminating variables that measure characteristics for which the groups are expected to differ is made. The mathematical objective of discriminant analysis is to weight and linearly combine the discriminating variables in some fashion so that the groups are forced to be as statistically distinct as possible.

($\log_e$) minor spall	($\log_e$) severe spall	($\log_e$) minor punchout	($\log_e$) severe punchout	($\log_e$) patch	group
5.2	.5	1.4	.5	.5	2.0
5.2	.5	0	.5	0	2.0
0	0	0	0	0	2.0
5.0	.5	.9	0	.5	2.0
5.0	.5	0	0	.5	2.0
0	0	0	0	0	2.0
5.5	0	0	0	0	2.0
4.4	0	0	0	0	2.0
5.1	0	0	0	0	2.0
4.7	0	0	0	0	2.0
5.2	.7	.3	0	.2	2.0
5.2	.7	.3	.1	0	2.0
0	0	.3	0	0	2.0
5.3	.5	.5	.1	.3	2.0
5.3	.5	.3	0	.1	2.0
0	0	1.1	0	0	2.0
5.6	1.3	2.9	.3	1.0	2.0
5.6	1.3	1.9	0	0	2.0
0	0	2.1	0	0	2.0
5.5	.9	3.2	1.2	.6	2.0
5.6	.9	3.0	.3	0	2.0
0	0	.4	.1	0	2.0
5.2	1.5	1.9	1.0	2.5	1.0
5.1	1.2	1.6	.7	2.2	1.0
5.9	1.4	2.3	0	1.9	1.0
6.2	1.6	4.2	.8	1.7	1.0
6.2	0	3.3	.9	1.5	1.0
4.7	2.2	4.0	0	1.4	1.0
5.2	2.7	4.0	.3	2.3	1.0
5.2	2.5	3.2	.2	1.9	1.0
0	0	1.5	1.1	3.2	1.0
0	0	2.0	1.0	3.3	1.0
0	0	2.6	.9	3.0	1.0
0	0	2.7	.4	3.5	1.0
5.1	3.9	1.4	1.8	3.5	1.0
5.3	4.1	2.4	1.6	4.0	1.0
6.2	4.1	1.5	1.4	2.0	1.0
6.1	3.1	2.4	2.4	3.0	1.0
5.6	1.1	3.1	2.0	2.7	1.0
5.8	2.1	2.7	.2	2.2	1.0
6.2	1.9	2.4	1.9	2.4	1.0
6.2	4.2	2.2	1.9	2.6	1.0
5.6	5.0	1.1	.8	2.2	1.0

FIGURE 4 Sample of data used in discriminant analysis.

In this study, the historical data are separated into two groups, overlaid and nonoverlaid pavements. Each group of data consists of several distress manifestations, as mentioned in the earlier section, which represent the distress condition of a specific section in a specific survey year. By using discriminant analysis, one or more composites or discriminant functions, the distress variables, are derived so that the composites yield boundaries that minimize the overlap in the distribution of the a discriminant scores of the different groups. The discriminant score is the value of the composite function for a particular datum. Ideally, the discriminant scores for cases within a particular group are similar. The maximum number of discriminant functions that can be derived is either one less than the number of groups or equal to the number of discriminant variables, if there are more groups than variables. Therefore, only one function is derived in this study.

The inputs of the discriminant analysis are historical condition survey data, including various distress types and their corresponding groups. The outcomes of the analysis are the discriminant function and its relative magnitudes that can be used as a distress index. In addition, the percentages of analyzed data that were correctly classified into either correspondent group are given. Once the equation is developed, data for any new section can be assigned to one of the predetermined groups by calculating its discriminant score and comparing it with the boundary between the groups.

In the development of the discriminant function, the discriminant subprogram of the statistical package SPSS was used *(9)*.

At this stage, it is important to mention several assumptions inherent to the approach used in this study.

1. The distress variables are considered normally distributed.
2. The SDHPT district's decisions for overlaying the sections were assumed to be correct and consistent.
3. The total cost of overlaying a pavement when it should not be overlaid is equal to the total cost of not overlaying a pavement when it should be.

