4.0 CLASSICAL REGRESSION ANALYSIS TO PREDICT THE DECREASE IN PSE VALUES

4.1 Multiple Regression Analysis

The major objective of this research was to objectively and quantitatively determine the PSE values of the pavements since the last rehabilitation action. However, the *decrease* in PSE value was taken as the dependent variable because it somewhat represents a "normalized" value. Classical multiple regression analysis was performed to estimate the *decrease* in the PSE (PSE) values. One of the most important aspects of classical regression analysis is the selection of independent variables which are strong indicators of the dependent variable. The selection was done in two steps *(Ott 1993)*:

(i) Enumerating the independent variables, and
(ii) Evaluating and selecting independent variables subjectively or by analyzing correlation.

4.2 Selection of Independent Variables for the Prediction of Decrease in the PSE Values

Extensive literature search was done to select the independent variables to predict the *decrease* in the PSE values. Expert opinion was also sought for this purpose. Since PSE ratings are based on the condition of the base and surface, as indicated by the maintenance costs, subgrade failures, and ability of the section to provide an adequate surface for the prevailing traffic, the following variables were selected to reflect those conditions:

1. Age of the pavement *since the last rehabilitation action* (in years),
2. Cumulative ESAL's that have passed over the pavement *since the last action*,
3. AC layer thickness (in inches),
4. PSE value assigned to the pavement *immediately after the last action*,
5. Decrease in structural number (SN), and
6. Distress level due to transverse cracking.

The selected variables were plotted on scatter plots against the dependent variable, PSE values, and were inspected for possible trends. Also, correlation coefficients for different pairs were determined. It was apparent from the scatter plot that age and SN were not linearly related to PSE values. In the case of age, the rationale is that PSE values do not decrease at the same rate with time. During the initial years this rate is lower, but after a certain period, the PSE values start to decrease drastically. A trial-and-error approach was followed to determine the transformed functional form for an independent variable *(Chowdhury 1998)*. After several trials, the variable age was transformed to $(\text{age})^{1.5}$. For the relationship between the dependent variable, PSE, and the independent variable, age, the Pearson's correlation coefficients improved from 0.35 to 0.68 for the FDBIT and 0.39 to 0.56 for the PDBIT pavements, when the transformation was performed. Similarly, the variable, decrease in structural number, SN, was transformed to exp(SN) to improve the correlation coefficient of the relationship from 0.49 to 0.61 for the FDBIT and 0.48 to 0.55 for the PDBIT pavements, respectively. The variable AC layer thickness was dropped from the PDBIT model as a predictor since the thickness of this type of pavement was not designed to carry the expected traffic. Another important fact to note is that the variables age and cumulative ESALs have a very high correlation between themselves (correlation coefficient of 0.65 for FDBIT and 0.58 for PDBIT). Therefore, only one of them, (age), was included in the model to avoid possible multicolinearity or overspecification of the model *(Chowdhury 1998)*.

Transverse cracking was included in the model as a binary variable. Transverse cracking on the pavements in Kansas is measured by the number of equivalent roadway-width cracks. According to the KDOT PMS rating guide *(KDOT 1996)*, the crack severity is categorized using three severity codes:

Code 1: No roughness, 6 mm (0.25 in.) or wider with no secondary cracking; or any width with secondary cracking less than 1.2 m (4 ft) per lane.

Code 2: Any width crack with noticeable roughness due to depression or bump. Also includes cracks that have greater than 1.2 m (4 ft) of secondary cracking, but no roughness.

Code 3: Any width crack with significant roughness due to depression or bump. Secondary cracking will be more severe than code 2.

Different combinations of the coded cracks will result in different distress levels due to transverse cracking *(KDOT 1996)*. Distress levels due to transverse cracking are defined as shown in Table 4.1.

Table 4.1 Distress Levels Due to Transverse Cracks

DISTRESS LEVELS	**TRANSVERSE CRACK CODES**		
	CODE 1	**CODE 2**	**CODE 3**
DL 1	< 3	0	0
DL 2	3	< 3	< 2
DL 3	ANY NO.	3	2

4.3 Criteria Used to Select a Model

The following criteria were used to select a model:

(i) *Minimize mean sum square errors (MSE)*: The smallest MSE will result in the narrowest confidence intervals and largest test statistics. The model with the smallest MSE involving the least number of independent variables can generally be considered as the best model *(Ott 1993)*.

(ii) *Maximize the Coefficient of Determination (R^2):* R^2 is a measure of how well the estimated model fits the observed data. The best model selected is generally the one with the largest R^2.

(iii) *Minimum increase of R^2 :* The best model is selected as the model associated with the smallest increase in R^2 with the addition of an extra variable.

(iv) *Mallows C_p statistic:* The best model is usually thought to have a C_p value closest to p, where, p is the number of regression coefficients. Models associated with C_p greater than p are usually thought to be biased or misspecified models *(Ott 1993)*.