ANALYSIS AND RESULTS

The discriminant function developed in the analysis to discriminate between groups is of the form

$$Z_i = \sum_{j=1}^{m} a_j z_{ij}; \quad for\ i = 1, \ldots n \tag{1}$$

where

Z_i	=	discriminant score of the *i*th pavement section datum,
a_i	=	weighting coefficient,
z_{ij}	=	standardized value for the *i*th discriminating variable with respect to the *j*th distress manifestation,
n	=	number of data, and
m	=	number of the discriminant variables.

The standardized values z_i are calculated as follows:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_{\bar{x}_j}}; \quad \text{for } i = 1, \ldots n; \quad j = 1, \ldots m \tag{2}$$

where

x_{ij}	=	value of the distress manifestation *j* for the datum *i* being classified,
x_j	=	mean value of the distress manifestation *j* for all *n* data, and
$\sigma_{\bar{x}_j}$	=	standard deviation of $\bar{x}_j$.

As can be seen in Equation 1, the discriminant function is linear, but it may not produce a realistic situation. However, the statistical program requires a linear form and any nonlinear transformation of the discriminating variables should be made before this program is used. Several transformation models have been tested, including multiple linear, second-degree polynomial, and natural logarithm. The logarithm method, which shows the best results of analysis, is also the most commonly used transformation for growth-type data, for example, distress evolution, and for cases in which the mean is proportional to the standard deviation, which is the case in this study.

TABLE 1 CONSTANTS TO BE USED IN EQUATIONS 1 AND 3, WITH MINOR SPALLING AND SEVERE SPALLING

i	Distress Manifestation	a_i	$\bar{x}_i'$	$\sigma_{x'_i}$
1	Minor spalling (MSP)	−0.04248	3.558	2.5075
2	Severe spalling (SSP)	−0.09866	1.4191	1.6301
3	Minor punchout (MPUNT)	0.05373	1.0853	1.0502
4	Severe punchout (SPUNT)	0.47223	0.3015	0.5044
5	Patch (PATCH)	0.72323	0.6313	0.8281

Based on the findings of the analysis, Equation 2 is modified to the following form (although Equation 1 remains unchanged).

$$z_{ij} = \frac{\ln(x_{ij} + 1) - \bar{x}_j'}{\sigma_{\bar{x}_j'}} \quad (3)$$

where

x_{ij} = same as defined before;

$\bar{x}_j' = \sum_{j=1}^{n} \ln(x_{ij} + 1)/n$;

n = number of data, both overlaid and nonoverlaid groups; and

$\sigma_{\bar{x}_j'}$ = standard deviation for $\bar{x}_j'$.

Table 1 summarizes the parameters to be used in Equation 3. The variable PATCH is the sum of asphalt cement, portland cement concrete, and failure patches. Inclusion of the minor and severe spalling terms in the equation would have been misleading because of their counter signs. In addition, the two terms have relatively small values of coefficients compared to the other three variables. Thus, another equation was developed without considering the terms of minor and severe spalling. Table 2 presents the coefficients, mean values, and standard deviations of analyzed discriminant variables used in the improved discriminant equation. This equation can be further simplified by introducing the total means and deviations of the distress variables in Equations 1 and 3 to obtain

$$Z = -1.02544 + 0.01872(\text{MPUNT}) + 1.04429(\text{SPUNT}) + 1.09347(\text{PATCH}) \quad (4)$$

where

Z = discriminant score (or Z score),

MPUNT = ln(minor punches per mile + 1),

SPUNT = ln(severe punchouts per mile + 1), and

PATCH = ln(total patches per mile + 1).

There were 882 data points for the population of historical surveyed data of the CRCP network in Texas included in the derivation of the discriminant equation. This data set comprises 56 overlaid and 826 nonoverlaid sections.

Hence, for any particular pavement, data on each distress manifestation should be substituted into Equation 4 in order to obtain a value of Z, the score for that pavement.

As noted in Equation 4, the score has a minimum value of −1.02544 and it increases with the quantity of various distresses. It was always thought that pavements in good condition should have higher scores than those in poor condition. Thus, the signs of constant terms and coefficients are reversed and Equation 4 is rewritten as Equation 5 with the same variable definitions:

$$Z' = 1.02544 - 0.01872(\text{MPUNT}) - 1.04429(\text{SPUNT}) - 1.09347(\text{PATCH}) \quad (5)$$

The Z' scores for all the pavements in the original (historical) data are calculated and the mean scores of each group are also computed. Table 3 summarizes the mean scores of each group and the probability of correct prediction by the discriminant equation.

It should be emphasized that the individual Z' score will not have the same distribution pattern about each group mean because of the different characteristics. The historical distress record of any specific pavement section always from its best condition, that is, no distress, and increases with time and traffic until the unacceptable condition before overlay. Thus, there exists a high bound, the best condition, in the Z' score distribution of the nonoverlaid group, whereas the Z' score of the overlaid group tends to be distributed normally. A frequency distribution for each of the two groups is plotted (against the Z' score) on one continuous horizontal axis (Figure

TABLE 2 STATISTICAL PARAMETERS OF THE CRCP DATA USED IN EQUATIONS 1 AND 3, WITHOUT MINOR SPALLING OR SEVERE SPALLING

i	Distress Manifestation	a_i	$\bar{x}_i'$	$\sigma_{x'_i}$
1	Minor punchout (MPUNT)	0.01869	1.0853	1.502
2	Severe punchout (SPUNT)	0.44885	0.3015	0.5044
3	Patch (PATCH)	0.72391	0.6313	0.8281

TABLE 3 GROUP MEANS OF DISCRIMINAT Z" SCORE AND NUMBER OF CASES CORRECTLY PREDICTED BY THE DISCRIMINANT EQUATION 5

Group	Mean	Number of Cases	Number of Correct Classifications	Percent
Overlaid	−3.1736	56	56	92.9
Nonoverlaid	0.2151	826	757	91.6
Total	0.0	882	809	91.72

5). The shadowed area indicates the overlap of the two distributions. In the discriminant analysis, the grand mean of the groups will always be zero, which falls between the two group means but is not necessarily the average of these two means. A special case happens only when the groups have an equal amount of data and each has a normal distribution of the Z′ score. Overlapping of the Z′ scores between the two groups is not preventable. However, this area can be reduced by transforming the input data or it can be balanced for both sides by seeking a specific value of the Z′ score that will give 50 percent probability to assigning a pavement having a Z′ score with the overlapping area to the nonoverlaid and the overlaid groups. For the equation formed by using the data for the whole population of Figure 4, this specific Z′ score is calculated to be −1.60.

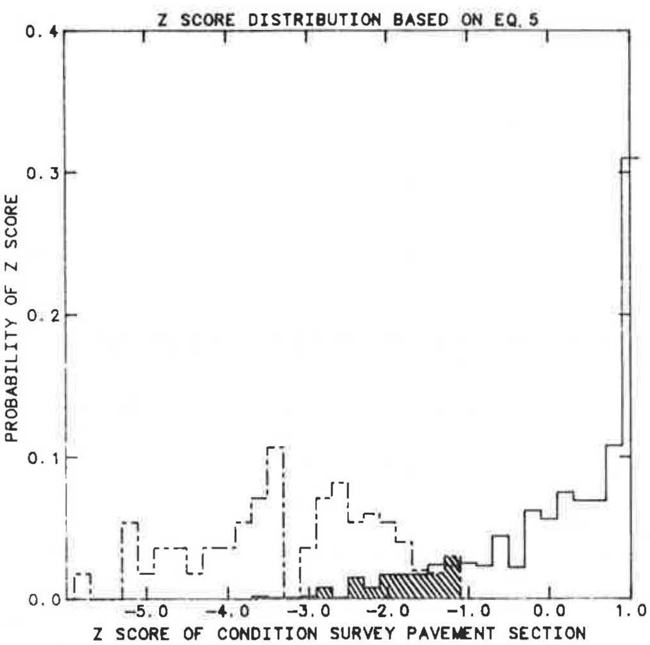

FIGURE 5 Frequency distribution of Z′ scores for the data set used in the discriminant analysis.