4.4 Models Obtained and the 'Model Utility' Tests

FDBIT Pavements: Detailed analyses and summary statistics of the model development have been described by Chowdhury *(1998)*. For FDBIT pavements, the selected models are:

<u>Distress Level 1</u>

$$PSE = 0.216 * (AGE)^{1.5} - 20.82 * \exp[\ SN] + 0.138 * TH + 0.328 * PSE + 17.65 * DL1 \qquad (4.1)$$

<u>Distress Level 2</u>

$$PSE = 0.216 * (AGE)^{1.5} - 20.82 * \exp[\ SN] + 0.138 * TH + 0.328 * PSE + 18.06 * DL2 \qquad (4.2)$$

<u>Distress Level 3</u>

$$PSE = 0.216 * (AGE)^{1.5} - 20.82 * \exp[\ SN] + 0.138 * TH + 0.328 * PSE + 18.38 * DL3 \qquad (4.3)$$

where,
PSE= Predicted *decrease* in the PSE value,
AGE= Age of the pavement *since the last rehabilitation action* (in years),
TH = AC layer thickness (in inches),
PSE= PSE value assigned to the pavement *immediately after the last action*,
SN= Decrease in structural number, and
DL_i= Distress level due to transverse cracking (i = 1, 2 and 3).

The p-values for the parameters imply that all the variables are significant at a level of more than 95%. The ANOVA shown in Table 4.2 for the models implies that the model has an F-value of 37 and its significance value is 0.0001. Since the selected model has a high F-value and a very low p-value, it satisfactorily passes the model utility test, which indicates that the model is helpful and adequate in predicting the dependent variable. Also the estimated root mean square error ( ) value for the model is 0.47, which reveals the fact that the selected model will predict the decrease in PSE values at a variability of ±2 or ±0.94 with a confidence of 99%.

It should be noted that the decrease in structural number, SN, values can be computed from the FWD data following the methodology described in Chapter 3 or can be estimated using Equations 3.6 & 3.7 developed previously in Chapter 3.

PDBIT Pavements : For PDBIT pavements, the selected models are:

<u>Distress Level 1</u>

$$PSE = 0.024 * (AGE)^{1.5} - 1.145 * \exp[\ SN] + 0.171 * PSE + 0.229 * DL1 \quad \textit{(4.4)}$$

<u>Distress Level 2</u>

$$PSE = 0.024 * (AGE)^{1.5} - 1.145 * \exp[\ SN] + 0.171 * PSE + 0.958 * DL2 \quad \textit{(4.5)}$$

<u>Distress Level 3</u>

$$PSE = 0.024 * (AGE)^{1.5} - 1.145 * \exp[\ SN] + 0.171 * PSE + 0.2.27 * DL3 \quad \textit{(4.6)}$$

The variables in the above equations have been described before. The p-values for the parameters imply that all the variables are significant at a level of more than 95%. The ANOVA shown in Table 4.3 for the models implies that the model has an F-value of 132 and its significance value is 0.0001. Since the selected model has a high F-value and a very low p-value, it satisfactorily passes the model utility test, which indicates that the model is helpful and adequate in predicting the dependent variable. Also the estimated root mean square error ( ) value for the model is 0.47, which reveals the fact that the selected model will predict the decrease in PSE values at a variability of ±2 or ±0.94 with a confidence of 99%.

Table 4.2 SAS ANOVA Results for the Model Developed for FDBIT Pavements

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Prob > F
Model	7	59.413	8.487	37.011	0.0001
Error	20	4.586	0.229		
Total	27	64.000			

Root MSE: 0.478 *R-square: 0.7835*
Dep. Mean: 1.259 Adj. R-sq: 0.7717
C.V. 38.028

Parameter Estimates

Variable	Deg. of Freedom	Parameter Estimate	Standard Error	T for Ho: Parameter = 0	Prob > {T}
$(AGE)^{1.5}$	1	0.21668	0.239	0.906	0.0105
exp[SN]	1	-20.820	29.999	-0.694	0.0512
THICKNESS	1	0.138	0.049	2.785	0.0114
PSE	1	0.328	0.109	2.989	0.0073
DL1	1	17.655	30.628	0.576	0.0487
DL2	1	18.064	30.636	0.590	0.0197
DL3	1	18.381	30.636	0.600	0.0185

Table 4.3 SAS ANOVA Results for the Model Developed for PDBIT Pavements

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Prob > F
Model	6	138.178	23.029	131.67	0.0001
Error	39	6.821	0.174		
Total	45	145.000			

Root MSE: 0.412 *R-square: 0.8665*
Dep. Mean: 1.444 Adj. R-sq: 0.855
C.V. 28.953

Parameter Estimates

Variable	Deg. of Freedom	Parameter Estimate	Standard Error	T for Ho: Parameter = 0	Prob > {T}
(AGE)1.5	1	0.0246	0.0182	1.352	0.0184
exp[SN]	1	-1.145	0.5559	-2.061	0.0460
PSE	1	0.171	0.0619	2.766	0.0086
DL1	1	0.229	0.4534	0.506	0.0415
DL2	1	0.958	0.4292	2.233	0.0314
DL3	1	2.227	0.4439	5.017	0.0010