Therefore, if Z′ of any pavement is less than −1.60, there is a strong probability that the pavement will be a good candidate for an overlay. Similarly, a pavement with a Z′ value larger than −1.60 has a large probability of being in good condition so that no overlay is necessary. Under the analysis, when Z′ has the value of −1.60, the overlap area is equally divided in two parts. The right half represents the probability α that a pavement that should be overlaid is misclassified into the nonoverlaid group. Likewise, the left half indicates the probability β that a pavement is classified into the overlaid group although it is still better than the acceptable level. In order to simplify decision making about which value the Z′ should be, an assumption was made. It was assumed that the total cost, including agency and user costs, of overlaying a pavement when it should not be overlaid is equivalent to the total cost of not overlaying a pavement when it should be. The Z′ score of −1.60 is, therefore, considered to be the appropriate value to separate the two groups in this study. This decision results in α = β = minimum possible value = 7.4 percent.

Equation 5 can be modified so that the Z scores are compared to zero rather than to −1.60, by using the equation

$$\begin{aligned} Z'' &= Z' + 1.60 \\ &= 2.62544 - 0.01872(\text{MPUNT}) \\ &\quad - 1.04429(\text{SPUNT}) - 1.09347(\text{PATCH}) \end{aligned} \quad (6)$$

or also dividing by 2.62544 so that the equation is of the form

$$\begin{aligned} Z'' = 1.0 &- 0.0071(\text{MPUNT}) \\ &- 0.3978(\text{SPUNT}) - 0.4165(\text{PATCH}) \end{aligned} \quad (7)$$

where Z″ = modified zero score, and MPUNT, SPUNT, and PATCH are as defined in Equation 4.

If $Z'' \geq 0$, pavement does not need overlay; if $Z'' < 0$, pavement needs overlay immediately. A plot of the Z″ frequency distributions of the two groups, based on Equation 7, is represented in Figure 6.

In Equation 7, the most important variable that affects the Z″ score is PATCH (patching), followed by SPUNT (severe punchout), and MPUNT (minor punchout). The required numbers of patches and severe punchouts, respectively, that will cause a pavement section to be overlaid when the other distress variables are all zeros are as follows:

Distress Manifestation	*Number/Lane/Mile*
Severe punchouts	11.3
Patches	10.0

If (minor punchout) = (severe punchout) = 0, then the number of patches/lane/mile to cause overlay of a pavement is exp(1/0.4165) = 10.0.

The number of minor punchouts is not presented herein because a situation in which a pavement section is overlaid due to millions of minor punchouts but no severe punchouts or patches is not realistic.

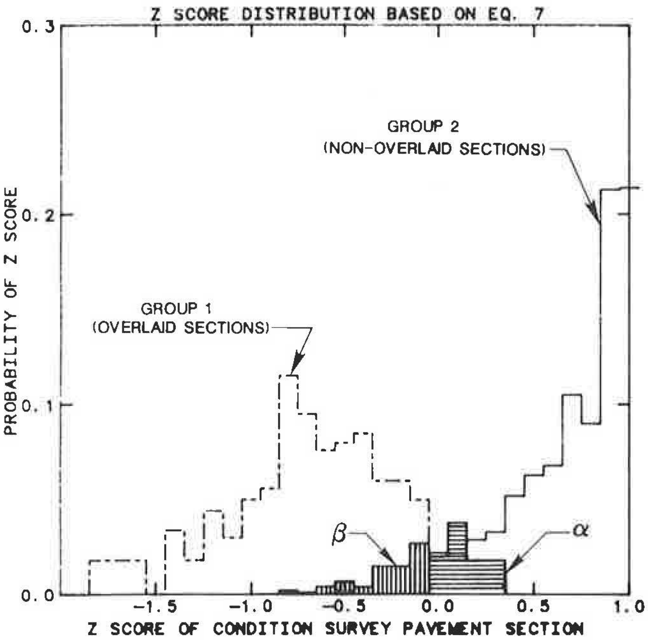

FIGURE 6 Modified frequency distribution of Z' scores for the nonoverlaid and overlaid groups.

The final equation correctly classified 92.6 percent ($1 - \alpha$) of the 882 cases. The cases used to test the prediction capability of the discriminant equations were the same as the ones used to develop the equation.

SUMMARY

Because of the rapid expansion of expenditures in pavement maintenance and rehabilitation, improvements to the PMS have become more and more important. A good evaluation system for pavement performance can not only efficiently use the available annual maintenance budgets at the network level but also economically plan the best rehabilitation timing of any specific pavement at the project level. This study is, therefore, focused on the derivation of distress index and rehabilitation criteria for the CRCP network in Texas. Several approximate methods aimed at developing a distress index have been studied. Discriminant analysis was chosen because its technique is appropriate for the available data. The historical condition survey data of the comprehensive CRCP network were used in the discriminant analysis. The logarithmic method was selected to transform the original distress data before applying the discriminant technique because it resulted in the best fit for the data groupings.

After several modifications, the final equation used to calculate the distress index Z score is

$$Z'' = 1.0 - 0.0071(\text{MPUNT}) - 0.3978(\text{SPUNT}) - 0.4165(\text{PATCH})$$

where

Z'' = distress index (or Z score),
MPUNT = ln(minor punchouts per mile + 1),
SPUNT = ln(severe punchouts per mile + 1), and
PATCH = ln(total patches per mile + 1).

The criterion for major rehabilitation is that a pavement would be classified as a candidate to be overlaid if its distress index Z'' is negative.

CONCLUSIONS

The principal conclusions derived from this study concerning the development of distress index and rehabilitation criteria for CRCP at the network level are summarized as follows.

1. PSI values did not correlate with the rehabilitation decision.
2. A distress index Z'' was developed using 10 years of condition survey data with the discriminant analysis method.
3. Punchouts and patches were the primary distress manifestations included in the distress index equation.
4. Z'' ranged from +1.0 to −1.8 with zero as the rehabilitation criterion with a confidence level of 92.6 percent.
5. The Z'' equation provides a ranking method for rehabilitation needs for network analysis and maintenance programming.
6. It is recommended that the economics analysis of the overlay timing be studied. The specific number of 1.60 in Equation 6 can be revised by changing the ratio of α to β. This ratio is proportional to the value of the cost of overlaying a pavement when it should not be overlaid divided by the cost of not overlaying a pavement when it should be.
7. Although the distress index and rehabilitation criterion developed in this study were based on the condition survey data of Texas, it is strongly believed that the results can be applied to other states. However, the rehabilitation criterion may be shifted from zero (Figure 6) to new points for different states because the SDHPT district's decision to overlay the pavement sections is highly dependent on budget constraints that vary from state to state.

ACKNOWLEDGMENT

The authors are pleased to acknowledge the combined support of the Center for Transportation Research at the University of Texas at Austin and the Texas State Department of Highways and Public Transportation, in cooperation with FHWA. Special acknowledgment is made to the staff of the Center for Transportation Research of the University of Texas at Austin for their assistance in preparing the drafts and figures of this paper.

REFERENCES

1. R. M. Thorndike. *Correlational Procedure for Research.* Garden Press, Inc., New York, 1972.
2. W. N. Carey and P. E. Irick. The Pavement Serviceability-Performance Concept. *Bulletin 250,* HRB, National Research Council, Washington, D.C., 1960.
3. M. de V. Guiterrez and B. F. McCullough. *Rigid Pavement Network Rehabilitation Scheduling Using Distress Quantities.* Research Report 249-5. Center for Transportation Research, The University of Texas at Austin, 1983.

4. M. de V. Guiterrez and B. F. McCullough. *Summary Report for 1978 CRCP Condition Survey in Texas.* Research Report 177-20. Center for Transportation Research, The University of Texas at Austin, 1969.
5. J. P. Machado, B. F. McCullough, and W. R. Hudson. *CRCP: Predictions of Distress Quantities.* Research Report 177-8. Center for Transportation Research, The University of Texas at Austin, 1977 (unpublished).
6. R. D. Pedigo and W. R. Hudson. *Simplified Pavement Management at the Network Level.* Report NA-3/1. ARE, Inc., Austin, Tex., 1981.
7. N. R. Draper and H. Smith. *Applied Regression Analysis.* John Wiley and Sons, New York, 1966.
8. L. T. Oehler and L. F. Holbrook. *Performance of Michigan's Postwar Concrete Pavements.* Research Report 711. Michigan Department of State Highways, Lansing, June 1970.
9. M. M. Tatsuako. *Multivariate Analysis: Techniques for Education and Psychological Research.* John Wiley and Sons, New York, 1971.
10. N. G. Nie et al. *Statistical Package for the Social Sciences (SPSS),* 2nd ed. McGraw-Hill Book Co., New York, 1970.

Publication of this paper sponsored by Committee on Pavement Rehabilitation.